

# SCIENTIFIC REPORTS



OPEN

## Interactions between egocentric and allocentric spatial coding of sounds revealed by a multisensory learning paradigm

Giuseppe Rabini<sup>1</sup>, Elena Altobelli<sup>2</sup> & Francesco Pavani<sup>1,2,3</sup>

Although sound position is initially head-centred (egocentric coordinates), our brain can also represent sounds relative to one another (allocentric coordinates). Whether reference frames for spatial hearing are independent or interact remained largely unexplored. Here we developed a new allocentric spatial-hearing training and tested whether it can improve egocentric sound-localisation performance in normal-hearing adults listening with one ear plugged. Two groups of participants (N = 15 each) performed an egocentric sound-localisation task (point to a syllable), in monaural listening, before and after 4-days of multisensory training on triplets of white-noise bursts paired with occasional visual feedback. Critically, one group performed an allocentric task (auditory bisection task), whereas the other processed the same stimuli to perform an egocentric task (pointing to a designated sound of the triplet). Unlike most previous works, we tested also a no training group (N = 15). Egocentric sound-localisation abilities in the horizontal plane improved for all groups in the space ipsilateral to the ear-plug. This unexpected finding highlights the importance of including a no training group when studying sound localisation re-learning. Yet, performance changes were qualitatively different in trained compared to untrained participants, providing initial evidence that allocentric and multisensory procedures may prove useful when aiming to promote sound localisation re-learning.

Accurate processing of sound source coordinates is central to our ability to perceive the 3D structure of the surrounding auditory scene<sup>1–3</sup>, discern signal from noise<sup>4</sup> and orient attention in space<sup>5,6</sup>. Sound localisation relies on the interpretation of auditory cues (interaural time and level differences, as well as monaural spectral cues) deriving from the interactions between sound waves, the head and external ears<sup>7,8</sup>. Initial coding of sound position occurs in head-centred coordinates<sup>9</sup>, in an *egocentric* reference frame dependent on the head position on the trunk and listener's position in the environment. However, efficient interactions with the auditory environment, require encoding of sound position also in a reference frame that remains stable when our head and our body change position. Such reference frame is termed *allocentric* (or world-centred) and it encodes the position of sounds with respect to other sound sources or external landmarks in the multisensory environment. This allocentric reference frame allows judgements on the relative positions of sounds, independent of the listener's position.

Coding of auditory space in different reference frames has previously been investigated in humans using electrophysiology, taking advantage of the Mismatch Negativity (MMN) response. In two studies using virtual sounds delivered through headphones<sup>10</sup> or free-field sounds<sup>11</sup>, Altmann and colleagues found evidence for head-centred (craniocentric) coding at the early processing stage corresponding to the MMN. However, a similar study conducted with free-field sounds by Schechtman and colleagues found that head-dependent and head-independent<sup>11</sup> sound representations may in fact coexist, already at the processing stage captured by the MMN. The latter result is in agreement with research on ferrets, indicating that a majority of neurons in the auditory cortex is tuned to egocentric sound representation, and yet a restricted subsample of cells showing head-independent position coding also exists<sup>10</sup>.

<sup>1</sup>Center for Mind/Brain Sciences (CIMEC), University of Trento, Trento, Italy. <sup>2</sup>Department of Psychology and Cognitive Sciences (DiPSCo), University of Trento, Trento, Italy. <sup>3</sup>IMPACT, Centre de Recherche en Neurosciences Lyon (CRNL), Bron, France. Correspondence and requests for materials should be addressed to G.R. (email: [giuseppe.rabini@unitn.it](mailto:giuseppe.rabini@unitn.it))

To the best of our knowledge, only one study has examined egocentric and allocentric sound coding using behavioural methods, focusing however on a special population. Studying sighted and blind individuals, Gori and colleagues<sup>12</sup> showed that congenitally blind people are highly accurate when asked to localise single sounds (an egocentric task), but they are impaired in a task that require a judgment on the relative distance between pairs of sounds (i.e., the auditory space bisection task, an allocentric task). This study revealed that sound localisation abilities in allocentric and egocentric reference frames can dissociate, at least in visually impaired people. In general, however, the issue of dissociation vs. interaction of egocentric and allocentric sound coding in humans remained largely unexplored. In the present study, we aimed to investigate this interaction in behaviour by testing if a training of sound localisation abilities performed in allocentric coordinates (auditory space bisection task) can improve spatial hearing performance in a pointing to sound task performed in egocentric coordinates.

Research on sound localisation training over the last decades has consistently revealed the auditory system plasticity when dealing with altered auditory cues<sup>13</sup>. Studies in humans<sup>14–18</sup> as well as other animals<sup>19</sup>, showed that learning new correspondences between auditory cues and coordinates in space is possible, particularly when training paradigms are employed. The theoretical relevance of these findings resides in the identification of a perceptual ability which remains plastic beyond sensitive periods in the early stages of development<sup>20</sup>. Moreover, these observations have great clinical implications, as they open the possibility of training spatial hearing in adults with hearing loss, when binaural hearing is restored through hearing aids or cochlear implants. Several learning paradigms aimed at improving spatial hearing have adopted a multisensory approach<sup>16</sup>. This approach builds on the notion that multisensory training can promote unisensory learning<sup>21</sup>. In the case of audio-visual training, the hypothesis is that the brain can calibrate the association between auditory cues and spatial locations in external space, taking advantage of the more reliable visuo-spatial information. Support for this approach emerged from converging results in monaurally plugged hearing humans<sup>22</sup> and in ferrets with bilateral cochlear implants<sup>23</sup>. To date, however, all audio-visual training paradigms have been implemented in egocentric reference frame<sup>16</sup>, with no attempt to exploit the same principles in the context of allocentric coding of sounds.

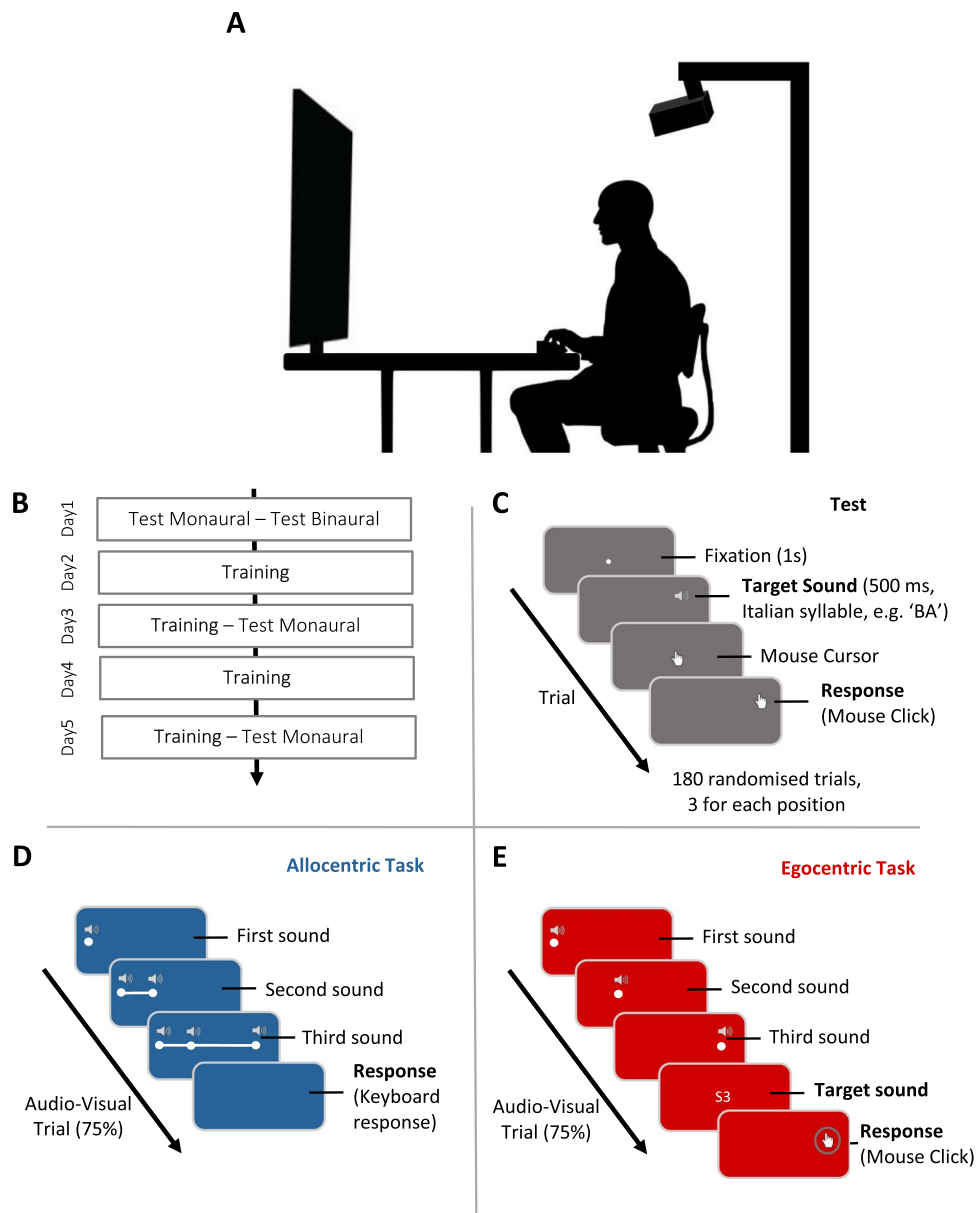
In audio-visual training paradigms the two multisensory events originate from identical positions in space (i.e., auditory and visual stimuli are spatially congruent). This perceptual coupling has been accomplished using either simultaneous or delayed multisensory stimulation. For instance, the audio-visual training developed by Strelnikov and colleagues<sup>22</sup> consisted in single sounds delivered in free-field, and paired with simultaneous and spatially congruent visual events presented using LEDs. Participants were asked to localise each sound, using a laser pointer controlled by a central knob. Notably, in order to avoid a complete reliance on visual information, the training session comprised occasional (15%) auditory-only stimulations. Differently, Majdak and co-workers<sup>24</sup>, implemented a virtual reality setup in which participants localised a single sound delivered through headphones in virtual auditory space, using a virtual reality laser pointer. No concomitant visual stimulation was delivered at this first stage. After the response, a red rotating cube was presented as feedback, to indicate the actual sound position (delayed spatially congruent stimulation). Finally, the sound was replayed once more from the same location, this time spatially and temporally coupled with the visual stimulus. Note that in both training studies, target sounds were localised with respect to the participant's body and required, in this sense, egocentric coding. An allocentric training would have require instead localisation of the sound with respect to a different object or landmark, or alternatively, the presentation of multiple sounds in different positions with a subsequent evaluation of their relative spatial location.

In the present work, we directly compared the efficacy of two multisensory trainings – one based on *allocentric* spatial processing, the other based on *egocentric* spatial processing – in promoting re-learning of correspondences between auditory cues and space. Following previous works, we altered spatial hearing by plugging one ear. Monaural listening deteriorates sound localisation performance<sup>25–27</sup> and proved an effective model for the study of auditory space re-learning<sup>28</sup>. One group of participants was trained using a multisensory (audio-visual) allocentric task, modelled on the auditory space bisection proposed by Gori and collaborators<sup>29</sup>. Three bursts of white noise were presented in sequence, paired with spatial and temporally matching visual stimuli. Participants in this group judged the relative position of the central sound with respect to the other two (see Fig. 1D). A second group of participants was trained using exactly the same audio-visual triplets, but was asked to perform an egocentric localisation task. A written instruction presented at the end of the trial instructed participants to point to one of the three sounds in the triplet. As in Strelnikov *et al.*<sup>22</sup>, in both training protocols audio-visual stimulation occurred in 75% of trials, whereas the remaining 15% consisted in auditory-only stimulation. Crucially, both groups were tested before and after training for their ability to localise a single auditory target delivered in front space. This *egocentric* sound localisation test was performed in both azimuth and elevation, because changes in performance following monaural plugging have been reported along the horizontal as well as vertical dimensions<sup>30,31</sup>. To assess training generalisation effects the stimulation used in pre- and post-training phases was a spoken syllable, instead of the white noise bursts used during training. Finally, a control group that did not perform any training was also included in the study to assess performance changes unrelated to training.

If allocentric and egocentric auditory spatial coding interact with one another, training benefits should transfer across *different* reference frames. This predicts that allocentric training should improve performance also in the subsequent egocentric task. Alternatively, partial segregation between allocentric and egocentric auditory spatial coding predicts that only the group that performed both training and test within the *same* reference frame (i.e., the egocentric training group) can improve performance.

## Material and Methods

**Participants.** Forty-five normal hearing participants took part in the study and were assigned to one of three experimental groups: Allocentric Training Group (AlloT,  $N = 15$ , mean age = 23,  $SD = 2.7$ , 10 females), Egocentric Training Group (EgoT,  $N = 15$ , mean age = 24,  $SD = 3.9$ , 12 females), Control Group (Control,  $N = 15$ , mean age = 21,  $SD = 2$ , 12 females). All participants reported no history of auditory or neurological disease and



**Figure 1.** Apparatus, overall experimental procedure and tasks. **(A)** Participants sat in front of a matrix ( $5 \times 12$ ) of loudspeakers mounted on a rectangular panel, on which visual stimuli were projected when appropriate. Response were collected through mouse cursor or keyboard. **(B)** The experiment lasted five consecutive days, with testing sessions at day 1, day 3 and day 5, and training sessions from day 2 to day 5. **(C)** During testing sessions, participants completed an egocentric sound localisation task. In each trial, a spoken Italian syllable was presented from one speaker of the stimulation matrix and participants indicated the source of the sound by left-clicking with the mouse cursor in the correspondent position. **(D)** The AlloT Group performed the acoustic space bisection task. Three consecutive sounds (white noise) were presented at different horizontal positions and participants indicated if the second sound was closer to the first or third presented sound. During audio-visual trials (75% of total) a white dot appeared on screen to mark the exact position of each sound, together with lines connecting the dots to depict the spatial relationships between sounds. **(E)** The EgoT group received exactly the same audio-visual stimulation (with the only exception of lines connecting the dots) but was instructed to localise only one of the three presented sounds using the mouse.

had normal or corrected to normal vision. Pure tone audiometry (250, 500, 1000, 2000, 4000 Hz) was used to screen participants for hearing loss, and served to assess the efficacy of the ear-plug in each participant (see Supplementary Materials and Fig. 1S). The study was conducted according to the research ethics regulation of the University of Trento and in line with the Declaration of Helsinki (1964, amended in 2013). Participants read and signed an informed consent before taking part in the experiment and received monetary reimbursement (7 €/hour) or course credits for participation.

**Apparatus.** The experiment was administered in a reverberant room ( $4.2 \times 5.0$  m; mean background noise = 50 dB SPL). The experimental setup consisted in a custom-made rectangular wooden panel ( $95 \times 60$  cm) covering a visual angle of  $43^\circ$ , in which 60 round-shaped speakers (5 cm diameter of Mylar; Pro Signal ABS-210-RC range 350–20,000 Hz, 8K, 1W RMS Power) were mounted, arranged in a matrix of 5 rows and 12 columns (see Fig. 2S). A white acoustically transparent fabric was placed in the frontal part of the panel, to prevent vision of the sound-sources and to serve as screen for projecting all visual stimulations. Visual stimuli were projected using an LCD projector (LG HW300G; resolution:  $1280 \times 1024$ ) connected to the stimulation PC (Dell Precision T3400). The panel was mounted on a wooden support and placed on a table 120 cm from participant's head, with its centre approximately at ear level (Fig. 1A). The integrated sound card of the PC, connected to an external loudspeaker for sound amplification, was used for delivering all audio signals. Each signal was switched between 30 relays (NEC MR62-4.5 USB) by using the digital output of two National Instrument boards (NIusb-6259) to activate the desired speaker. Auditory stimulation was controlled using custom-made Matlab scripts (Mathworks R2015b, 32-bit).

**Procedure.** The overall experimental protocol lasted five consecutive days (Fig. 1B). Hearing thresholds were measured by audiometric test for each participant on day 1 for right and left ear in normal hearing condition and left ear in plugged condition. In addition, hearing threshold of the plugged ear was assessed at the beginning of each testing day (see Supplementary Materials) to ensure that plugging was comparable across sessions.

On day 1, participants performed the test procedure in binaural (Binaural) and monaural hearing condition (Monaural day 1), in separate consecutive blocks. The same testing in monaural condition was repeated after two days (Monaural day 3) and on the last day (Monaural day 5). During the test procedure participants were administered a single-sound localisation task. We presented a stimulus (single spoken Italian syllable: counterbalanced number of occurrence between 'BA', 'RO' and 'GU', 500 ms duration) from a position of our stimulation setup and participants were required to localise the heard sound by moving a mouse cursor (hand icon, projected on the wide screen) to the location of the perceived sound, and validate their response with a left-click (Fig. 1C). In each testing phase we presented a total of 180 trials, 3 for each possible position in randomised order. The entire session lasted about 15 minutes and was divided into two blocks of 90 trials each.

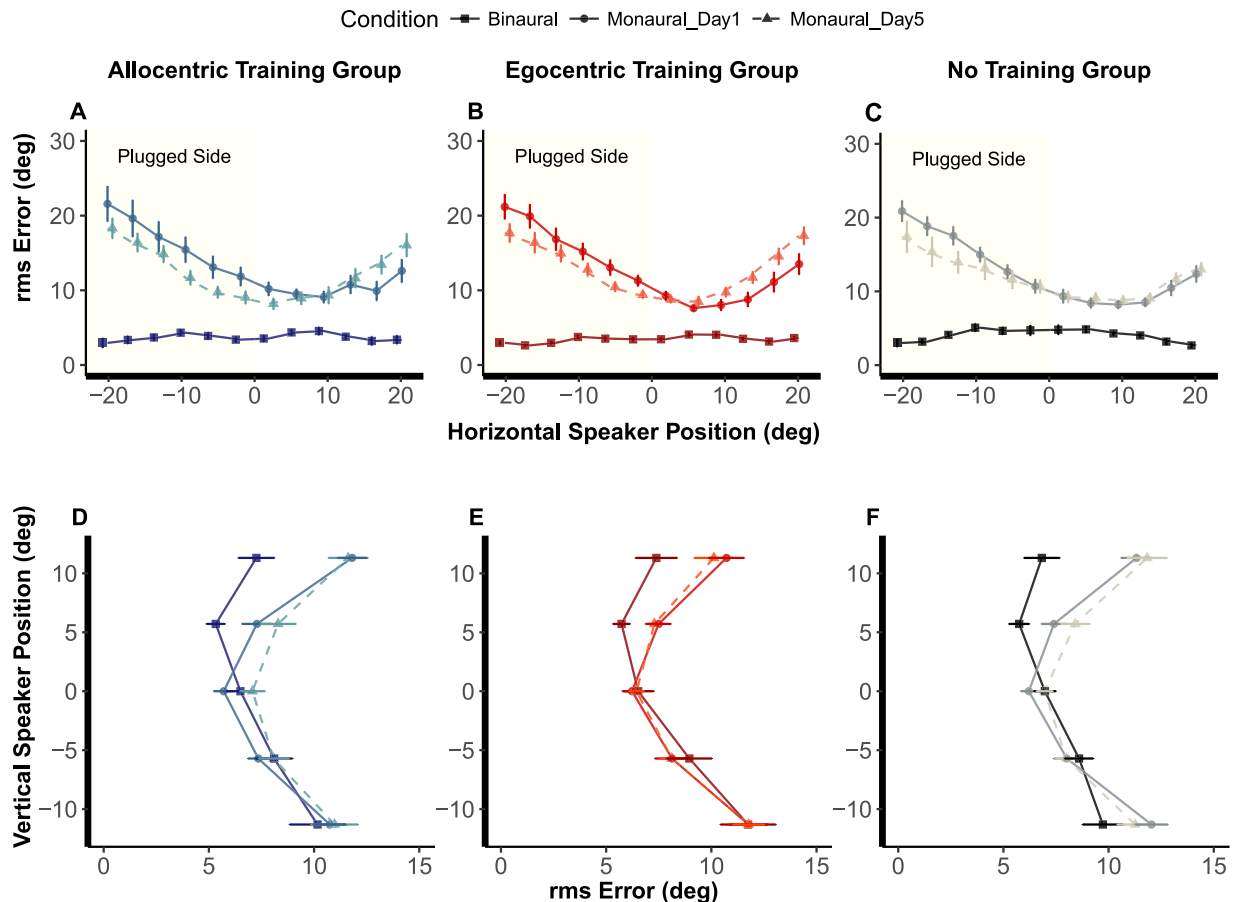
Training occurred from day 2 to day 5 (four sessions of training in total), always with the left ear plugged (see Supplementary Materials). For the AlloT group, training was modelled on the auditory space bisection task proposed by Gori and colleagues<sup>29</sup>. Three consecutive sounds (white noise bursts, 500 ms each) were presented at different horizontal positions, at identical elevation. In each trial, sounds originated from three positions with same elevation but different azimuth. The first sound of the triplet was delivered with equal probability from the speaker located at  $-20^\circ$ ,  $-17^\circ$  or  $-13^\circ$  to the left of the participant, whereas the third sound was always delivered 9 speakers to the right of the first (e.g., first sound at  $-20^\circ$ , third sound at  $13^\circ$ ). Critically, the second sound was delivered at one of the intermediate positions between the first and the third, from immediately adjacent to the first to immediately adjacent to the third. In each trial, participants had to indicate whether the second sound, was closer to the first or to the third one. Responses were collected using the computer keyboard. Our choice to change the position of the first sound (unlike the classic procedure by Gori and collaborators<sup>29</sup>) aimed to reduce the possibility that participants solved the allocentric task through egocentric strategies<sup>32</sup>, paying attention exclusively to the second sound while assuming that the first and third remained at identical positions throughout the experiment. Elevation of sound triplets changed unpredictably across trials (5 possible elevation levels), with each elevation sampled with equal probability within each block. The session was divided in 4 blocks of 60 trials each (total of 240 trials), for an overall duration of about 25 minutes.

In 75% of trials, a white dot (10 pixel diameter) appeared at the time of the auditory stimulation, indicating exactly the location of source location. To emphasise the spatial relation between the sounds a red line (2 pixels width) was presented to connect consecutive sounds (Fig. 1D). In the remaining 25% of trials, sounds were presented without visual stimulation to prevent participants from solving the task only relying on visual stimulation. The percentage of audio-visual stimulation was based on the work of Strelnikov and colleagues<sup>22</sup>.

For the EgoT group, in each trial participants listened to three consecutive sounds (white noise bursts, 500 ms each) exactly as detailed above for the AlloT group. However, after presentation of the last sound (third) a visual instruction appeared on screen indicating which of the three sounds localise (e.g. 'S1', to indicate the first sound). Participants were required to point to target position using the mouse cursor that appeared on the white panel only at the moment of the response. As in the allocentric training, in 75% of trials a white dot indicated the position of each sound (Fig. 1E). No red lines appeared connecting the dots. In the remaining 25% of trials, sounds were presented without visual stimulation. Again, participants received a total of 240 trials per session, divided in 4 blocks of 60 trials each. A single training session lasted approximately 25 minutes.

Three methodological aspects are worth noting. First, the auditory stimulation in both training procedures was completely identical, and the difference between training procedures concerned only the visual stimulation and the task. Although each individual sound likely underwent an initial egocentric coding in both training procedures, critical to the allocentric training task was the spatial comparison between sounds, which entails an allocentric spatial representation. Second, to observe potential generalisation effects of the training paradigms we employed different stimuli and tasks between training and testing sessions, in line previous works<sup>33–35</sup>. Third, unlike most of the previous works<sup>18</sup>, we introduced in the design a control group that did not take part in any training procedure (No Training group, NoT), but performed only testing of sound localisation at day 1, day 3 and day 5.

**Data analysis.** Our stimulation setup allowed the presentation of auditory stimuli both in azimuth and elevation, therefore we have been able to observe different degrees of performance in the two dimensions (see Tables 1S and 2S for cumulative error measurements in all conditions). In the present work, we investigated performance



**Figure 2.** Performance in the sound localisation test. Rms errors for each listening condition (binaural, monaural day 1 and monaural day 5; for clarity performance on day 3 is not shown, see Table 1S and 2S for mean values) are displayed as a function of speaker position, separately for each group, for azimuth (A–C) and elevation (D–F). Thicker axes highlight horizontal speaker position (A–C) and vertical speaker position (D–F), respectively. Error bars show the standard error.

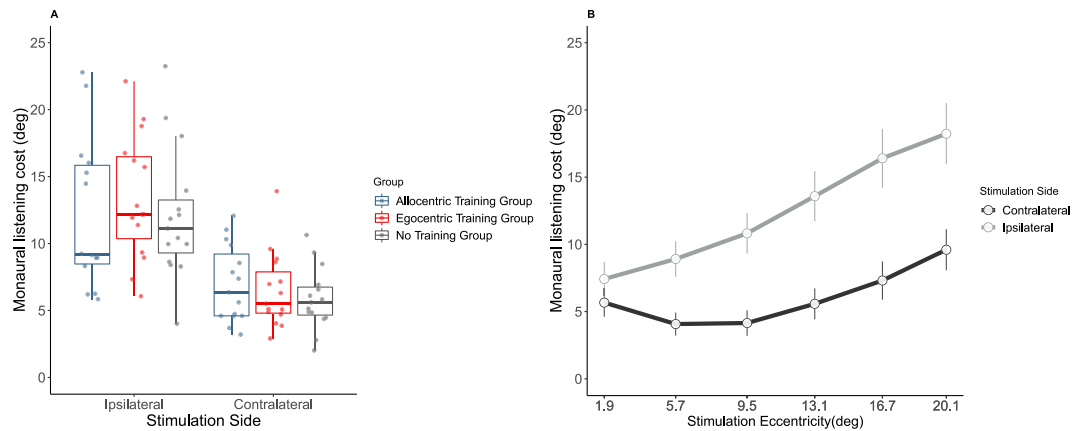
through rms error and signed error, following Hartmann<sup>36</sup> (see also refs<sup>6,37</sup>). The rms error represents the root mean squared difference between speaker position and subject's response. The Signed error represents mean difference between response and correct sound-source position and it could be either positive or negative. When considering azimuth, positive values indicate a rightward bias; when considering elevation, positive values indicate an upward bias. We calculated each error separately for each response component: azimuth and elevation. In case of sphericity assumption violations we adopted the Greenhouse–Geisser method of correction.

## Results

Figure 2 shows rms error at the beginning of the protocol for both binaural and monaural listening conditions (day 1, shown with bold lines) and at the end of the protocol (day 5, shown with dashed lines). Performance is presented separately for each of the three groups as a function of horizontal speaker position (Fig. 2A–C) and vertical speaker position (Fig. 2D,F). In the following sections we describe (1) the ear-plug effects on sound localisation on day 1; (2) performance changes at day 5; (3) performance changes at day 3.

**Ear-plug effects on sound localisation on day 1.** To examine sound localisation performance on day 1, we entered rms errors in a mixed Analysis of Variance (ANOVA) with LISTENING CONDITION (binaural, monaural), STIMULATION SIDE with respect to the plugged ear (ipsilateral, contralateral), and STIMULATION ECCENTRICITY with respect to listener's body midline (2°, 6°, 9°, 13°, 17°, 20°) as within-participant variables, and GROUP (AlloT, EgoT and NoT) as between-participants variable.

This analysis revealed a 2-way interaction between LISTENING CONDITION and STIMULATION SIDE ( $F(1,42) = 44.57, p < 0.001, \eta^2 = 0.69$ ) and a significant 3-way interaction between LISTENING CONDITION, STIMULATION SIDE and STIMULATION ECCENTRICITY ( $F(1.87, 78.46) = 16.69, p < 0.001, \eta^2 = 0.28$ ). None of these interactions was further modulated by GROUP (all  $F$ -values  $< 1$ ). To study these interactions, we calculated the cost of monaural compared to binaural listening (from now on termed '*monaural listening cost*') as the difference in rms error between the two conditions (positive values indicate worse performance in monaural listening). We found that monaural listening cost was overall larger for sounds delivered on the side ipsilateral to the plug ( $12.4^\circ \pm 5.4$ ) compared to the contralateral side ( $5.9^\circ \pm 2.6$ ; see Fig. 3A). Interestingly, monaural listening cost



**Figure 3.** (A) Monaural listening cost in the horizontal plane as a function of Stimulation Side (Ipsilateral/Contralateral to the plug) and experimental group. (B) Monaural listening cost in the horizontal plane, irrespective of group, as a function of Stimulation Eccentricity and Stimulation Side. Error bars represent confidence intervals of the mean.

varied substantially between-participants: some participants were minimally affected by the ear plug, whereas others localised all sounds towards the open ear. Finally, as revealed by the 3-way interaction reported above, monaural listening cost increased as a function of eccentricity, particularly on the side ipsilateral to the plug (see Fig. 3B). The omnibus ANOVA also revealed the main effects of LISTENING CONDITION, STIMULATION SIDE and STIMULATION ECCENTRICITY (all F-values > 30.00), as well as the two-way interactions STIMULATION SIDE\*STIMULATION ECCENTRICITY and LISTENING CONDITION\*STIMULATION ECCENTRICITY (all F-values > 14.00), which were subsidiary to the higher order interactions described above.

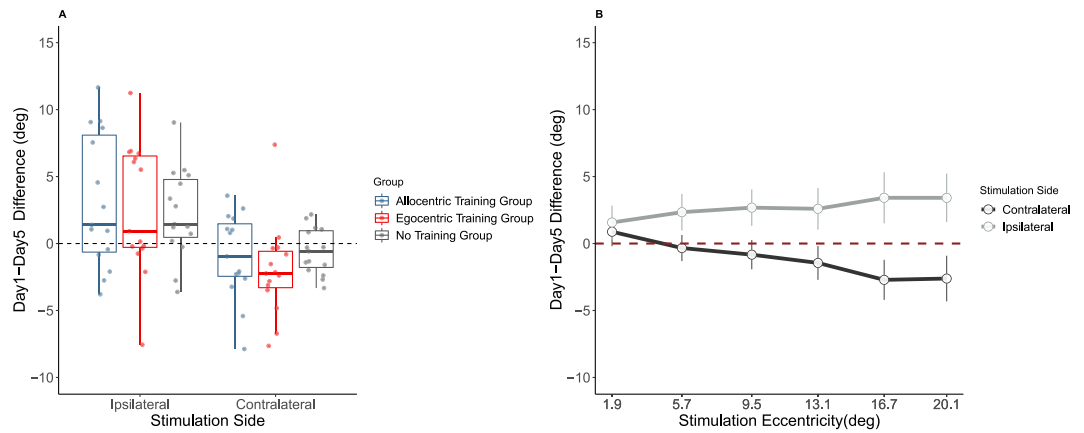
A similar ANOVA on mean signed error was run to examine systematic biases in sound localisation. This analysis revealed again a 3-way interaction between LISTENING CONDITION, STIMULATION SIDE and STIMULATION ECCENTRICITY ( $F(1.87, 78.46) = 16.69$ ,  $p < 0.001$ ,  $\eta^2 = 0.28$ ), which was not modulated by GROUP ( $F(3.74, 78.46) = 0.48$ ,  $p = 0.74$ ,  $\eta^2 = 0.02$ ). To study this interaction, we calculated the bias induced by monaural compared to binaural listening (from now on termed '*monaural listening bias*') as the difference in signed error between the two conditions (positive values indicate a bias towards the side contralateral to the plug). This bias is shown in Supplementary Fig. 3S as a function of eccentricity. It emerged as strong bias to respond towards the hearing side for sounds originating from the space ipsilateral to the plug (average bias:  $13.58 \pm 8.03$ ; for all eccentricities t-test comparisons against zero  $p < 0.001$ ). Interestingly, we also found a mild bias towards the body midline when responding to most eccentric sounds contralateral to the plug ( $-5.31 \pm 4.66$ ,  $-6.99 \pm 5.13$ ,  $-8.39 \pm 5.04$ , for the three most eccentric positions, respectively; all t-test comparisons against zero  $p < 0.001$ ).

In the vertical plane, sound localisation in binaural listening was overall less precise compared to localisation in azimuth (rms error: azimuth =  $3.7^\circ \pm 1.7$ ; elevation =  $7.6^\circ \pm 3.2$ ; paired t-test,  $t(44) = 12.47$ ,  $p < 0.001$ , Cohen's  $d = 1.86$ ). To study whether alteration of binaural cues impacted also on sound elevation perception we entered rms error along the vertical dimension in a mixed ANOVA with LISTENING CONDITION (binaural, monaural), VERTICAL SPEAKER POSITION ( $-11^\circ$ ,  $-6^\circ$ ,  $0^\circ$ ,  $6^\circ$ ,  $11^\circ$ ) as within-participants factors, and GROUP (AlloT, EgoT and NoT) as between-participants factor. This analysis revealed that rms error increased from binaural to monaural listening ( $8.3^\circ \pm 2.0$ ,  $9.3^\circ \pm 1.1$ , respectively;  $F(1, 42) = 14.33$ ,  $p < 0.001$ ,  $\eta^2 = 0.25$ ). Localisation errors increased particularly for sounds above ear level ( $+5^\circ$ :  $5.6^\circ \pm 1.7$  vs.  $7.4^\circ \pm 2.3$ ;  $+11^\circ$ :  $7.1^\circ \pm 3.4$  vs.  $11.3^\circ \pm 2.9$ ), resulting in a significant two-way interaction between LISTENING CONDITION and VERTICAL SPEAKER POSITION ( $F(1.46, 61.33) = 15.95$ ,  $p < 0.001$ ,  $\eta^2 = 0.27$ ). As for the analysis on the horizontal dimension, performance decrement determined by monaural plugging was comparable across groups (no main effect or interaction involving the GROUP factor emerged; all F values < 1).

A similar ANOVA on Signed Error in the vertical dimension, revealed an interaction between LISTENING CONDITION and VERTICAL SPEAKER POSITION ( $F(2.15, 90.31) = 55.98$ ,  $p < 0.001$ ,  $\eta^2 = 0.54$ ). Sounds from above the horizontal midline were localised downward in binaural listening ( $-2.9^\circ \pm 3.7$ ) and even more so in monaural listening ( $-7.3^\circ \pm 3.0$ ;  $t(44) = 7.36$ ,  $p < 0.001$ , Cohen's  $d = 1.1$ ). By contrast, sounds from below the midline were localised upward in both binaural ( $7.7^\circ \pm 4.6$ ) and monaural ( $7.7^\circ \pm 3.3$ ) listening conditions alike ( $t(44) = 0.07$ ,  $p = 0.942$ ).

**Change in performance at day 5.** Changes in sound localisation performance in azimuth and elevation between day 1 and day 5 are shown in Fig. 2, separately for the trained groups and the no-training group (compare solid and dashed lines depicting monaural listening conditions at day 1 and day 5, respectively).

To study the difference in monaural listening performance between the first and the last testing day, we entered rms errors in a mixed Analysis of Variance (ANOVA) with DAY (day 1, day 5), STIMULATION SIDE, STIMULATION ECCENTRICITY and GROUP as before. This analysis revealed a significant 2-way interaction between DAY and STIMULATION SIDE ( $F(1, 42) = 23.82$ ,  $p < 0.001$ ,  $\eta^2 = 0.36$ ), as well as a 3-way interaction between DAY, STIMULATION SIDE and STIMULATION ECCENTRICITY ( $F(2.81, 118.10) = 9.56$ ,  $p < 0.001$ ,  $\eta^2 = 0.18$ ). None of the above interactions was further modulated as a function of group (all F-values < 1).



**Figure 4.** (A) Day1-Day5 Difference in the horizontal plane as a function of Stimulation Side (Ipsilateral/Contralateral to the plug) and experimental group. (B) Day1-Day5 Difference in the horizontal plane irrespective of group as a function of Stimulation Eccentricity and Stimulation Side. Error bars represent confidence intervals of the mean. Positive values indicate performance improvement, negative values performance decrement.

To study these interactions, we calculated performance change as the difference in rms error between day 1 and day 5 (from now on termed '*Day1-Day5 Difference*'). On the side ipsilateral to the plug, the rms error was smaller on day 5 ( $13.54 \pm 4.18$ ) compared to day 1 ( $16.21 \pm 5.31$ ; t-test:  $t(44) = 4.10$ ,  $p < 0.001$ , Cohen's  $d = 0.61$ , corrected for multiple comparison; Day1-Day5 Difference =  $2.67 \pm 4.37$ ). By contrast, on the side contralateral to the plug, the rms error significantly increased from day 1 ( $9.87 \pm 2.44$ ) to day 5 ( $11.05 \pm 2.73$ ; t-test:  $t(44) = 2.73$ ,  $p = 0.02$ , Cohen's  $d = 0.41$ , corrected for multiple comparisons using Bonferroni; Day1-Day5 Difference =  $-1.18 \pm 2.89$ ; see Fig. 4A). Day1-Day5 Difference also changed as a function of eccentricity on both sides, but particularly on the side ipsilateral to the plug (see Fig. 4B). The omnibus ANOVA revealed also the significant main effects of STIMULATION SIDE and STIMULATION ECCENTRICITY (all F-values  $> 30.00$ ) and the two-way interaction STIMULATION SIDE\*STIMULATION ECCENTRICITY ( $F(1.83, 76.73) = 12.56$ ,  $p < 0.001$ ,  $\eta^2 = 0.23$ ), both subsidiary to the higher order interactions described above.

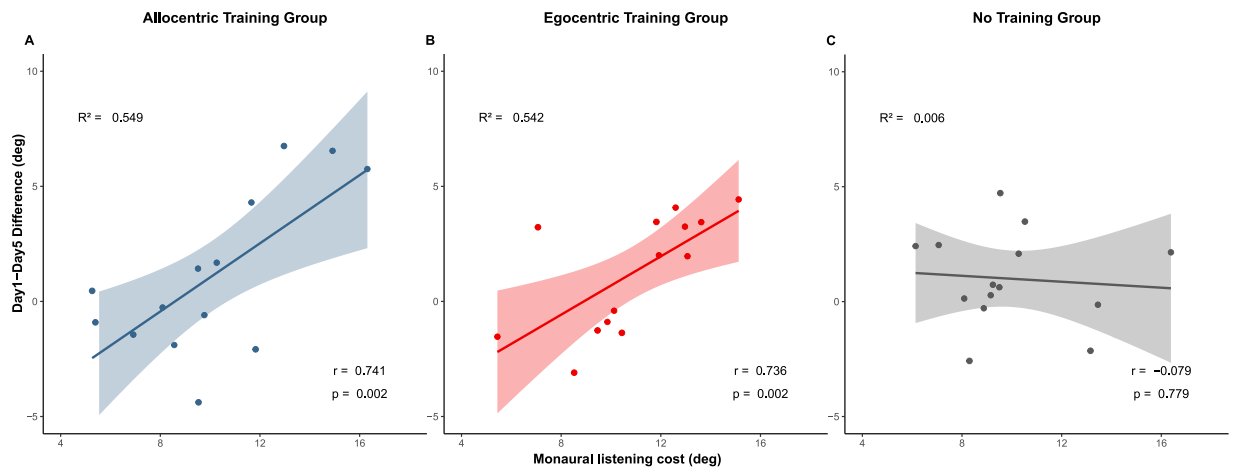
A similar ANOVA performed on mean signed error was conducted to study if any systematic bias changed between day 1 and day 5. This analysis revealed the significant main effects of DAY ( $F(1,42) = 25.5$ ,  $p < 0.001$ ,  $\eta^2 = 0.375$ ), STIMULATION SIDE ( $F(1,42) = 673.21$ ,  $p < 0.001$ ,  $\eta^2 = 0.94$ ), STIMULATION ECCENTRICITY ( $F(4.05, 170.25) = 12.49$ ,  $p < 0.001$ ,  $\eta^2 = 0.22$ ), and the two-way interaction STIMULATION SIDE\*ECCENTRICITY ( $F(1.85, 170.25) = 77.74$ ,  $p < 0.001$ ,  $\eta^2 = 0.92$ ). No higher order interaction emerged (all F-values  $< 1$ ), as well as no main effect or interactions involving the between participants factor GROUP (all F-values  $< 1.1$ ).

Overall, signed error changed between day 1 and day 5 ( $5.24 \pm 5.37$ ,  $2.06 \pm 4.09$ , respectively, t-test:  $t(44) = 5.14$ ,  $p < 0.001$ , Cohen's  $d = 0.77$ ). Similar to monaural listening at day 1, results on monaural listening at day 5 confirmed localisation bias toward the hearing side for sounds ipsilateral to the plug ( $9.99 \pm 7.21$ ), and a bias towards the midline for sounds originating from the space contralateral to the plug ( $-5.89 \pm 7.12$ ).

Interestingly, sound localisation performance in the vertical plane remained unchanged between day 1 and 5 (see lower panels in Fig. 2). A mixed ANOVA on rms error, with DAY (day 1, day 5) and VERTICAL SPEAKER POSITION ( $-11^\circ$ ,  $-6^\circ$ ,  $0^\circ$ ,  $6^\circ$ ,  $11^\circ$ ) as within-participants factors, GROUP (Allot, EgoT and NoT) as between-participants factor, revealed no significant main effect of DAY ( $F(1.527, 61.081) = 1.54$ ,  $p = 0.224$ ,  $\eta^2 = 0.04$ ) or GROUP ( $F(2,40) = 0.39$ ,  $p = 0.68$ ,  $\eta^2 = 0.02$ ), nor any two-way or three way interactions (all F-values  $< 1.4$ ).

**Change in performance at day 3.** Having established that monaural listening performance in rms error changed between the first and the last day of our protocol, we investigated whether this dependent variable changed even after two days, namely at day 3. To this aim, we entered azimuthal rms errors in a mixed Analysis of Variance (ANOVA) with monaural testing DAY (day1, day 3), STIMULATION SIDE with respect to the plugged ear (ipsilateral, contralateral), and STIMULATION ECCENTRICITY with respect to listener's body midline ( $2^\circ$ ,  $6^\circ$ ,  $9^\circ$ ,  $13^\circ$ ,  $17^\circ$ ,  $20^\circ$ ) as within-participant variables, and GROUP (Allot, EgoT and NoT) as between-participants variable (see Tables 1S and 2S for mean values).

This analysis revealed a 2-way interaction between DAY and STIMULATION SIDE ( $F(1,40) = 12.18$ ,  $p = 0.001$ ,  $\eta^2 = 0.23$ ) and the 3-way interaction DAY\*STIMULATION SIDE\*stimulation eccentricity ( $F(2.58, 103.26) = 6.53$ ,  $p < 0.001$ ,  $\eta^2 = 0.14$ ). None of these interactions were modulated by GROUP (all F-values  $< 1$  s). To study these interactions, we calculated localisation performance difference in the horizontal plane between the first (day 1) and the third (day 3) testing day in monaural listening as the difference in rms error between the two testing days (from now on termed '*Day1-Day3 Difference*'). On the ipsilateral side, rms error reduced on day 3 ( $13.27 \pm 4.54$ ) compared to day 1 ( $16.21 \pm 5.31$ ; t-test:  $t(42) = 4.18$ ,  $p < 0.001$ , Cohen's  $d = 0.64$ , corrected for multiple comparisons using Bonferroni; Day1-Day3 Difference =  $2.50 \pm 3.92$ ). On the side contralateral to the plug, however, no change in rms error emerged from day 1 ( $9.87 \pm 2.44$ ) to day 3 ( $10.32 \pm 2.38$ ) (t-test:  $t(42) = 1.03$ ,  $p = 0.62$ , Cohen's  $d = 0.16$ , corrected for multiple comparison; Day1-Day3 Difference =  $0.03 \pm 2.24$ ). Day1-Day3 Difference increased as a function of eccentricity in the plugged side, whereas in the unplugged side reached



**Figure 5.** (A–C) Relation between monaural listening cost and Day1–Day5 Difference as a function of group. The trained groups (AlloT and EgoT, panels A and B, respectively) show a positive and significant correlation between monaural listening cost at day 1 and the performance difference between day 1 and day 5. No such relation emerged for the No Training group (C).

significance only for one eccentricity (see Supplementary Fig. 4S). The main effects of DAY, STIMULATION SIDE and stimulation eccentricity (all F-values > 10.00) as well as the two-way interactions STIMULATION SIDE\* STIMULATION ECCENTRICITY ( $F(1.84, 73.57) = 13.50, p < 0.001, \eta^2 = 0.25$ ) reached significance, but were subsidiary to the higher order interactions described above.

In the vertical plane localisation performance did not change between day 1 and day 3. *Rms error* in the vertical plane was studied using an ANOVA with DAY (day 1, day 3) and VERTICAL SPEAKER POSITION ( $-11^\circ, -6^\circ, 0^\circ, 6^\circ, 11^\circ$ ) as within-participant factor, and GROUP (AlloT, EgoT and NoT) as between-participants factor. The analysis revealed no significant main effects or interaction involving the GROUP variable (all F-values < 1.5).

**Follow-up analyses.** The first follow-up analysis stem from the observation that monaural listening cost in horizontal dimension at day 1 was substantially different across participants (Fig. 3A). We asked whether such variability at day 1 could predict performance changes observed at the end of our protocol (day 5). To this aim, we analysed the relation between *monaural listening cost* at day1 and *Day1–Day5 Difference* using Pearson correlations. Correlation plots are shown in Fig. 5, separately for the three groups. In both training groups a significant relation between monaural listening cost and Day1–Day5 Difference emerged: the more participants decreased their performance in azimuth due to monaural plugging on day 1, the more they changed localisation performance after training (AlloT Group:  $r = 0.74, p = 0.002$ , EgoT Group:  $r = 0.74, p = 0.002$ ). By contrast, Day1–Day5 Difference was unrelated to monaural listening cost in the No Training group ( $r = -0.079, p = 0.78$ ). An ANCOVA with DAY1-DAY5 DIFFERENCE as dependent variable, GROUP (AlloT, EgoT and NoT) as independent variable and MONAURAL LISTENING COST as covariate, showed a significant 2-way interaction between GROUP and MONAURAL LISTENING COST ( $F(2,39) = 5.69, p = 0.007, \eta^2 = 0.14$ ).

The second follow-up analysis stem from the observation that Day1–Day5 Difference on horizontal plane was positive on the plugged side (suggesting performance improvement), but negative on the unplugged side (suggesting performance decrement). One interpretation for this unexpected finding is that recalibration of auditory space between day 1 and 5 emerged as overall shift of responses towards the unplugged side – rather than a re-learning of auditory-cue space correspondences. If this is the case, Day1–Day5 Difference on the plugged side should correlate with Day1–Day5 Difference in the unplugged side across participants. We tested this hypothesis using correlations (Pearson) and found no relation between ipsilateral improvement and contralateral decrease, neither when treating all participants together ( $r = 0.03, p = 0.86$ ) nor when splitting them as a function of group (AlloT:  $r = 0.27, p = 0.34$ ; EgoT:  $r = -0.15, p = 0.58$ ; NoT:  $r = -0.07, p = 0.080$ ).

## Discussion

In the present study, we examined the relation between allocentric and egocentric reference frames in spatial hearing, in the context of multisensory guided training of sound localisation. First, we confirmed that the method adopted for auditory cues alteration (monaural ear plugging) can effectively impair sound localisation abilities in the horizontal dimension. In addition, we extended this classic finding by showing that sound localisation impairments emerged also in the vertical dimension, and by documenting substantial inter-participants variability in the monaural listening cost determined by the ear-plug. Second, after four days of protocol (day 5), we measured performance improvements on single sound localisation on the side ipsilateral to the plug (left), which emerged together with concomitant performance decrements on contralateral space (right). Notably, these performance changes emerged only in the horizontal dimension, whereas monaural sound localisation in the vertical dimension remained entirely unaffected. Third, we found that performance changes occurred with comparable strength in all tested groups – i.e., irrespective of whether participants were trained or untrained. From a theoretical perspective, this finding implies that listeners can change their sound localisation performance even in absence of error feedback signal. From a methodological point of view, it stresses the importance of including a



no training group when assessing the efficacy of sound localisation training methods. Finally, we proposed a new audio-visual protocol based on allocentric coding of sound, extending the principles of multisensory stimulation previously implemented only using egocentric procedures. In the following sections, we will discuss each of these novel contributions in details.

In agreement with previous findings we show performance decrement from binaural to monaural listening (ear-plugging), mostly in the space ipsilateral to the plug. However, the impact of monaural plugging on sound localisation performance was markedly different between participants. Inter-individual variability of the consequences of ear plugging has been reported previously<sup>38</sup>, but its origin remains to be ascertained. It could reflect inter-individual efficacy of the plug, different sensitivity to auditory cues alteration, or a combination of the two. An interesting finding of the present study is the impact of this inter-individual variability of the plug-effect (which we termed ‘monaural listening cost’) on performance change after four days of training (which we termed ‘Day1-Day5 Difference’). Although sample size of each experimental group was limited, it is noteworthy that in our two trained groups the magnitude of training effects was linearly related with the amount of sound localisation deficit induced by the plug (see Fig. 5A–C). Participants who were more affected by the ear plug changed their performance post-training to a greater extent, compared to participants who were minimally affected by monaural occlusion. Intriguingly, this positive relation was completely absent in the control group – which just repeated the test after 2 and 4 days. Given the limited number of participants in each group ( $N = 15$ ), these correlation results should be taken with caution. Yet, positive relations between earplug effect and training benefit have previously been reported in humans with similar sample sizes<sup>18</sup> (but see ref.<sup>39</sup> for contrasting results). In particular, in a study on spatial hearing re-learning after monaural ear-moulding, Van Wanrooij and Van Opstal<sup>18</sup> showed that participants whose spectral cues were disrupted the most achieved better post-training results ( $N = 11$ ), and learned at faster rates compared to participants who were minimally affected by monaural ear-moulding. The functional model proposed by Van Opstal (2016) assumes that a new correspondence between auditory cues and space is stored whenever a discrepancy is detected between the spatial outcome of an existing sound-space correspondence (SSC) and the actual stimulus position, leading to the consolidation of new SSC with repeated exposures. In this context, the observed relation between monaural listening cost and Day-Day5 Difference might be interpreted as the consequence of the perceived discrepancy between the response and the actual sound position. This discrepancy was only experienced during training and was clearly largest for those participants who were maximally affected by the plug. Indeed, no such relation between ear-plug effect and change in performance between day 1 and day 5 was detected in the no-training group (which never received feedback).

This finding suggests that qualitative differences might exist between the trained groups and the no-training group, despite the quantitatively comparable performance. Performance changes in individuals with altered auditory cues who received no error-feedback signal has been recently documented using other methods. Zonooz and colleagues<sup>40</sup> tested sound localisation in the vertical plane using spectrally poor (band limited) sounds in trained vs. untrained participants. While performance improvements were larger in participants trained with visual feedback, even those participants who did not receive performance-related feedback showed sound-localisation improvements in vertical plane. Performance changes in the no-training group are compatible with the theoretical framework provided by Keating and King<sup>15</sup>. The authors proposed that adaptation to altered sound localisation cues can result from two different (though not mutually exclusive) mechanisms: cue-reweighting and cue-remapping. Cue-reweighting is a fast adaptation process, in which the relative contribution (i.e., weight) of the available auditory cues is changed to accommodate the intervening modulations (e.g., ear-plugging, as here). Instead, cue-remapping is a slowly emerging adaptation phenomenon in which a new correspondence between available auditory cues and space is progressively encoded. The observed differences between trained and untrained participants may relate to these two different adaptation mechanisms. Specifically, the no-training group may have primarily changed its performance exploiting a fast cue-reweighting mechanism. By contrast, the sensory feedback provided in the audio-visual trainings may have triggered a slower process of sensory recalibration, in which new sound-space correspondences started to emerge from a combination of cue-reweighting and cue-remapping processes. This suggests that longer and more consistent training procedure may be essential to achieve proper cue-relearning. In this sense, it is possible that our short training (25 minutes per sessions, for 4 consecutive days) captured only the intermediate phase of the entire adaptation process, in which cue-reweighting – but not a full and stable cue-remapping mechanism – has been accomplished. This could also partially explain the combination of ipsilateral improvement and contralateral decrement measured in the trained group.

The present study also introduced a novel training approach based on audio-visual coding of sounds in *allocentric* coordinates. Previous research exploring the effects of training in the context of reversibly altered auditory cues almost exclusively adopted an egocentric frame of reference (for a review see ref.<sup>16</sup>). One notable exception is a study conducted by Gori and colleagues<sup>41</sup>, conducted with blindfolded hearing participants tested in binaural listening. The procedure entailed an auditory spatial bisection task similar to the one adopted here, as well as training sessions aimed at improving allocentric coding using multisensory feedback. Specifically, the experimental set-up comprised nine speakers, each associated with one vibrotactile stimulator placed on the forearm of the participant. Speakers and vibrotactile stimulators were kept spatially aligned throughout the task. The tactile feedback group listened to triplets of acoustic stimuli (as in the present work), paired with delayed (200 ms) spatially-congruent tactile stimulation. During training, participants were required to attend the multisensory stimulation but performed a different task, which consisted in detecting an occasional higher tone. Performance in the auditory bisection task improved in the tactile feedback group, but not in participants who received only a verbal feedback about sound position, or participants who received no feedback.

While this audio-tactile study highlights the potentials of multisensory stimulation in promoting allocentric auditory space perception, two differences with the current work are noteworthy. First participants were not explicitly required to perform an allocentric coding of sounds during training, unlike our audio-visual allocentric

task. Second, in Gori *et al.*<sup>41</sup> stimuli and task remained identical throughout the experiment, limiting our understanding of generalisation of training effects. Conversely, our study introduces, for the first time, an allocentric training procedure based on *relative* position of multiple sounds for the purpose of changing auditory space mapping. Importantly, the structure of our paradigm, with differences between training and test sessions both in terms of stimulation (spoken syllables, rather than white noise bursts) and task (localisation of a single stimulus, rather than egocentric/allocentric processing of triplets of sounds) allows assessing learning generalisation effects. Generalisation effects have strong implications for translational application of training procedures to applied or clinical settings, and are thus considered a key requirement when assessing the efficacy of training procedures<sup>33,34,42,43</sup>. Our novel allocentric training procedure also extends to allocentric coding the multisensory training principles<sup>44–47</sup> for the purpose of acoustic space re-learning. We provided participants a visual counterpart not only of the position of the sounds, but also of the spatial relationship between them. The rationale for adding lines between the dots (Fig. 1D) was to emphasise visuo-spatial encoding of the acoustic scene. In addition, we reasoned that they could promote the use of visuo-spatial mental imagery<sup>48</sup> during the response phase, in agreement with the evidence that vision can facilitate allocentric representations of space<sup>49</sup>. We acknowledge, however, that the effectiveness of the added lines connecting the dots should not be taken for granted, as a proper control condition is missing in the present paradigm.

In conclusion, our work constitutes a first attempt to address a largely unexplored aspect of sound localisation abilities – namely the interaction between egocentric and allocentric representations. Addressing this domain is important because it could potentially introduce a novel approach to auditory spatial training. In vision, studies on egocentric vs. allocentric space coding have shown partial dissociation<sup>50–55</sup> as well as interactions<sup>56–60</sup> between these representations. Likewise, integration and interaction between different reference frames has been proposed to be necessary for an efficient tactile localisation<sup>61</sup>. To the best of our knowledge, our study provides the first investigation aimed at assessing such interaction between reference frames in the auditory domain. It has been recently proposed that egocentric and allocentric representations may be linked in a hierarchical manner, with egocentric processing involving only a subpart of the wider neural network subtending allocentric processing<sup>62</sup>. If this is the case, exploiting allocentric training could foster space re-learning to a greater extent and prove more effective compared to training based on egocentric space representations. Testing this hypothesis was beyond the scopes of the present work, and it implies testing the efficacy of egocentric and allocentric training on sound localisation tests conducted on both reference frames (i.e., egocentric and allocentric). Yet, this follow-up study could exploit the current findings to further understand the contribution of allocentric auditory space processing in sound localisation re-learning.

## Data Availability

A public data repository is available at the following link: <https://osf.io/rczbg/>.

## References

- Bregman, A. S. *Auditory Scene Analysis*. (MIT Press, 1990).
- Bregman, A. S. *Auditory scene analysis: the perceptual organization of sound*. (MIT Press, 1990).
- Middlebrooks, J. C. & Simon, J. Z. Ear and Brain Mechanisms for Parsing the Auditory Scene. in *The Auditory System at the Cocktail Party* 1–6, [https://doi.org/10.1007/978-3-319-51662-2\\_1](https://doi.org/10.1007/978-3-319-51662-2_1) (Springer, Cham, 2017).
- Yost, W. A. Spatial release from masking based on binaural processing for up to six maskers. *Cit. J. Acoust. Soc. Am.* **141**, 1985 (2017).
- Neupane, K. R., Woyke, J. & Wilde, J. Effect of Initial Strength of Honey Bee Colonies (*Apis mellifera*) Supered in Different Ways on Maximizing Honey Production in Nepal. *Journal of Apicultural Science* **56**, (Boston University, 2014).
- Pavani, F. *et al.* Spatial and non-spatial multisensory cueing in unilateral cochlear implant users. *Hear. Res.* **344**, 24–37 (2017).
- Middlebrooks, J. C. & Green, D. M. Sound Localization by Human Listeners. *Annu. Rev. Psychol.* **42**, 135–159 (1991).
- Middlebrooks, J. C. *Sound localization. Handbook of Clinical Neurology* **129**, (Elsevier B.V., 2015).
- Blauert, J. *Spatial hearing: the psychophysics of human sound localization*. (MIT Press, 1997).
- Town, S. M., Brimijoin, W. O. & Bizley, J. K. Egocentric and allocentric representations in auditory cortex. *PLoS Biol.* **15**, e2001878 (2017).
- Schechtman, E., Shrem, T. & Deouell, L. Y. Spatial Localization of Auditory Stimuli in Human Auditory Cortex is Based on Both Head-Independent and Head-Centered Coordinate Systems. *J. Neurosci.* **32**, 13501–13509 (2012).
- Gori, M., Sandini, G., Martinoli, C. & Burr, D. C. Impairment of auditory spatial localization in congenitally blind human subjects. *Brain* **137**, 288–293 (2014).
- Wright, B. A. & Zhang, Y. A review of learning with normal and altered sound-localization cues in human adults. *Int. J. Audiol.* **45**, 92–98 (2006).
- Carlile, S. The plastic ear and perceptual relearning in auditory spatial perception. *Front. Neurosci.* **8**, 1–13 (2014).
- Keating, P. & King, A. J. Sound localization in a changing world. *Curr. Opin. Neurobiol.* **35**, 35–43 (2015).
- Mendonça, C. A review on auditory space adaptations to altered head-related cues. *Front. Neurosci.* **8**, 1–14 (2014).
- Shinn-Cunningham, B. G., Durlach, N. I. & Held, R. M. Adapting to supernormal auditory localization cues. I. Bias and resolution. *J. Acoust. Soc. Am.* **103**, 3656 (1998).
- Van Wanrooij, M. M. & Van Opstal, A. J. Relearning Sound Localization with a New Ear. *J. Neurosci.* **25**, 5413–5424 (2005).
- Kacelnik, O., Nodal, F. R., Parsons, C. H. & King, A. J. Training-Induced Plasticity of Auditory Localization in Adult Mammals. *PLoS Biol.* **4**, e71 (2006).
- Ponton, C. W. *et al.* Plasticity in the adult human central auditory system: evidence from late-onset profound unilateral deafness. *Hear. Res.* **154**, 32–44 (2001).
- Shams, L., Wozny, D. R., Kim, R. & Seitz, A. Influences of multisensory experience on subsequent unisensory processing. *Front. Psychol.* **2**, 1–9 (2011).
- Strelnikov, K., Rosito, M. & Barone, P. Effect of audiovisual training on monaural spatial hearing in horizontal plane. *PLoS One* **6**, 1–9 (2011).
- Isaiah, A., Vongpaisal, T., King, A. J. & Hartley, D. E. H. Multisensory Training Improves Auditory Spatial Processing following Bilateral Cochlear Implantation. *J. Neurosci.* **34**, 11119–30 (2014).
- Majdak, P., Walder, T. & Laback, B. Effect of long-term training on sound localization performance with spectrally warped and band-limited head-related transfer functions. *J. Acoust. Soc. Am.* **134**, 2148–2159 (2013).
- Flannery, R. & Butler, R. A. Spectral cues provided by the pinna for monaural localization in the horizontal plane. *Perception & Psychophysics* **29** (1981).

26. Musicant, A. D. & Butler, R. A. The influence of pinnae-based spectral cues on sound localization. *Cit. J. Acoust. Soc. Am.* **75**, 1648 (1984).
27. Slattery, W. H. & Middlebrooks, J. C. Monaural sound localization: Acute versus chronic unilateral impairment. *Hear. Res.* **75**, 38–46 (1994).
28. Musicant, A. D. & Butler, R. A. *Monaural localization: An analysis of practice effects.* *Perception & Psychophysics* **28** (1980).
29. Gori, M., Sandini, G. & Burr, D. Development of Visuo-Auditory Integration in Space and Time. *Front. Integr. Neurosci.* **6**, 77 (2012).
30. Doucet, M.-E. *et al.* Blind subjects process auditory spectral cues more efficiently than sighted individuals. *Exp. Brain Res.* **160**, 194–202 (2005).
31. Van Wanrooij, M. M. & Van Opstal, A. J. Sound Localization Under Perturbed Binaural Hearing. *J. Neurophysiol.* **97**, 715–726 (2007).
32. Filimon, F. Are All Spatial Reference Frames Egocentric? Reinterpreting Evidence for Allocentric, Object-Centered, or World-Centered Reference Frames. *Front. Hum. Neurosci.* **9**, 648 (2015).
33. Wright, B. A. & Zhang, Y. A review of the generalization of auditory learning. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **364**, 301–11 (2009).
34. Mendonça, C., Campos, G., Dias, P. & Santos, J. A. Learning Auditory Space: Generalization and Long-Term Effects. *PLoS One* **8**, 1–14 (2013).
35. Watson, C. J. G., Carlile, S., Kelly, H. & Balachandar, K. The Generalization of Auditory Accommodation to Altered Spectral Cues. *Sci. Rep.* **7**, 11588 (2017).
36. Hartmann, W. M. Localization of sound in rooms. *J. Acoust. Soc. Am.* **74**, 1380–1391 (1983).
37. Grantham, W., Ashmead, D., Ricketts, T. A., Haynes, D. S. & Labadie, R. F. Horizontal-Plane Localization of Noise and Speech Signals by Postlingually Deafened Adults Fitted With Bilateral Cochlear Implants. *Ear Hear.* **28**, 524–541 (2007).
38. Kumpik, D. P., Kacelnik, O. & King, A. J. Adaptive reweighting of auditory localization cues in response to chronic unilateral earplugging in humans. *J. Neurosci.* **30**, 4883–94 (2010).
39. Trapeau, R. & Schönwiesner, M. Adaptation to shifted interaural time differences changes encoding of sound location in human auditory cortex. *Neuroimage* **118**, 26–38 (2015).
40. Zonooz, B., Arani, E. & Opstal, A. J. Van. Learning to localise weakly-informative sound spectra with and without feedback. *Sci. Rep.* **8**, 17933 (2018).
41. Gori, M., Vercillo, T., Sandini, G. & Burr, D. Tactile feedback improves auditory spatial localization. *Front. Psychol.* **5**, 1–7 (2014).
42. Mendonça, C. *et al.* On the Improvement of Localization Accuracy with Non-Individualized HRTF-Based Sounds. *J. Audio Eng. Soc.* **60**, 821–830 (2012).
43. Wright, B. A., Sabin, A. T., Zhang, Y., Marrone, N. & Fitzgerald, M. B. Enhancing Perceptual Learning by Combining Practice with Periods of Additional Sensory Stimulation. *J. Neurosci.* **30**, 12868–12877 (2010).
44. Lehmann, S. & Murray, M. M. The role of multisensory memories in unisensory object discrimination. *Cogn. Brain Res.* **24**, 326–334 (2005).
45. Lovelace, C. T., Stein, B. E. & Wallace, M. T. An irrelevant light enhances auditory detection in humans: a psychophysical analysis of multisensory integration in stimulus detection. *Cogn. Brain Res.* **17**, 447–453 (2003).
46. Kim, R. S., Seitz, A. R. & Shams, L. Benefits of stimulus congruency for multisensory facilitation of visual learning. *PLoS One* **3** (2008).
47. Shams, L. & Seitz, A. R. Benefits of multisensory learning. *Trends Cogn. Sci.* **12**, 411–417 (2008).
48. Cornoldi, C. & Vecchi, T. *Visuo-spatial Working Memory and Individual Differences*, <https://doi.org/10.4324/9780203641583> (Psychology Press, 2004).
49. Pasqualotto, A., Spiller, M. J., Jansari, A. S. & Proulx, M. J. Visual experience facilitates allocentric spatial representation. *Behav. Brain Res.* **236**, 175–179 (2013).
50. Feigenbaum, J. D. & Morris, R. G. Allocentric Versus Egocentric Spatial Memory After Unilateral Temporal Lobectomy in Humans. *Neuropsychology* **18**, 462 (2004).
51. Grimsen, C., Hildebrandt, H. & Fahle, M. Dissociation of egocentric and allocentric coding of space in visual search after right middle cerebral artery stroke. *Neuropsychologia* **46**, 902–914 (2008).
52. Iachini, T., Ruggiero, G. & Ruotolo, F. Does blindness affect egocentric and allocentric frames of reference in small and large scale spaces? *Behav. Brain Res.* **273**, 73–81 (2014).
53. Kesner, R. P., Farnsworth, G. & DiMattia, B. V. Double dissociation of egocentric and allocentric space following medial prefrontal and parietal cortex lesions in the rat. *Behav. Neurosci.* **103**, 956–961 (1989).
54. Pizzamiglio, L., Guariglia, C. & Cosentino, T. Evidence for Separate Allocentric and Egocentric Space Processing in Neglect Patients. *Cortex* **34**, 719–730 (1998).
55. Ruggiero, G., Ruotolo, F. & Iachini, T. The role of vision in egocentric and allocentric spatial frames of reference. *Cogn. Process.* **10**, 719–730 (2009).
56. Burgess, N. Spatial memory: how egocentric and allocentric combine. *Trends Cogn. Sci.* **10**, 551–557 (2006).
57. Fink, G. R. *et al.* Performing allocentric visuospatial judgments with induced distortion of the egocentric reference frame: an fMRI study with clinical implications. *Neuroimage* **20**, 1505–1517 (2003).
58. Klatzky, R. L. Allocentric and Egocentric Spatial Representations: Definitions, Distinctions, and Interconnections. in *Spatial cognition* 1–17, [https://doi.org/10.1007/3-540-69342-4\\_1](https://doi.org/10.1007/3-540-69342-4_1) (Springer, Berlin, Heidelberg, 1998).
59. Neggers, S. F. W., Schölvinc, M. L. & van der Lubbe, R. H. J. Quantifying the interactions between allo- and egocentric representations of space. *Acta Psychol. (Amst.)* **118**, 25–45 (2005).
60. Neggers, S. F. W., der Lubbe, V., Ramsey, R. H. J. & Postma, N. F. A. Interactions between ego- and allocentric neuronal representations of space. *Neuroimage* **31**, 320–331 (2006).
61. Heed, T., Buchholz, V. N., Engel, A. K. & Röder, B. Tactile remapping: from coordinate transformation to integration in sensorimotor processing. *Trends Cogn. Sci.* **19**, 251–258 (2015).
62. Zaehle, T. *et al.* The neural basis of the egocentric and allocentric spatial frame of reference. *Brain Res.* **1137**, 92–103 (2007).

## Acknowledgements

We thank Stefano Targher for having built our stimulation set-up and for his help during the implementation of the experiment, together with Massimo Vescovi. Francesco Pavani was supported by a grant of the French Agence Nationale de la Recherche (ANR-16-CE17-0016, VIRTUALHEARING3D) and by a prize of the Fondation Medisite (France).

## Author Contributions

G.R. and F.P. designed the study. G.R. implemented the paradigm. G.R. and E.A. collected and analysed the data. G.R., F.P. and E.A. interpreted the data. G.R. and F.P. wrote the manuscript. We are grateful to Prof. Tobias Heed and one anonymous reviewer for their thoughtful and constructive comments on a previous version of this manuscript.

## Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-019-44267-3>.

**Competing Interests:** The authors declare no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019