



HAL
open science

Capturing Visual Attention With Perturbed Auditory Spatial Cues

Chiara Valzolgher, Mariam Alzaher, Valérie Gaveau, Aurélie Coudert, Mathieu Marx, Eric Truy, Pascal Barone, Alessandro Farnè, Francesco Pavani

► **To cite this version:**

Chiara Valzolgher, Mariam Alzaher, Valérie Gaveau, Aurélie Coudert, Mathieu Marx, et al.. Capturing Visual Attention With Perturbed Auditory Spatial Cues. *Trends in Hearing*, 2023, 27, 10.1177/23312165231182289 . hal-04303851

HAL Id: hal-04303851

<https://hal.science/hal-04303851>

Submitted on 28 Nov 2023

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.

Capturing Visual Attention With Perturbed Auditory Spatial Cues

Trends in Hearing
Volume 27: 1–10
© The Author(s) 2023
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/23312165231182289
journals.sagepub.com/home/tia



Chiara Valzolgher^{1,2} , Mariam Alzahr^{3,4}, Valérie Gaveau², Aurélie Coudert⁵, Mathieu Marx^{3,4}, Eric Truy^{2,5}, Pascal Barone³, Alessandro Farnè^{1,2,6} and Francesco Pavan^{1,2,7}

Abstract

Lateralized sounds can orient visual attention, with benefits for audio-visual processing. Here, we asked to what extent perturbed auditory spatial cues—resulting from cochlear implants (CI) or unilateral hearing loss (uHL)—allow this automatic mechanism of information selection from the audio-visual environment. We used a classic paradigm from experimental psychology (capture of visual attention with sounds) to probe the integrity of audio-visual attentional orienting in 60 adults with hearing loss: bilateral CI users ($N = 20$), unilateral CI users ($N = 20$), and individuals with uHL ($N = 20$). For comparison, we also included a group of normal-hearing (NH, $N = 20$) participants, tested in binaural and monaural listening conditions (i.e., with one ear plugged). All participants also completed a sound localization task to assess spatial hearing skills. Comparable audio-visual orienting was observed in bilateral CI, uHL, and binaural NH participants. By contrast, audio-visual orienting was, on average, absent in unilateral CI users and reduced in NH listening with one ear plugged. Spatial hearing skills were better in bilateral CI, uHL, and binaural NH participants than in unilateral CI users and monaurally plugged NH listeners. In unilateral CI users, spatial hearing skills correlated with audio-visual-orienting abilities. These novel results show that audio-visual-attention orienting can be preserved in bilateral CI users and in uHL patients to a greater extent than unilateral CI users. This highlights the importance of assessing the impact of hearing loss beyond auditory difficulties alone: to capture to what extent it may enable or impede typical interactions with the multisensory environment.

Keywords

selective attention, auditory cues, spatial hearing, cochlear implant, binaural hearing, multisensory, unilateral hearing loss

Received 15 September 2022; Revised 25 May 2023; accepted 29 May 2023

Introduction

Selective attention is a fundamental tuning process that improves perception (Mehrpour et al., 2020). In everyday life, abrupt sensory events in the environment capture and orient selective attentional resources, leading to increased perceptual processing of other stimuli occurring in the same portion of space (Bashinski & Bacharach, 1980; Posner, 1980; Posner et al., 1978). Since the 1980s, a quintessential paradigm for the study of visual-attention orienting has been the Posner cueing task. In this task, the participant's attention is captured by a visual event presented for a brief period before a visual target. Participants are instructed to promptly respond to the target, and their response times are typically faster when the target appears on the same side as the preceding visual stimulus (congruent), as compared to when the two visual events occur on opposite sides (incongruent) (Posner, 1980; for a review see also Carrasco, 2011).

This attention-orienting mechanism can also occur across sensory modalities (Hillyard et al., 2016; Spence & Driver,

1997). For instance, when asked to discriminate the elevation of a visual target, an auditory event originating in the same versus opposite side of the space just before the target's appearance facilitates correct responses (Ho & Spence, 2005; Lee & Spence, 2015; Spence & Driver, 1997; Spence

¹Center for Mind/Brain Sciences - CIMEC, University of Trento, Rovereto, Italy

²Integrative, Multisensory, Perception, Action and Cognition Team, Lyon Neuroscience Research Center, Lyon, France

³Centre de Recherche Cerveau & Cognition, Toulouse, France

⁴Hospices Civils, Toulouse, France

⁵Hospices Civils de Lyon, Lyon, France

⁶Neuro-immersion, Lyon, France

⁷Centro Interuniversitario di Ricerca « Cognizione, Linguaggio e Sordità », Rovereto, Italy

Corresponding Author:

Chiara Valzolgher, Center for Mind-Brain Sciences – CIMEC, University of Trento, Corso Bettini 31, 38068 Rovereto (TN), Italy.
Email: chiara.valzolgher@unitn.it



& Santangelo, 2009 for a review). This shows that lateralized sounds can orient visual attention (for a description of the neural mechanisms subtending this behavioural effect, see: Feng et al., 2017; Romei et al., 2012; Störmer et al., 2009). In other words, sounds on a congruent side with respect to a subsequent visual target improve visual processing compared to sounds occurring on an incongruent side. In several circumstances, these multisensory selective-attention effects can emerge beyond voluntary control, revealing a substantial degree of automaticity (Koelewijn et al., 2010; Mazza et al., 2007; McDonald et al., 2000 for discussion).

Perturbed auditory spatial cues can impact audio-visual attention orienting (Shinn-Cunningham & Best, 2008). One notable example comes from people with deafness using cochlear implants (CI), a neural prosthesis that substitutes the natural ear by electrically stimulating the acoustic nerve (Moore & Shannon, 2009; Wilson, 2019). Unilateral CI users (uCI), who experience a substantial alteration of auditory spatial cues, with consequent difficulties in localizing sounds, do not benefit from audio-visual-attention orienting (Pavani et al., 2017). Rapid and effective attention orienting towards the speaker also allows access to visual information relevant to speech understanding (i.e., lip reading; Dorman et al., 2020; van Hoesel, 2015).

In the present study, we aimed to characterize how different conditions of hearing loss and assisted hearing resulting in perturbed auditory spatial cues impact information selection from the audio-visual environment. Specifically, we studied audio-visual attention orienting in three populations with hearing loss: uCI, bilateral CI users (bCI), and people with unilateral hearing loss (uHL). bCI are a model of binaural auditory processing obtained through assisted artificial hearing, whereas uCI are a model of an asymmetric artificial hearing experience. It is now well-established that bCI, for whom partial recovery of binaural cues is possible, show better sound localization abilities than uCI (e.g., Murphy et al., 2011; Seeber & Fastl, 2008; van Hoesel & Tyler, 2003). Similar to uCI, people with uHL also experience strong asymmetrical processing of auditory cues. However, their residual hearing experience is natural, because the acoustic inputs are not conveyed by a technological device, and they can often exploit monaural spectral shape cues at the hearing ear, which are key elements of natural listening (Van Wanrooij & Van Opstal, 2004, 2005). For comparison with these hearing-impaired groups, we also included a group of normal-hearing (NH) people listening with both ears or with one ear temporarily plugged (to degrade the binaural experience). Note that NH people with one ear plugged experience monaural listening as the result of an acute alteration, while uCI as well as uHL people experience hearing asymmetry longer.

Overall, we expected the ability to orient visual attention through sounds to be more effective in those groups in which auditory spatial cues are most preserved. This is because decreased accuracy or precision in spatial hearing makes

the correspondence between the auditory cue and the subsequent visual target more difficult to appreciate. In addition, any bias in localization could result in orienting attention towards the wrong location, effectively disrupting the congruence. Specifically, for uCI, for whom binaural and monaural auditory spatial cues are substantially perturbed, we expected to replicate the difficulty in using sounds to capture visual attention, as reported previously (Pavani et al., 2017). That is, we predicted no facilitation of processing (in terms of response times and accuracy) for visual targets preceded by sounds appearing on the same side of space, compared to visual targets preceded by sounds appearing on the opposite side (the so-called audio-visual *cueing effect*). For bCI, we expected a partial recovery of this audio-visual-attention orienting ability, due to better access to binaural auditory spatial cues. For people with uHL, we anticipated two alternative scenarios: either a reliable audio-visual cueing effect, due to their partially preserved auditory spatial cues allowing sufficient analysis of sound direction; or a reduced or absent audio-visual cueing effect, due to the asymmetry of the hearing experience. Finally, the NH groups served as a baseline reference for binaural listening when tested with both ears free, as well as for the monaural listening condition when tested with one ear plugged.

Materials and Methods

Participants

Twenty bCI participants, 20 uCI participants (including 11 bimodal listeners), and 20 uHL were recruited to participate in the study. One uHL and two uCI users did not complete the study (one uHL and one uCI abandoned the experiment, and the other uCI did not match all criteria to participate in the study). Mean ages of the resulting participants were: bCI (45.6, SD = 13.1, 7 males), uCI users (46.3, SD = 16.0, 9 males), uHL (52, SD = 11.8, 8 males), with no age difference between groups ($F(2,54) = 1.26, p = 0.29$). Eleven uCI users wore one hearing aid, in the non-implanted ear, and were tested in this bimodal listening condition to retain their everyday listening experience during the study (hearing threshold in the contralateral ear, either aided or unaided: pure tone average [PTA] = 59.7, SD = 23.7, range = 35–120). Detailed information about all CI users and uHL, with PTA thresholds (aided and unaided), are reported in Supplemental Materials (see Tables S1-S2-S3). We also recruited 20 NH participants, tested in binaural and monaural listening conditions (mean age = 29.4, SD = 10.5, 5 males). The mean age of NH was different from uHL ($t(37) = 6.29, p < .001$), bCI ($t(38) = 4.29, p < .001$), and uCI ($t(28.85) = 3.79, p < .001$).

All had normal or corrected-to-normal vision and reported no motor or vestibular deficits, nor any history of neurological or psychiatric disorders. NH, bCI, and uCI participants were recruited and tested in the otorhinolaryngology

department of the civil Hospital Edouard Herriot (HEH) in Lyon (France). uHL participants were recruited and tested in the otorhinolaryngology department of the University Hospital of Purpan (CHU, Purpan) in Toulouse (France). Before starting the experiment, all participants signed an informed consent document, which had received ethical approval from the national ethics committee in France (Ile de France X, N° ID RCB 2019-A02293-54) and registered as clinical trial (clinicaltrials.gov: NCT04183348). The present work focuses on the data collected in the context of a broader sound localization training protocol during the pre-training session only (full results are published in Valzolgher, Todeschini et al., 2022; Valzolgher, Gatel, et al., 2022; Valzolgher, Bouzaid, et al., 2023; and Alzaher et al., 2023, for NH, bCI, uCI, and uHL participants, respectively). Although the sample size was determined by the research question of the training studies, it is noteworthy that the sample size in each of the groups matches the numerosity of a previous study that investigated a similar experimental question (Pavani et al., 2017, $N = 17$).

Stimuli, Procedure, and Apparatus

All participants performed an audio-visual attentional-orienting task (10 min) and a sound-localization task (15 min). CI users and uHL completed the tasks once, whereas NH participants completed the tasks in both binaural and monaural listening conditions. Monaural listening was obtained by occluding the right ear with an ear plug (3 M PP 01 002; attenuation values from the manufacturer: high frequencies = 30 dB SPL; medium frequencies = 24 dB SPL; low frequencies = 22 dB SPL), and a monaural ear muff (3 M 1445, modified to cover only the right ear; attenuation values from the manufacturer: high frequencies = 32 dB SPL; medium frequencies = 29 dB SPL; low frequencies = 23 dB SPL).

Visual Attention Capture With Auditory Spatial Cues. We assessed audio-visual attentional orienting ability of participants (see Koelewijn et al., 2010) by testing them in a visual discrimination task with lateralized cueing sounds (AV cueing). Participants were asked to discriminate the elevation of a visual target presented in the upper or lower hemifield with respect to the horizontal meridian passing through visual fixation (1.15° above or below the meridian), either in the left or right hemifield (10° from fixation). Crucially, the visual target was always preceded by a sound presented either from the same (congruent) or the other (incongruent) side of the space (emitted by one of two speakers, located 20° to the left and right of the fixation). When performing AV cueing, participants sat at a table with the head braced on a chinrest (see Figure 1A).

Each trial started with a white fixation cross appearing in the center of a screen, located in front of the participants, and remaining visible until response. After a random delay (450–

600 ms), an auditory cue (white noise; 60 dB SPL, as measured from head position) lasting 100 ms was emitted by one of the two speakers positioned on the sides of the screen. At sound offset, the visual target was presented. This consisted of a filled white circle (20-pixel radius, 0.5° of visual angle) appearing on a black background for 140 ms. In half of the trials, the visual target and the sound cue appeared in the same hemispace (congruent trials), while in the other half, the visual target and the sound cue appeared in opposite hemispaces (incongruent trials). Hence, sound position was not predictive of either visual target side or elevation. Participants were asked to fixate throughout the task and to indicate as quickly and accurately as possible the elevation of visual targets using the up/down arrows keys on an external keyboard with their right index/middle finger (2000 ms timeout). At the end of each block, participants received feedback on accuracy (percentage of correct responses) and mean response time (in ms). They were explicitly told that sounds were entirely task-irrelevant. The experiment started with eight practice trials, followed by 128 trials of both congruent and incongruent audio-visual conditions, in randomized order. It lasted approximately 10 min.

Sound Localization Task. A measure of the participants' sound localization abilities was also conducted in relation to the AV cueing task. Note, however, that this sound localization task involved multiple sound sources, using hybrid virtual reality (i.e., with real sounds delivered in a visual virtual-reality scenario) and with unconstrained head posture (participants were free to move the head after sound onset). These methodological choices reflected the general aims of the training paradigm. We include this measure in the current manuscript because it provides useful information about the varied sound localization skills across groups. When performing the sound localization task, participants were asked to wear a head-mounted display (HMD, Vive Enterprise), and they were immersed in a virtual room that matched the size of the real one they were in, but devoid of objects. Real sounds were played in free-field by a single speaker (JBL GO Portable), moved by the experimenter, and positioned by following visual instructions on a monitor in one of eight possible positions, with four different azimuths in front space (-67.5° , -22.5° , 22.5° , or 67.5° ; with respect to the participant's body midline), two different elevations (5° above and 10° below ear-level to increase uncertainty), and a single distance (55 cm from the center of the head). All pre-determined positions were computed for each trial based on the initial head position (for a detailed description of the experimental setup, see Coudert et al., 2022; Valzolgher, Alzhaler et al., 2020; Valzolgher, Verdelet et al., 2020). Participants were told that sound targets could be delivered anywhere in the 3D space around them and were instructed to indicate the sound position using the head as a pointer. At the beginning of each trial, participants were asked to direct their gaze in front of them by aligning the head with a central fixation

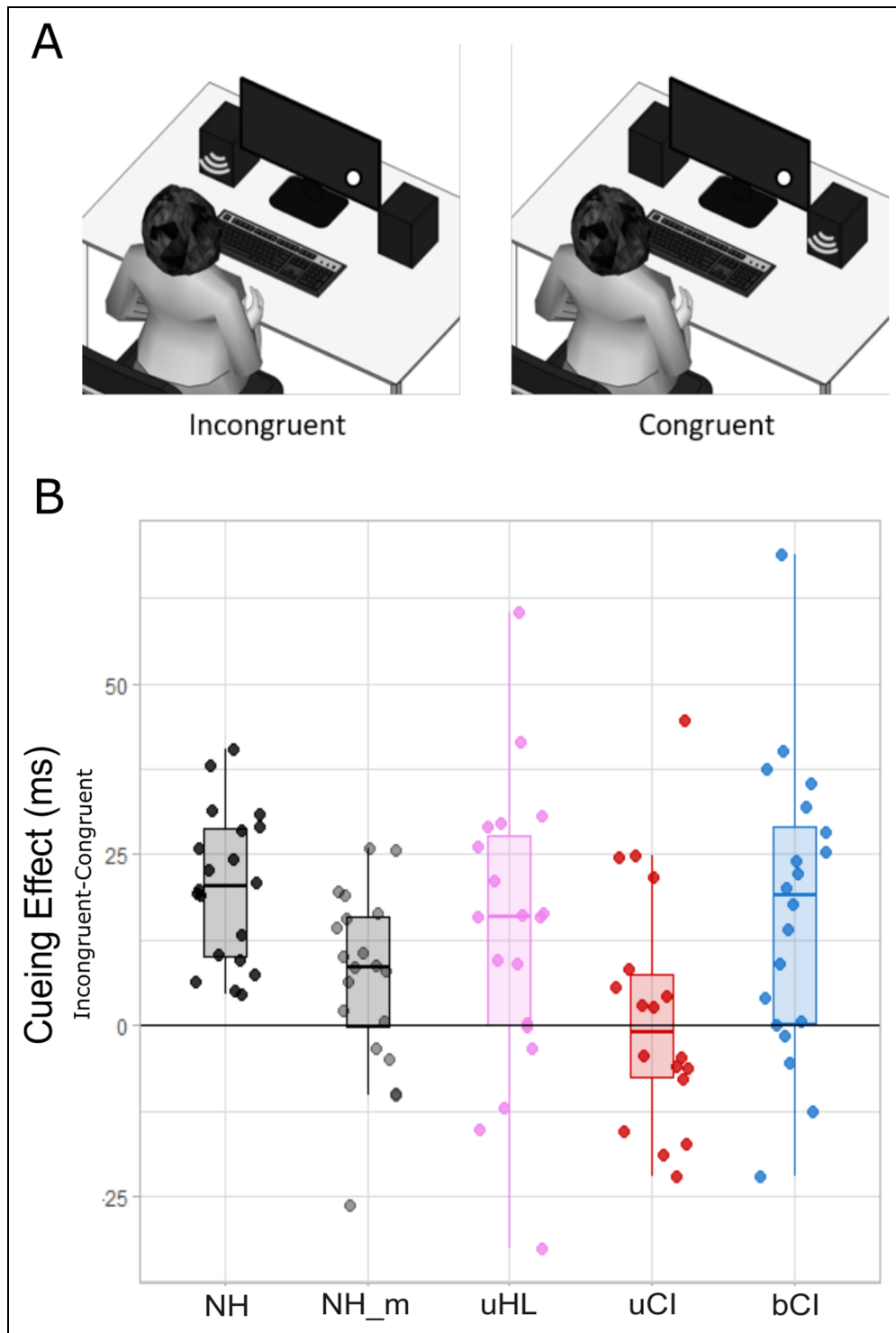


Figure 1. AV cueing. (A) Experimental setting: left: example of the incongruent audio-visual condition (the sound and the visual stimulus are presented on the opposite sides of the space); right: example of the congruent condition (the sound and the visual stimulus are presented on the same side of the space). (B) Cueing effect (ms) is calculated for each participant as the difference in response times between incongruent and congruent conditions. Normal hearing in black, NH; normal hearing with one ear plugged in grey, NH_m; unilateral hearing loss in pink, uHL; unilateral cochlear implant users in red, uCI and bCI users in blue.

cross. When the correct posture was reached, the fixation cross turned from white to blue and the sound was delivered.

The sound consisted of white noise of 3 s, amplitude-modulated at 2.5 Hz, and delivered at about 65 dB SPL, as

measured from the participant's head. We choose to use white noise in the sound localization training protocol, as it is robustly localized (see also Valzolgher, Alzhaler, et al., 2020). During sound emission, participants were free to move their heads and, at the end of the sound, they were instructed to point with their heads in the direction of the sound (i.e., they were instructed to point with the nose in the direction of the sounds) and to validate their response by pressing a button on a controller. The sound localization task comprised 40 trials (5 repetitions for each of the 8 sound positions), delivered in random order. Five practice trials were also completed but discarded from the analyses. The task lasted approximately 15 min.

Statistical Analysis

To study the AV cueing task, we used reaction time (RT) in milliseconds (ms) and accuracy for visual discrimination. Specifically, we examined to what extent the auditory cue on the congruent side with respect to the visual target improves performance when compared to the incongruent side. The RT and absolute error distributions were checked via quantile-to-quantile plot inspection, and deviant trials were excluded from analysis (respectively about 1% of the RT of correct trials and 4% of absolute error). Furthermore, we corrected the skewness of the distributions by log-transforming the variables when necessary (Baayen & Milin, 2010). To study sound localization, we focused instead on absolute error in azimuth (deg), signed error (deg), and left-right discrimination (%). For the localization task, 0.34% of trials were lost due to lack of data tracking or subject error in the procedure.

Analyses were conducted using linear mixed-effect modeling (LME), except for accuracy data for which a binomial generalized linear mixed-effect model (GLME) was adopted. All models were run using R (version 1.4.1106), employing the R-packages *emmean*, *lme4*, *lmerTest*, *car*, in R Studio (Bates et al., 2014; Fox & Weisberg, 2021). Throughout the Results section, we report the value of Chi-Square obtained by the deviance table extracted using the R function *ANOVA* (package *car*) and the post-hoc comparisons obtained using the R function *emmeans* (package *emmean*) which included Tukey correction by default. When variables were calculated by collapsing the trials (i.e., variable error), we instead adopted non-parametric comparisons between groups, such as the Kruskal-Wallis test or Wilcoxon Test, with Dunn post-hoc test corrected with Holm. Data can be retrieved from osf.io/pw5xg.

Results

To examine to what extent capturing visual attention changes with perturbed auditory spatial cues, we studied RT and accuracy in the AV cueing as a function of group. We entered visual-discrimination RTs for correct trials into an LME model with congruency (congruent vs. incongruent) and group (binaural NH, uHL, bCI, and uCI) as fixed

effects, and participant intercept as random effect. NH participants showed a well-documented capture of visual attention by sounds: they responded faster when sounds and visual targets were delivered on the same side compared to the opposite (congruent = 223 ± 53 ms vs. incongruent = 243 ± 58 ms). This resulted in a significant difference between the two conditions (pairwise comparison: $z = 7.97$, $p < .001$; Figure 1B). From now on, we will refer to this difference as the cueing effect or CE.

This CE was also evident and statistically significant at the single-group level in uHL (congruent = 340 ± 100 ms vs. incongruent = 354 ± 98 ms; CE = 14 ms; $z = 3.25$, $p = .001$) and in bCI users (congruent = 294 ± 84 ms vs. incongruent = 311 ± 86 ms; CE = 17 ms; $z = 4.88$, $p < .001$). By contrast, it decreased substantially in uCI users (congruent = 339 ± 98 ms vs. incongruent = 341 ± 98 ms; CE = 2 ms; $z = 0.46$, $p = 0.65$). This resulted in a significant interaction between congruency and group ($X^2(2) = 35.14$, $p < .001$); the main effect of group ($X^2(1) = 20.72$, $p < .001$) and the main effect of congruency ($X^2(1) = 63.52$, $p < .001$) were also significant¹. When NH participants tested monaurally were compared with the other groups, two notable findings emerged: first, the NH CE also decreased compared to the CE of NH tested binaurally (congruent: 237 ± 64 ms; incongruent: 244 ± 62 ms; CE = 7 ms; $X^2(1) = 12.94$, $p < .001$); second, they became more comparable to uCI users ($z = 0.45$, $p = .65$), whereas CE was more prominent in uHL ($z = 3.13$, $p = .002$) and bCI ($z = 4.81$, $p < .001$). This was confirmed by a significant main effect of group in the LME model comparing monaural NH listeners with uHL, bCI, and uCI users ($X^2(1) = 13.82$, $p = .003$).

Accuracy in the AV cueing task was very high overall (96.9%). We nonetheless studied accuracy as a function of group to exclude possible speed-accuracy trade-offs in our findings (e.g., RT advantages at the expense of accuracy). Accuracy data were analyzed using a binomial GLME, with similar fixed and random effects as the model used for RT data. We found that binaural NH participants were more accurate in congruent than incongruent trials (97.5% vs. 96.0%; CE = 1.5%; within-group pairwise comparison, $z = 2.20$, $p = .02$). By contrast, no such difference emerged for uHL (congruent: 96.3%; incongruent: 96.5%; CE = 0.2%, $z = 0.37$, $p = 0.71$), bCI (congruent: 97.7%; incongruent: 97.5%; CE = 0.2%, $z = 0.42$, $p = 0.68$) or uCI users (congruent: 96.3%; incongruent: 97.9%, CE = -1.6%, $z = 2.49$, $p = .01$). This resulted in a significant interaction between congruency and population ($X^2(2) = 11.36$, $p = .01$); the main effect of congruency also reached significance ($X^2(1) = 4.87$, $p = .03$). When NH participants were tested monaurally, the difference between congruent and incongruent trials was no longer present ($z = 0.48$, $p = .63$), though a comparison with the binaural listening condition yielded a marginally significant difference (congruency and listening condition, $X^2(2) = 3.52$, $p = .06$), and their CE did not differ from that measured among uHL, bCI, and uCI users

(all $ps > .19$). In brief, no speed-accuracy trade-off emerged in our findings.

Having established that capturing visual attention changed as a function of group, we examined if CE in the groups with asymmetric hearing (monaural NH, uCI, and uHL) was most prominent (or only present) for stimuli ipsilateral to their best-hearing side (i.e., the unplugged-ear side for NH-m, the CI side for uCI, and the best-ear side for uHL). The rationale for this additional analysis was that asymmetric hearing could result in biases in perceived sound position towards the best hearing side. If this is the case, any effect of congruency should interact with the laterality of the sound with respect to the best hearing side. To this end, we ran further analyses considering group, side of sound (ipsilateral or contralateral to the best-hearing side) and condition (congruent vs. incongruent) as variables. No interaction between sound position and congruency emerged, meaning that the effect of congruency was not specific to the “best” side ($X^2(1) = 0.40$, $p = .52$). Unexpectedly, we observed a main effect of group ($X^2(1) = 20.91$, $p < .001$; uHL were slower than both NH with one ear plugged and uCI users), and a significant interaction between sound position and population ($X^2(1) = 8.28$, $p = .02$; uCI users were faster when sound came from the best-hearing side, while uHL were faster when sound came from the other side). These results were unexpected, and they are not discussed further (Figure 1B).

Sound Localization Abilities and the Relation With the AV Cueing Task

Participants’ sound localization abilities were also measured and examined in relation to AV cueing. Figure 2A illustrates the distribution of sound localization responses separately for each target position (vertical dashed lines) for NH (for either binaural or monaural listening), uHL, uCI, and bCI participants. Clearly, sound localization differed in the various groups with respect to accuracy (compare the peak of the distribution with the matching color dashed lines) and precision (width of the distribution). To describe sound localization skills of the different groups and to examine whether any relation exists between this ability and performance in the AV cueing task, we now focus on the horizontal absolute error. Similar analyses on other measures of errors describing different aspects of acoustic-space-perception ability are also available in Supplemental Materials: the proportion of left-right discrimination, the variable error, and signed error (Figure S1 A, B, and C).

We entered horizontal absolute errors into an LME model with group (NH with binaural hearing, uHL, bCI, and uCI users) as fixed effects, using participant and sound position as random effects, since they can introduce variability that we did not choose to consider at this stage. As expected, binaural NH listeners performed better ($4.6^\circ \pm 4.8^\circ$) than uHL participants ($24.2^\circ \pm 16.6^\circ$, $t = 7.96$, $p < .001$), uCI

users ($50.1^\circ \pm 22.5^\circ$, $t = 12.18$, $p < .001$) and bCI users ($24.2^\circ \pm 15.5^\circ$; $t = 8.82$, $p < .001$; main effect of group: $X^2(2) = 161.18$, $p < .001$). Furthermore, sound localization errors were smaller in bCI than in uCI users ($t = 4.06$, $p < .001$). When tested monaurally, NH participants performed worse ($21.1^\circ \pm 11.4^\circ$; $X^2(1) = 1560.83$, $p < .001$), as revealed by an LME considering only NH participants in the two listening conditions. A further LME model comparing monaural NH listeners with uHL, bCI, and uCI participants showed that monaural NH performed similarly to bCI users ($t = 0.22$, $p = .99$) and uHL ($t = 0.46$, $p = .97$), but better than uCI users ($t = 3.56$, $p = .004$; main effect of group: $X^2(1) = 20.96$, $p < .001$). Furthermore, the performance of bCI users and uHL did not differ ($t = 0.25$, $p = .99$), and both performed better than uCI (both $ps < .002$, Figure 2B).

To explore if attention-orienting effects were related to these sound localization skills, we ran correlation analyses between CE and sound localization error. No significant correlation emerged for NH participants (irrespective of listening conditions; $ps > .37$), for uHL ($ps > .09$), or for bCI users ($ps > .62$). By contrast, a significant negative correlation emerged for uCI users: the larger the localization error, the smaller the CE ($R = -0.55$, $p = .02$; Figure 2C). While none of these analyses remained significant when corrected for multiple comparisons using Bonferroni, it is worth noting that the correlation linking spatial abilities and CE in uCI users was present for three of four dependent variables we considered (see “other error measures” in Supplemental Materials).

Finally, to investigate the effect of uCI hearing experience on attentional ability, we ran correlation analyses between PTA threshold at the non-implanted ear (as an index of hearing asymmetry; 59.7 ± 23.7) and CE ($R = -0.39$, $p = .11$) or sound localization absolute error ($R = 0.22$, $p = .38$). Similarly, to investigate the effect of uHL severity, we ran correlation analyses between PTA threshold at the worse ear (103.8 ± 23.3) and CE ($R = -0.12$, $p = .62$) or sound localization absolute error ($R = 0.38$, $p = .11$). None of these analyses reached significance (see also Supplemental Materials). The low variability of PTA could affect the ability to test this relationship. These analyses will not be discussed further.

Discussion and Conclusion

In this study, we examined audio-visual-attention orienting in adults with hearing loss, asking to what extent perturbed auditory cues resulting from altered hearing experience and the use of unilateral or bilateral CI or uHL attenuate this fundamental multisensory mechanism. Results showed comparable audio-visual attention orienting in bCI, people with uHL, and NH participants tested in binaural listening. By contrast, audio-visual-attention orienting was, on average, absent in uCI and reduced in NH listeners performing the task with one ear plugged. Consistent with results observed

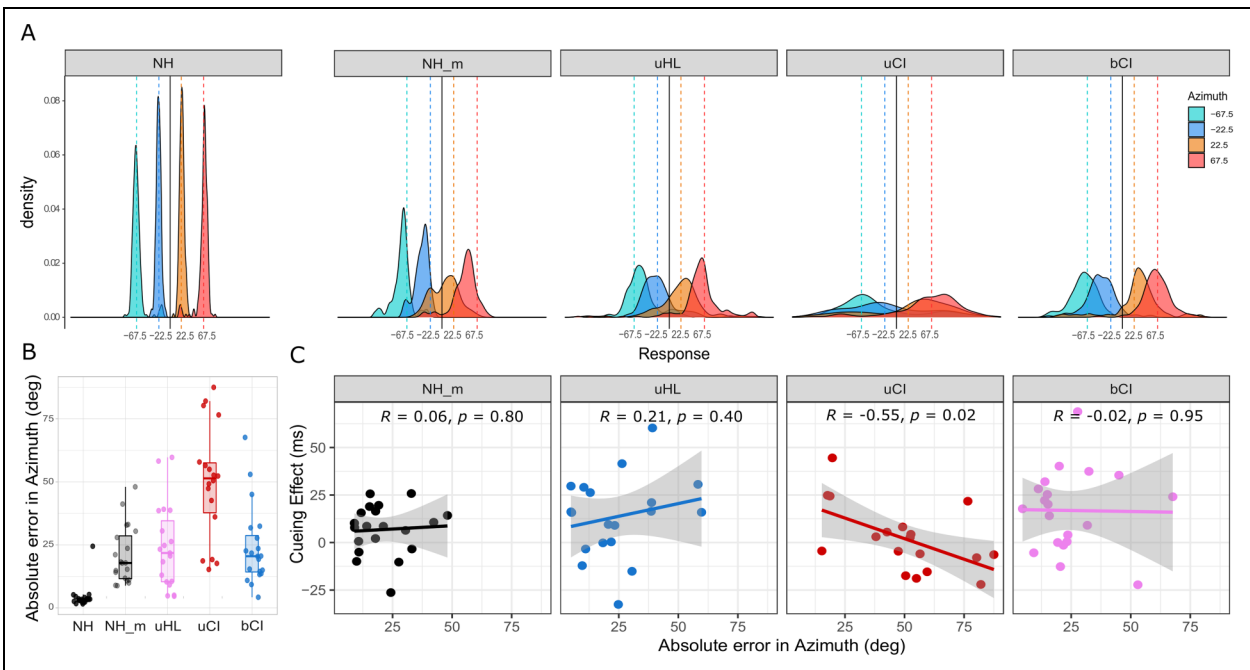


Figure 2. Sound-localization task. (A) All-trial-response density plotted using `geom_density` function in R-studio. This function draws a smoothed version of the histogram (kernel density estimate). Data are colored as a function of target azimuth positions (-67.5° , -22.5° , 22.5° , and 67.5°) and represented separately as a function of population. Note that the side of hearing loss for uHL and plugged side for NH listeners with one ear plugged was unified to the right for all participants. (B) Absolute error in azimuth as a function of population. Normal hearing in black, NH; normal hearing with one ear plugged in grey, NH_m; unilateral hearing loss in pink, uHL; unilateral cochlear implant users in red, uCI and bilateral cochlear implant users in blue, bCI. (C) Correlation between absolute error in azimuth (deg) and cueing effect (ms) in NH with one ear plugged, uHL, bCI, and uCI users.

in the AV cueing task, spatial hearing skills measured in the sound localization task were better in bCI, people with uHL, and binaural NH participants, compared to uCI and monaurally plugged NH listeners. These results indicate that the binaural hearing experience is crucial to restoring audio-visual-attention-orienting ability when the listening experience is restored via CI. They also suggest that having a natural listening experience (even if asymmetrical) allows sufficient analysis of the direction of the sound to enable audio-visual-attention orienting.

Visual-attention-orienting ability was easily altered in NH individuals when we manipulated their hearing experience using the ear plug. Despite the fact that they were significantly younger than uHL (29.4 ± 10.5 vs. 52 ± 11.8 , respectively) and that the PTA in the plugged ear was less severe than the PTA in the deaf ear of uHL ($+53$ dB when considering the combined effect of plug and ear-muff vs. 103.8 ± 23.3 , respectively), the immediate plug interfered significantly with audio-visual-attention orienting. This result suggests an important adaptation in uHL, which may exploit their lifelong asymmetrical hearing experience. Future studies could examine if, given enough time, NH participants wearing a plug recover their audio-visual attention orienting skills, perhaps in a similar way to how they can re-learn to localize sounds with new ears (e.g., Hofman et al., 1998;

Irving & Moore, 2011; Kumpik & King, 2019; Trapeau & Schönwiesner, 2015). Furthermore, future studies could also include a unisensory control condition to test if adaption might be purely auditory or instead involve multisensory processes (for instance, Pavani et al., 2017).

In hearing animals, sounds allow the detection of changes in the environment beyond the limitations of the visual field, and spatial hearing is essential for directing head and eyes towards novel events (Heffner, 1997). The importance of this audio-visual coupling in attention orienting is evident in the correlation between sound localization thresholds and the width of the best field of view (i.e., area of visual field perceived with the highest visual resolution, which can be operationalized as the retinal region with a density of ganglion cells of at least 75% of maximum) across mammals (Heffner & Heffner, 1992). Humans, whose best field of view is less than 2 degrees wide, rely on efficient spatial hearing to orient their visual attention in the environment, both explicitly (head and eye orienting) and implicitly (without eye movements, as here). Having preserved audio-visual links in attention (through natural hearing) or having them restored through CIs could promote a virtuous cycle in sensory processing. Rapidly directing visual attention to a sound could trigger those mechanisms that foster its processing (Best et al., 2007; Turri et al., 2021). The importance

of this cycle is particularly evident in complex audio-visual scenarios, such as concurrent conversations between multiple visible speakers in a room, a common and challenging auditory context. Rapidly engaging the relevant speaker (an audio-visual target) could favor the selection of its spoken message (the relevant auditory stream), as well as access to visual information carried by faces (e.g., lip reading), as demonstrated in previous studies on bilateral CI users (van Hoesel, 2015; Dorman et al., 2016). Moreover, complex audio-visual scenarios require a high level of attentional resources, especially when auditory signals are degraded, as for CI users (see Stacey et al., 2014). Potentiating sound localization to favor fast access to visual information may also contribute, especially for CI users, in reducing attention demands. Future studies aimed at enhancing sound localization abilities may also consider the consequences of their training on visuo-attentional abilities. Future studies could also control more effectively the heterogeneity of the uCI group by directly comparing people wearing a CI with a clear unilateral acoustic experience and people who benefit from bimodal stimulation. Our prediction is that bimodal listening could favor AV cueing effects while experiencing unilateral hearing through CI could dramatically reduce this audio-visual advantage. Preliminary evidence in this direction is evident in the correlation between CE and localization performance in uCI (see Figure 2 and Figure S1).

One limitation of the present work concerns the relation between the two tasks examined in this study. As mentioned in the Introduction, all participants tested in the present study were recruited in the context of a broader project aimed at assessing the effects of a training paradigm on sound localization. For this reason, the sound localization task was not originally conceived to study the relationship between sound localization abilities and audio-visual-attention-orienting mechanisms. Sound duration differed substantially in the two tasks and, most notably, head movements were prevented in the AV cueing task but not during the sound localization task. Our preliminary investigation of the link between AV cueing and sound localization provides initial evidence of the potential role of spatial-hearing abilities in mediating the cueing effect, specifically in uCI users. Nonetheless, future studies should examine this relationship using a sound localization task more closely matching the spatial-hearing demands of the AV task. For instance, sound localization abilities could be measured using shorter sounds and preventing head movements. In such a way, the auditory spatial information available in the AV cueing task would fully match that available when measuring spatial hearing.

Another scenario would be to change the AV cueing paradigm to allow head movements and to present acoustic stimuli with a longer duration. Although we adopted here a gold-standard measure of audio-visual-attention-orienting ability (Spence & Driver, 1997; for a review see Störmer, 2019), future studies could measure it by presenting semi-

naturalistic environments simulated in virtual reality, ideally using a variety of stimulation positions and tasks (see Hartley, 2022). These additional experiments would allow the testing of the consequences of being able to localize sounds for visual-orienting-attention ability in situations in which quickly directing visual attention is ecologically relevant to solving the task (e.g., to facilitate speech understanding). Furthermore, letting participants behave in a natural manner and tracking their gaze and body movements should allow measurement of both their spontaneous behavior and the implementation of motor-behavioral strategies that may influence attention-orienting abilities (see for instance, Brimijoin et al., 2012 and Hadley et al., 2021).

To conclude, our results highlight the importance of assessing the impact of hearing loss beyond auditory difficulties alone, capturing to what extent they may enable or prevent typical interactions with the multisensory environment. Having preserved audio-visual links in attention (through natural hearing) or restoring them through CIs could permit a more complex experience of the world in adults.

Acknowledgments

We are grateful to all participants; we thank Laura Ratenet and Julie Gatel for their help in coordinating recruiting of CI users, as well as the engineers of Neuro-immersion in Lyon for having developed the software for sound localization in VR. We also thank Giordana Torresani for graphical support (Figure 1A) and Ben Timberlake for English revisions.


Declaration of Conflicting Interests

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

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: C. V. was supported by a grant of the Università Italo-Francesca/Université Franco-Italienne, the Zegna Founder's Scholarship, and Associazione Amici di Claudio Demattè. F. P., A.F., P.B., and V.G. were supported by a grant of the Agence Nationale de la Recherche [ANR-16-CE17-0016, VIRTUALHEARING3D, France], by a prize of the Foundation Medisite (France), by the Neurodis Foundation (France) and by a grant from the Italian Ministry for Research and University [MUR, PRIN 20177894ZH]. The study was supported by a grant of the Agence Nationale de la Recherche [IHU CeSaMe ANR-10-UBHU-0003 and ANR 2019CE37 Blind Touch]. A.C. was supported by a grant of the Hospices Civils de Lyon. P.B. and M.M. were supported by a grant of the Agence Nationale de la Recherche [ANR-20-CE28-0016 AgeHear].

ORCID iD

Chiara Valzolgher  <https://orcid.org/0000-0002-1354-3288>

Supplemental Material

Supplemental material for this article is available online.

Note

- To control for the possibility that congruency effects were modulated by the overall reaction-time differences across groups, we re-analyzed our findings using normalized data as a dependent variable. Specifically, we calculated CE as percent reduction (%) of incongruent compared to congruent conditions. An ANOVA on percent reduction confirmed the observed group differences ($F(3,73)=6.52, p < .001$). In particular, the uCI group showed significantly lower CE ($0.45\% \pm 4.95$) compared to the NH group ($8.46\% \pm 4.51, t = 4.35, p < .001$) and the bCI group ($5.50\% \pm 6.67, t = 0.04, p = 0.05$). Meanwhile, the difference with respect to UHL was not significant ($4.01\% \pm 6.22, t = 0.24, p = .36$). In any case, the CE measured as percent reduction was significant for all groups (all p s $< .01$ on t -tests against zero), except for the uCI group ($t(17) = 0.39, p = .70$).

References

- Alzaher, M., Valzolgher, C., Verdelet, G., Pavani, F., Farnè, A., Barone, P., & Marx, M. (2023). Audiovisual training in virtual reality improves auditory spatial adaptation in unilateral hearing loss patients. *Journal of Clinical Medicine, 12*(6), 2357. <https://doi.org/10.3390/jcm12062357>
- Baayen, R. H., & Milin, P. (2010). Analyzing reaction times. *International Journal of Psychological Research, 3*(2), 12–28. <https://doi.org/10.21500/20112084.807>
- Bashinski, H. S., & Bacharach, V. R. (1980). Enhancement of perceptual sensitivity as the result of selectively attending to spatial locations. *Perception & Psychophysics, 28*(3), 241–248. <https://doi.org/10.3758/BF03204380>
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1-7. <https://doi.org/10.48550/arXiv.1406.5823>
- Best, V., Ozmeral, E. J., & Shinn-Cunningham, B. G. (2007). Visually-guided attention enhances target identification in a complex auditory scene. *Journal for the Association for Research in Otolaryngology, 8*(2), 294–304. <https://doi.org/10.1007/s10162-007-0073-z>
- Brimijoin, W. O., McShefferty, D., & Akeroyd, M. A. (2012). Undirected head movements of listeners with asymmetrical hearing impairment during a speech-in-noise task. *Hearing Research, 283*(1-2), 162–168. <https://doi.org/10.1016/j.heares.2011.10.009>
- Carrasco, M. (2011). Visual attention: The past 25 years. *Vision Research, 51*(13), 1484–1525. <https://doi.org/10.1016/j.visres.2011.04.012>
- 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 and Hearing, 43*(1), 192–205. <https://doi.org/10.1097/AUD.0000000000001090>
- Dorman, M. F., Liss, J., Wang, S., Berisha, V., Ludwig, C., & Natale, S. C. (2016). Experiments on auditory-visual perception of sentences by users of unilateral, bimodal, and bilateral cochlear implants. *Journal of Speech, Language, and Hearing Research, 59*(6), 1505–1519.
- Dorman, M. F., Natale, S., & Knickerbocker, A. (2020). Bilateral cochlear implants allow listeners to benefit from visual information when talker location is varied. *Journal of the American Academy of Audiology, 31*(07), 547–550. <https://doi.org/10.1055/s-0040-1709444>
- Feng, W., Störmer, V. S., Martinez, A., McDonald, J. J., & Hillyard, S. A. (2017). Involuntary orienting of attention to a sound desynchronizes the occipital alpha rhythm and improves visual perception. *NeuroImage, 150*, 318–328. <https://doi.org/10.1016/j.neuroimage.2017.02.033>
- Fox, J., & Weisberg, S. (2021). Using car and effects Functions in Other Functions. In *cran.pau.edu.tr*. Retrieved from <http://z.umn.edu/carbook>.
- Hadley, L. V., Whitmer, W. M., Brimijoin, W. O., & Naylor, G. (2021). Conversation in small groups: Speaking and listening strategies depend on the complexities of the environment and group. *Psychonomic Bulletin & Review, 28*(2), 632–640. <https://doi.org/10.3758/s13423-020-01821-9>
- Hartley, C. A. (2022). How do natural environments shape adaptive cognition across the lifespan? *Trends in Cognitive Sciences, 26*(12), 1029–1030. <https://doi.org/10.1016/j.tics.2022.10.002>
- Heffner, R. S. (1997). Comparative study of sound localisation and its anatomical correlates in mammals. *Acta Oto-Laryngologica, 117*(sup532), 46–53. <https://doi.org/10.3109/00016489709126144>
- Heffner, R. S., & Heffner, H. E. (1992). Visual factors in sound localisation in mammals. *Journal of Comparative Neurology, 317*(3), 219–232. <https://doi.org/10.1002/cne.903170302>
- Hillyard, S. A., Störmer, V. S., Feng, W., Martinez, A., & McDonald, J. J. (2016). Cross-modal orienting of visual attention. *Neuropsychologia, 83*, 170–178. <https://doi.org/10.1016/j.neuropsychologia.2015.06.003>
- Ho, C., & Spence, C. (2005). Assessing the effectiveness of various auditory cues in capturing a driver's visual attention. *Journal of Experimental Psychology: Applied, 11*(3), 157–174. <https://doi.org/10.1037/1076-898X.11.3.157>
- Hofman, P. M., Van Riswick, J. G., Opstal, V., & J, A. (1998). Relearning sound localisation with new ears. *Nature Neuroscience, 1*(5), 417–421. <https://doi.org/10.1038/1633>
- Irving, S., & Moore, D. R. (2011). Training sound localisation in normal hearing listeners with and without a unilateral ear plug. *Hearing Research, 280*(1-2), 100–108. <https://doi.org/10.1016/j.heares.2011.04.020>
- Koelewijn, T., Bronkhorst, A., & Theeuwes, J. (2010). Attention and the multiple stages of multisensory integration: A review of audiovisual studies. *Acta Psychologica, 134*(3), 372–384. <https://doi.org/10.1016/j.actpsy.2010.03.010>
- Kumpik, D. P., & King, A. J. (2019). A review of the effects of unilateral hearing loss on spatial hearing. *Hearing Research, 372*, 17–28. <https://doi.org/10.1016/j.heares.2018.08.003>
- Lee, J., & Spence, C. (2015). Audiovisual crossmodal cuing effects in front and rear space. *Frontiers in Psychology, 6*, 1086. <https://doi.org/10.3389/fpsyg.2015.01086>
- Mazza, V., Turatto, M., Rossi, M., & Umiltà, C. (2007). How automatic are audiovisual links in exogenous spatial attention? *Neuropsychologia, 45*(3), 514–522. <https://doi.org/10.1016/j.neuropsychologia.2006.02.010>
- McDonald, J. J., Teder-Sälejärvi, W. A., & Hillyard, S. A. (2000). Involuntary orienting to sound improves visual perception. *Nature, 407*(6806), 906–908. <https://doi.org/10.1038/35038085>
- Mehrpour, V., Martinez-Trujillo, J. C., & Treue, S. (2020). Attention amplifies neural representations of changes in sensory input at the

- expense of perceptual accuracy. *Nature Communications*, *11*(1), 1–8. <https://doi.org/10.1038/s41467-020-15989-0>
- Moore, D. R., & Shannon, R. V. (2009). Beyond cochlear implants: Awakening the deafened brain. *Nature Neuroscience*, *12*(6), 686–691. <https://doi.org/10.1038/nn.2326>
- Murphy, J., Summerfield, A. Q., O'Donoghue, G. M., & Moore, D. R. (2011). Spatial hearing of normally hearing and cochlear implanted children. *International Journal of Pediatric Otorhinolaryngology*, *75*(4), 489–494. <https://doi.org/10.1016/j.ijporl.2011.01.002>
- Pavani, F., Venturini, M., Baruffaldi, F., Artesini, L., Bonfioli, F., Frau, G. N., & van Zoest, W. (2017). Spatial and non-spatial multisensory cueing in unilateral cochlear implant users. *Hearing Research*, *344*, 24–37. <https://doi.org/10.1016/j.heares.2016.10.025>
- Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, *32*(1), 3–25. <https://doi.org/10.1080/00335558008248231>
- Posner, M. I., Nissen, M. J., & Ogden, W. C. (1978). Attended and unattended processing modes: the role of set for spatial location. In H. L. Pick, & E. Salzman (Eds.), *Modes of perceiving and processing information* (pp. 137–158). Erlbaum.
- Romei, V., Gross, J., & Thut, G. (2012). Sounds reset rhythms of visual cortex and corresponding human visual perception. *Current Biology*, *22*(9), 807–813. <https://doi.org/10.1016/j.cub.2012.03.025>
- Seeber, B. U., & Fastl, H. (2008). Localisation cues with bilateral cochlear implants. *The Journal of the Acoustical Society of America*, *123*(2), 1030–1042. <https://doi.org/10.1121/1.2821965>
- Shinn-Cunningham, B. G., & Best, V. (2008). Selective attention in normal and impaired hearing. *Trends in Amplification*, *12*(4), 283–299. <https://doi.org/10.1177/1084713808325306>
- Spence, C., & Driver, J. (1997). Audiovisual links in exogenous covert spatial orienting. *Perception & Psychophysics*, *59*(1), 1–22. <https://doi.org/10.3758/BF03206843>
- Spence, C., & Santangelo, V. (2009). Capturing spatial attention with multisensory cues: A review. *Hearing Research*, *258*(1–2), 134–142. <https://doi.org/10.1016/j.heares.2009.04.015>
- Stacey, P. C., Murphy, T., Sumner, C. J., Kitterick, P. T., & Roberts, K. L. (2014). Searching for a talking face: The effect of degrading the auditory signal. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(6), 2106–2111. <https://doi.org/10.1037/a0038220>
- Störmer, V. S. (2019). Orienting spatial attention to sounds enhances visual processing. *Current Opinion in Psychology*, *29*, 193–198. <https://doi.org/10.1016/j.copsyc.2019.03.010>
- Störmer, V. S., McDonald, J. J., & Hillyard, S. A. (2009). Cross-modal cueing of attention alters appearance and early cortical processing of visual stimuli. *Proceedings of the National Academy of Sciences*, *106*(52), 22456–22461. <https://doi.org/10.1073/pnas.0907573106>
- Trapeau, R., & Schönwiesner, M. (2015). Adaptation to shifted interaural time differences changes encoding of sound location in human auditory cortex. *NeuroImage*, *118*, 26–38. <https://doi.org/10.1016/j.neuroimage.2015.06.006>
- Turri, S., Rizvi, M., Rabini, G., Melonio, A., Gennari, R., & Pavani, F. (2021). Orienting auditory attention through vision: The impact of monaural listening. *Multisensory Research*, *35*(1), 1–28. <https://doi.org/10.1163/22134808-bja10059>
- Valzolgher, C., Alzhaler, M., Gessa, E., Todeschini, M., Nieto, P., & Verdelet, G., ... F. Pavani (2020). The impact of a visual spatial frame on real sound-source localisation in virtual reality. *Current Research in Behavioral Sciences*, *1*, 100003. <https://doi.org/10.1016/j.crbeha.2020.100003>
- Valzolgher, C., Bouzaid, S., Grenouillet, S., Gatel, J., Ratenet, L., & Murenu, F., ... F. Pavani (2023). Training spatial hearing in unilateral cochlear implant users through reaching to sounds in virtual reality. *European Archives of Oto-Rhino-Laryngology*, 1–12. <https://doi.org/10.1007/s00405-023-07886-1>
- Valzolgher, C., Gatel, J., Bouzaid, S., Grenouillet, S., Todeschini, M., & Verdelet, G., ... F. Pavani (2022). Reaching to sounds improves spatial hearing in bilateral cochlear implant users. *Ear and Hearing*, *44*(1), 189–198. <https://doi.org/10.1097/AUD.0000000000001267>
- Valzolgher, C., Todeschini, M., Verdelet, G., Gatel, J., Salemme, R., Gaveau, V., & Pavani, F. (2022). Adapting to altered auditory cues: Generalization from manual reaching to head pointing. *PloS One*, *17*(4), e0263509. <https://doi.org/10.1371/journal.pone.0263509>
- Valzolgher, C., Verdelet, G., Salemme, R., Lombardi, L., Gaveau, V., Farné, A., & Pavani, F. (2020). Reaching to sounds in virtual reality: A multisensory-motor approach to promote adaptation to altered auditory cues. *Neuropsychologia*, *149*, 107665. <https://doi.org/10.1016/j.neuropsychologia.2020.107665>
- van Hoesel, R. J., & Tyler, R. S. (2003). Speech perception, localisation, 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>
- van Hoesel, R. J. M. (2015). Audio-visual speech intelligibility benefits with bilateral cochlear implants when talker location varies. *Journal of the Association for Research in Otolaryngology*, *16*(2), 309–315. <https://doi.org/10.1007/s10162-014-0503-7>
- Van Wanrooij, M. M., & Van Opstal, A. J. (2004). Contribution of head shadow and pinna 1007 cues to chronic monaural sound localisation. *The Journal of Neuroscience*, *24*(17), 4163–4171. <https://doi.org/10.1523/JNEUROSCI.0048-04.2004>
- Van Wanrooij, M. M., & Van Opstal, A. J. (2005). Relearning sound localization with a new ear. *Journal of Neuroscience*, *25*(22), 5413–5424. <https://doi.org/10.1523/JNEUROSCI.0850-05.2005>
- Wilson, B. S. (2019). The remarkable cochlear implant and possibilities for the next large step forward. *Acoustics Today*, *15*(1), 53–61. <https://doi.org/10.1121/AT.2019.15.1.55>