



HAL
open science

Spatialisation auditive : de l'innovation en neurosciences aux perspectives cliniques

Valérie Gaveau

► **To cite this version:**

Valérie Gaveau. Spatialisation auditive : de l'innovation en neurosciences aux perspectives cliniques. Sciences cognitives. UCBL, 2024. tel-04757415

HAL Id: tel-04757415

<https://hal.science/tel-04757415v1>

Submitted on 28 Oct 2024

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.

Public Domain

Thèse d'Habilitation à Diriger des Recherches - Université Lyon 1

Numéro d'ordre XXX-XXX

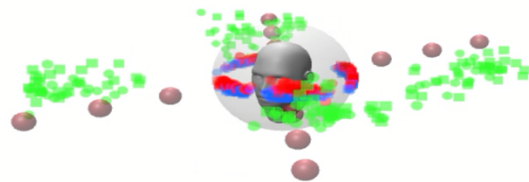
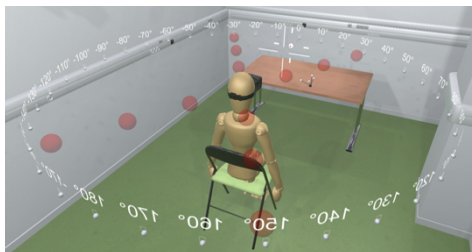
présenté par

GAVEAU Valérie

pour l'obtention du diplôme

D'Habilitation à Diriger des Recherches

Spatialisation auditive : de l'innovation en neurosciences aux perspectives cliniques



À Romeo & Matteo

Remerciements

Les travaux présentés dans cet ouvrage découlent d'une recherche menée sur une période de huit ans depuis le début de ma carrière en tant qu'enseignante-chercheuse à l'Université de Lyon1. Ces réalisations sont le fruit de collaborations précieuses avec des individus dotés d'une valeur scientifique et humaine exceptionnelle, englobant des chercheurs, des étudiants, des ingénieurs, des cliniciens, des administratifs et des patients. La recherche, telle que je la conçois, ne peut se concrétiser sans une collaboration étroite, où chaque personne apporte son expertise pour le bénéfice collectif, progressant dans la même direction tout en respectant les différences de chacun.

Cependant, mes activités de recherche et d'enseignement ont débuté bien avant mon accession au statut d'enseignante-chercheuse. Ma reconnaissance envers les personnes qui ont croisé ma route et qui ont été sources d'inspiration va au-delà de cette limite temporelle arbitraire, remontant à plus de vingt ans.

Je tiens à remercier ma gratitude envers tous les étudiants qui m'ont stimulée en me permettant de les guider dans leur parcours de recherche. Mes remerciements s'étendent également à ceux qui ont partagé et partagent avec moi les plaisirs de l'enseignement universitaire, à ceux qui ont contribué à créer une atmosphère de travail unique au sein de l'équipe Impact, à ceux qui ont cru et continuent de croire en moi. De nombreux noms me viennent à l'esprit, sans ordre particulier : Roméo, Claude, Anne-Emmanuelle, Jean-Louis, Eric K., Eric T., Sonia, Aurélie, Francesco, Alessandro, Anaël, Fred, Chiara, Rosella, Anna-Lisa, Clément, Laure, Jacintha, Annabelle, Karen, Fadila, Michel, Grégoire, Denis, Yuji, Dominique, Laura, Pascal, Nadine, Arleen, Jeanne-Laure, Clara, Muriel, Elisa, Martine, Leslie, Julie, Yves, Jacques... et les autres !

TABLE DES MATIERES

TABLE DES MATIERES	1
CURRICULUM VITAE	2
1. ACTIVITES DE RECHERCHE PASSEES	16
2. ACTIVITES DE RECHERCHE PRESENTES	19
Généralités sur la SA	19
CHAPITRE 1. Étudier les performances de localisation auditive : prérequis méthodologiques.	23
1. Prérequis n°1 : Étudier la spatialisation auditive dans les 3 dimensions de l'espace	24
☞ Contribution= une nouvelle approche pour l'étude de la SA (Gaveau et al. 2020)	26
☞ Contribution= Validation de l'utilisation d'un système de RV pour les études neuroscientifiques (Verdelet et al. 2020)	27
2. Prérequis n°2 : importance du contrôle du regard et du contrôle des informations visuelles	28
☞ Contribution= Contrôle de la posture d'écoute (Gaveau et al., 2022)	29
☞ Contribution= Impact de la scène visuelle dans les performances de SA (Valzolgher et al., 2020)	31
3. Prérequis n°3 : Favoriser l'écoute active	32
☞ Contribution= l'écoute active favorise les performances de SA (Coudert et al., 2021 ; Gaveau et al., 2022)	33
4. Prérequis n°4 : Favoriser les interactions avec l'espace	35
☞ Contribution= interagir dans son espace acoustique favorise la re-spatialisation auditive (Valzolgher et al., 2020)	35
CHAPITRE 2. Évaluation et rééducation de la SA & déficits auditifs	39
1. Performance de SA des patients porteurs d'implants cochléaires	40
☞ Contribution= Évaluation des capacités de localisation auditive des patients porteurs de deux prothèses implantables auditives dans les 3 dimensions de l'espace, bénéfice de l'écoute active (Coudert et al., 2021)	41
2. Contribution= Modèle d'entraînement à la spatialisation auditive suite à une altération de l'écoute binaurale et généralisation à l'espace sonore (Valzolgher et al., 2022)	41
3. Contribution= Entraînement à la spatialisation auditive de patients implantés cochléaires (Valzolgher et al. 2023a, 2023b)	43
4. Contribution= Entraînement intensif à la spatialisation auditive de patients implantés cochléaires, et généralisation à d'autres fonctions cognitives (Coudert et al. 2022)	43
5. Contribution= Évaluation des capacités spatiale auditive de patients implantés cochléaires par un questionnaire à passation rapide (Coudert et al. 2023)	44
3. PERSPECTIVES DE RECHERCHE	47
1. Impact du positionnement des microphones des implants cochléaires dans les performances de SA. Projet MicLo.	47
2. Re-spatialiser les sons chez les enfants implantés cochléaires avec une approche ludique. Projet Kid Train.	52
3. Impact du traumatisme sonore sur les capacités de SA chez les militaires traumatisés sonores. Projet SPHEAR.	56
4. Spatialisation auditive en réalité virtuelles chez des patients avec un premier épisode psychotique. Projet LocIZ.	58
CONCLUSION & REFLEXIONS PERSONNELLES	61
BIBLIOGRAPHIE	63
ANNEXES	70

Curriculum Vitae

DDN : 09/12/1971, 52 ans
Nationalité : Française
Adresse : Equipe IMPACT du Centre de Recherche en Neurosciences de Lyon (CRNL)
16 Avenue du Doyen Lépine- 69500 Bron
valerie.gaveau@univ-lyon1.fr

Institut des Sciences et Techniques de la Réadaptation (ISTR)
Université Claude Bernard de Lyon (UCBL), Faculté de Médecine Lyon Est
8 Av. Rockefeller, 69008 Lyon
valerie.gaveau@univ-lyon1.fr

Formations initiales

2013-2014 DIU "Formation des investigateurs aux essais cliniques des médicaments". UCBL
2001-2005 Doctorat en Neurosciences. Inserm U864 (bourse McDonnell Pwee). UCBL
1994 DEA de Neurosciences. UCBL
1993 Maîtrise de Biologie des Organismes et Populations. Univ. Franche-Comté
1992 Licence de Biologie des Organismes et Populations. Univ. Franche-Comté

Expériences professionnelles

2021- Maître de Conférences Hors Classe– ISTR- UCBL
2015-2020 Maître de Conférences échelon 04– ISTR- UCBL
2015 Ingénieur en Recherche clinique - Hôpital Neurologique P. Wertheimer- Lyon
2013 Chercheuse Post-doctorante - Hôpital H. Gabrielle- Lyon
2010-2012 Chercheuse Post-doctorante – IMPACT - CRNL- Lyon
2007-2010 Chercheuse Post-doctorante - INSERM U864- Lyon
2006-2007 Chercheuse Post-doctorante - Hôpital H. Gabrielle - Lyon
2005-2006 Chercheuse Post-doctorante - Département de Physiologie humaine de Bologne (Italie).
2000-2001 Assistant ingénieur- Hôpital H. Gabrielle, laboratoire de Kinésigraphie- Lyon.
1999-2000 Assistant ingénieur - Institut des Sciences Cognitives de Bron - Lyon
1997-1999 Assistant ingénieur - Neurobiologie Préclinique du C.H.S. Le Vinatier- Lyon.

Stages professionnels et formations professionnelles

2015,2021	Formation aux 'Bonnes pratiques cliniques' - Hospices Civiles de Lyon
2014	Attachée de recherche clinique - Pôle IMER, Hôpital Lyon Sud.
2011	Formation 'Introduction aux statistiques', Inserm ADR 05.
2010	Formation à l'utilisation en autonomie de l'IRM 1.5T du CERMEP sur primates vigiles
2009	Journées de Formation en Neuro-Imagerie « IRM fonctionnelle : une introduction »
2004	Ecole d'été du Collège de France « Neurophysiologie intégrative et computationnelle »
1995	Formation Expérimentation Animale Niveau I, UCBL

Publications scientifiques

📧 Journaux internationaux

1^{er} auteur = 12, *dernier auteur* = 3, *co-auteur* = 17. *H-Index* = 17, *i10-Index* = 22.

Italique = étudiant encadré.

1. Valzolgher C., Alzaher M., **Gaveau V.**, Coudert A., Marx M., Truy E., Barone, P., Farne A., Pavani F. Capturing visual attention with perturbed auditory spatial cues. *Trends in Hearing*. 2023 Jan-Dec;27:23312165231182289. doi: 10.1177/23312165231182289. PMID: 37611181.
2. Coudert A, Reilly KT, Truy E, Moulin A, **Gaveau V.** A new tool for the assessment of speech understanding and spatial hearing difficulties in children: the Kid-SSQ questionnaire. *Eur Arch Otorhinolaryngol*. 2023 Feb 10. doi: 10.1007/s00405-023-07846-9. Epub ahead of print. PMID: 36763152.
3. Valzolgher C, Bouzaid S, Grenouillet S, Gatel J, Ratenet L, Murenu F, Verdelet G, Salemme R, **Gaveau V.**, Coudert A, Hermann R, Truy E, Farnè A, Pavani F. Training spatial hearing in unilateral cochlear implant users through reaching to sounds in virtual reality. *Eur Arch Otorhinolaryngol*. 2023 Mar 11. doi: 10.1007/s00405-023-07886-1
4. Coudert A, Verdelet G, Reilly KT, Truy E, **Gaveau V.** Intensive Training of Spatial Hearing Promotes Auditory Abilities of Bilateral Cochlear Implant Adults: A Pilot Study. *Ear Hear*. 2023 Jan-Feb 01;44(1):61-76. doi: 10.1097/AUD.0000000000001256.
5. Valzolgher C, Gatel J, Bouzaid S, Grenouillet S, Todeschini M, Verdelet G, Salemme R, **Gaveau V.**, Truy E, Farnè A, Pavani F. Reaching to Sounds Improves Spatial Hearing in Bilateral Cochlear Implant Users. *Ear Hear*. 2023 Jan-Feb 01;44(1):189-198. doi: 10.1097/AUD.0000000000001267. Epub 2022 Aug 19. PMID: 35982520.
6. **Gaveau V.**, Coudert A, Salemme R, Koun E, Desoche C, Truy E, Farnè A, Pavani F. Benefits of active listening during 3D sound localization. *Exp Brain Res*. 2022 Nov;240(11):2817-2833. doi: 10.1007/s00221-022-06456-x. Epub 2022 Sep 7.
7. Valzolgher C, Todeschini M, Verdelet G, Gatel J, Salemme R, **Gaveau V.**, Truy E, Farnè A, Pavani F. Adapting to altered auditory cues: Generalization from manual reaching to head pointing. *PLoS One*. 2022 Apr 14;17(4):e0263509. doi: 10.1371/journal.pone.0263509.
8. Coudert A., **Gaveau V.**, et al. Spatial hearing difficulties in bilateral cochlear implant children improve with head movements. *Ear and Hearing*. *Ear Hear*. 2022 Jan/Feb;43(1):192-205. doi: 10.1097/AUD.0000000000001090

9. Valzolgher C, Verdelet G, Salemme R, Lombardi L, **Gaveau V.**, et al. Reaching to sounds in virtual reality: A multisensory-motor approach to promote adaptation to altered auditory cues. *Neuropsychologia*. 2020 Dec; 149:107665.
10. Valzolgher C., Alzhaler M., Gessa E., Todeschini M., Nieto P., Verdelet G., Salemme S, **Gaveau V.**, et al. The Impact of a Visual Spatial Frame on Real Sound-Source Localization in Virtual Reality. *Current Research in Behavioral Sciences* 1. Nov. 2020. 100003.
11. Priot AE, Revol P, Sillan O, Prablanc C, **Gaveau V.** Sensory Prediction of Limb Movement Is Critical for Automatic Online Control. *Front Hum Neurosci*. 2020 Oct 7; 14:549537.
12. Sarrazin V., Kandel M., V. **Gaveau V.**, et al. P267 Cerebellar Inhibition Disrupts Prism Adaptation by Impairing Feedforward Error Correction. *Clinical Neurophysiology* 131, n° 4. April 2020.: e165.
13. Verdelet, G., Salemme R., Desoche C., Volland F., Farne A., *Coudert A.*, Hermann R., Truy E., **Gaveau V.** and Pavani F. Assessing Spatial and Temporal Reliability of the Vive System as a Tool for Naturalistic Behavioural Research. In *2019 International Conference on 3D Immersion (IC3D)*, 1-8. Brussels, Belgium: IEEE, 2019.
14. Mikula, L. **Gaveau V.**, et al. Learned Rather than Online Relative Weighting of Visual-Proprioceptive Sensory Cues. *Journal of Neurophysiology* 119, n° 5. 2018: 1981-92. doi: 10.1152/jn.00338.2017. Epub 2018 Feb 21.
15. **Gaveau V.**, et al. Paradoxical adaptation of successful movements: The crucial role of internal error signals. *Conscious Cogn*. 2018 Sep; 64:135-145.
16. Luauté J, Dubois A, Heine L, *Guironnet C*, Juliat A, **Gaveau V.**, et al. Electrodermal reactivity to emotional stimuli in healthy subjects and patients with disorders of consciousness. *Ann Phys Rehabil Med*. 2018 Nov;61(6):401-406. doi: 10.1016/j.rehab.2018.04.007. Epub 2018 May 18.
17. Panouillères MT, **Gaveau V.**, et al. Oculomotor Adaptation Elicited By Intra-Saccadic Visual Stimulation : Time-Course of Efficient Visual Target Perturbation. *Front Hum Neurosci*. 2016 Mar 9 ; 10:91.
18. Pisella L., **Gaveau V.**, et al. Separate Pathways for Automatic and Intentional Visuo-Manual Reach Transformations: New Perspectives for Hemiparesia Rehabilitation Using Moving Objects. *Int J Phys Med Rehabil* 3:5, 2015
19. **Gaveau V.**, et al. Visuomotor adaptation needs a validation of prediction error by feedback error. *Front Hum Neurosci*. Nov 4 :8-880, 2014
20. **Gaveau V.**, et al. Automatic online control of motor adjustments in reaching and grasping. *Neuropsychologia*. Mar 55:25-40, 2014
21. O'Shea J, **Gaveau V.**, et al. Kinematic markers dissociate error correction from sensorimotor realignment during prism adaptation. *Neuropsychologia*. Mars 55:15-24, 2014
22. Panouillères M., **Gaveau V.**, et al. Our brain adapts to what it sees during eye movements. *PLoS One*.8(1), 2013
23. Monfardini E., **Gaveau V.**, et al. Social learning as a way to overcome choice-induced preferences? Insights from humans and rhesus macaques. *Frontiers in Neuroscience*, 6:127, 201
24. Laurence Havé., **Gaveau V.**, Luauté J. Interest of automatic motivity involvement in hemiplegic patients rehabilitation. *Annals of Physical and Rehabilitation Medicine* 54, October 2011
25. Gerardin P., **Gaveau V.**, et al. Integration of Visual Information for Saccade Production. *Human Movement Science*, 30(6):1009-21, 2011.
26. Blangero A, **Gaveau V.**, et al., A hand and a field effect in on-line motor control in unilateral optic ataxia. *Cortex*. May 44(5):560-8, 2008.
27. **Gaveau V.**, et al. Saccadic control and eye-hand coordination in optic ataxia. *Neuropsychologia*, 46:475-486, 2008
28. **Gaveau V.**, et al. Self-generated saccades do not modify gain of adapted reactive saccades. *Experimental Brain Research*, 162: 526–531, 2005
29. **Gaveau V.**, et al., On-Line Modification of Saccadic Eye Movements By Retinal Signals. *Neuroreport*, 14, 2003

30. Desmurget M., **Gaveau V.**, et al. On-line motor control in patients with Parkinson's disease. *Brain*, 127, 2004.
31. Ravel N., Chabaud P., Martin C., **Gaveau V.**, et al. Olfactory learning modifies the expression of odour-induced oscillatory responses in the gamma (60-90 Hz) and beta (15-40 Hz) bands in the rat olfactory bulb. *European Journal of Neuroscience*, 17, 2003
32. Karoumi B., Saoud M., d'Amato.T., Rosenfeld F., Denis P., Gutknecht C., **Gaveau V.**, et al. Poor performance in smooth pursuit and antisaccadic eye-movement tasks in healthy siblings of patients with schizophrenia. *Psychiatry Research*, 101, 2001

☞ Ouvrages publiés et notes scientifiques (en anglais)

1. Pisella L., Striemer C., Blangero A., **Gaveau V.**, et al. Perceptual deficits in optic ataxia? In Rossetti Y, Haggard P. & Kawato M eds. *Sensorimotor foundations of higher cognition: Attention & Performance XXII*. Oxford New York: Oxford University Press, 47-72, 2007
2. **Gaveau V.**, et al. From semantic analogy to theoretical confusion? Commentary on "The emulation theory of representation: motor control, imagery, and perception" by Grush. *Behavioral and Brain Sciences*, 27(3): 404, 2004
3. **Gaveau V.** et Desmurget D. Do movement Planning and control represent independent modules? Commentary on "Separate visual representations in the planning and control of action" by Glover. *Behavioral and Brain Sciences*, 27(1):35, 2004
4. **Gaveau V.**, et al. Eye-Hand Coordination in Reaching Movements. *The Handbook of Brain Theory and Neural Networks 2ed.*, MIT Press, 2003.

☞ Ouvrages publiés et notes scientifiques (en français)

1. Pavani F., Farnè A., Truy E. and **Gaveau V.** Perception auditive spatiale. In *Surdit , actualit s, innovations et espoirs- Perception auditive*. Elsevier-Masson, de Truy E., Lescanne E., Loundon N., Roman S. 2018.
2. Barone P., Pavani F., Farn  A., Truy E., **Gaveau V.**, et al. Int gration multisensorielle chez le patient d ficient auditif. In *Surdit , actualit s, innovations et espoirs- Perception auditive*. Elsevier-Masson, de Truy E., Lescanne E., Loundon N., Roman S. 2018.
3. **Gaveau V.**, et al. R le des r gulations automatiques du contr le moteur : mis en place d'un protocole de r education des h mipar sies. *Kin sith rapie Scientifique*, 542 : 5-12, 2013

☞ Communications Internationales orale et s minaires

1. **Gaveau V.** A novel approach to study sound localization: SPHERE. *7th European Congress of NeuroRehabilitation (France)*. 2023
2. **Gaveau V.**, Coudert A. SPHERE, a new tool for evaluation and rehabilitation of spatial hearing. *NeuroFrance International Meeting (France)*. 2023
3. **Gaveau V.** Virtual reality as a tool for naturalistic approach of auditory spatial perception. *14th Stereopsia (Belgique)*. 2022
4. Coudert A., **Gaveau V.** Truy E. Sound perception in 3D space: strategies and deficits in adults with cochlear implants. *4th European Symposium of Paediatric Cochlear Implantation (Hongrie)*. 2019
5. **Gaveau V.** Cartographie 3D et entrainement multisensori-moteur pour l'audition spatiale. *Entretiens Jacques Cartier (Canada)*. 2017
6. **Gaveau V.**, et al. Fast visual eye control deficit in optic ataxia. *Insitute of advanced studies of Bologna University (Italie)*. 2006
7. **Gaveau V.**, et al. On-line modification of saccadic eye movements by retinal signals. *2nd Gainfield Conference (Allemagne)*. 2002

8. **Gaveau V.**, et al. On-line control of movement: psychophysical approaches. *Mc Donnell Pwe foundation (Espagne)*. 2003

☞ **Communications Internationales écrites avec publication des actes et comité de sélection**

1. **Gaveau V**, **Coudert A.** et al. SPHERE: a Novel Approach to 3d and Active Sound Localization. A naturalistic approach to spatial hearing studies. SIGGRAPH (Canada). 2022
2. **Coudert A.**, **Gaveau V.** Truy E. Sound perception in 3D space: strategies and deficits in adults with cochlear implants. 14th European Symposium of Paediatric Cochlear Implantation (Roumanie). 2019.
3. **Gaveau V.** Pavani F., Coudert A., et al. Active sound localisation in 3D space in unilateral and bilateral cochlear implant users: a new approach to spatial hearing assessment through immersive virtual reality. 10th International Symposium on Objective Measures in Auditory Implants. (Israël). 2018
4. Pavani F., **Gaveau V.**, et al. Active sound localisation in 3D space in unilateral and bilateral cochlear implant users: a new approach to spatial hearing assessment through immersive virtual reality. International Federation of Oto-Rhino-Laryngological Societies. (France). 2017
5. Panouillères M., **Gaveau V.**, et al. Our brain sees and learns during fast eye movements. *The Neural Control of Movement Society (Italie)*. 2012
6. Panouillères M., **Gaveau V.**, et al. Our brain sees and learns during fast eye movements. *8TH FENS Forum of Neuroscience. (Spain)*. 2012
7. **Gaveau V.**, et al. Movement automatic guidance by direct internal model. *The Neural Control of Movement Society (Italie)*. 2012
8. Kennedy J., Pastor D., **Gaveau V.**, et al. Dual processes in prism adaptation - Evidence from computational modelling and transcranial direct current stimulation. *The Neural Control of Movement Soc. (Italie)*. 2012
9. Hadj-bouziane F., Monfardini E., **Gaveau V.**, et al. Left, right, but not both: A monkey fMRI study of the neuronal bases of visual competition. *Society for Neuroscience, (USA)*. 2011
10. **Gaveau V.**, et al. Visual aftereffects transfer to saccadic touch localisation. *International Congress of Neuro-ophthalmology. (France)*. 2010
11. **Gaveau V.**, et al. Functional link between Action (saccade) and tactile coding of finger position is modified by the adaptation of visually-guided saccade. *Attention and Awareness, Rovereto. (Italie)*. 2009
12. Blangero A., **Gaveau V.**, et al. Parietal contribution in on-line control of visually-guided movements. *Progress in Motor Contro. (France)*. 2009
13. Gerardin P., **Gaveau V.**, et al. Reactive Saccades Dynamics: Visual Integration and Visual Context. *Vision Sciences Society Annual Meeting (USA)*, 2008
14. Pisella L., Blangero A., **Gaveau V.**, et al. Optic ataxia, a pure visuo-manual deficit ?. *ESF Research Conference on "3D Sensory and Motor Space", (Spain)*, 2007
15. Gaveau V. Oculo-manual coordination and motor control. Workshop on Visuomotor Coordination and Control. (France). 2007
16. Pisella L., Blangero A., **Gaveau V.**, et al. Effects of lesions of medial extrastriate cortex in humans. *7th IBRO World Congress of Neurosciences, (Australie)*, 2007
17. **Gaveau V.**, et al. Control of corrective saccade: limits of visual updating. *The Neural Control of Movement Society, (Espagne)*, 2007.
18. Blangero A., **Gaveau V.**, et al. Impaired fast update of visuo-spatial representation for both action and perception in optic ataxia ?. *European Societies of Neuropsychology, (France)*, 2006
19. Blangero A, **Gaveau V.**, et al. Visuo-spatial and visuo-manual deficit in optic ataxia revealed by online control motor performance. *'Attention & Performance XXII', (France)*, 2006
20. Xie J, Thobois S, Desmurget M, **Gaveau V.**, et al. On-line movement guidance in patients with Parkinson's Disease. 8th Inter. Congress of Parkinson's Disease & movement disorders, (Italie), 2004
21. Revol P., **Gaveau V.**, et al. The quick visual guidance of hand pointing and internal models. *6th World Congress of Neuroscience (IBRO), (Tchécoslovaquie)*, 2003

22. **Gaveau V.**, et al. Saccadic suppression only affects consciousness. *Euroconferences and EBBS workshop on Cognitive and Neural Bases of Visuomotor Control, Lalonde, et 3^{ème} Forum Européen des Neurosciences, (France). 2002*
23. **Gaveau V.**, et al. Saccadic online correction. *2nd Gainfield Conference 'Space coding in primates', (Allemagne), 2002.*
24. **Gaveau V.** et Joseph J-P. Modulation of FEF & LIP neuronal activities by visual attention in a target selection task. *2nd Meeting of European Neuroscience Association, (France). 1996.*

☞ Communications Nationales écrites avec publication des actes et comité de sélection

1. **Gaveau V.**, Desoche C. Le handicap visuel à travers l'immersion en réalité virtuelle. *Société Française de Simulation en Santé, Lyon, 2022*
2. Havé L., Priot A-E., Luauté J., Delporte, L., Revol. P, **Gaveau V.** The contribution of automatic online motor control in a rehabilitation program for hemiplegic patients. Conference NeuroFrance 2021
3. Coudert A., **Gaveau V.**, et al. Rôle des mouvements de tête dans la perception spatiale des sons chez les patients implantés cochléaires ». OTOForum, Grenoble. 2021
4. **Gaveau V.**, et al. Erreur terminale et erreur de prédiction dans l'adaptation prismatique. *11^{ème} Colloque de la Société des Neurosciences Françaises, Lyon-Grenoble 2013*
5. **Gaveau V.** Rôle des régulations automatiques du contrôle moteur : mise en place d'un protocole de rééducation. *15^{ème} J. Nat. Hautevilloise Rech. & l'étude en Kinésithérapie, Hauteville-Lompnès, 2013*
6. **Gaveau V.**, et al. Comprendre et réduire le handicap moteur : Le rôle des régulations automatiques de l'action. *25^{ème} congrès de la Société Française de Médecine Physique et de Réadaptation, Marseille, 2010*
7. Bosco A., Breveglieri R., **Gaveau V.**, et al. Movimenti di raggiungimento di un bersaglio nello spazio 3-D: studio cinematografico sul macaco sveglio. *Società Italiana di Fisiologia, Siena, 2007*
8. Munuera J., **Gaveau V.**, et al. Contrôle en ligne et coordination œil-main lors des mouvements d'atteinte. *'Voir, agir, communiquer : déficit et handicap', Lille 3, 2004*
9. **Gaveau V.**, et al. Modèles internes prédictifs et mécanismes de rétrocontrôles du geste. *14^{ème} Rencontres Régionales de la Recherche en Rhône-Alpes, Saint-Etienne, 2003*
10. Cavézian C., Desmurget M., **Gaveau V.**, et al. Etude de la coordination visuo-manuelle chez des patients schizophrènes et des patients parkinsoniens : ajustement en ligne du mouvement. *6^{ème} Colloque de la Société des Neurosciences Françaises, Rouen 2003*
11. Gervais R., Hudry J., **Gaveau V.**, et al. Mise en évidence de réponses oscillatoires aux odeurs dans les noyaux amygdaliens chez des patients épileptiques cérébro-implantés. *5^{ème} Colloque de la Société des Neurosciences Françaises, Toulouse 2001*
12. Ravel N., Chabaud P., **Gaveau V.**, et al. Modulation des activités oscillatoires de type bêta (15-40 Hz) et gamma (60-90 Hz) dans le système olfactif du rat au cours d'un apprentissage de type Go-NoGo. *5^{ème} Colloque de la Société des Neurosciences Françaises, Toulouse 2001*

☞ Communications orales Nationales

1. Coudert A., Percher M., **Gaveau V.** Spatialisation auditive et aides auditives : évaluation et rééducation. 1ère Journée scientifique de l'ISTR. Lyon. 2023
2. **Gaveau V.** Virtual Reality as a tool for naturalistic approach of perception. EdTech Lyon. CRNL 2022
3. Coudert A., **Gaveau V.** and Truy E. Sound localization in virtual reality in cochlear implant users. Journée annuelle de l'école doctorale en Neurosciences et Cognition de Lyon (NSCo). 2021.
4. Coudert A., Gaveau V., et al. Capacités de localisation spatiale des enfants bi-IC : une déficience en 3D. Congrès annuel de la Société Française d'ORL. Paris. 2020
5. Coudert A., Gaveau V. et al. Perception des sons en 3D chez des enfants bi-implantés cochléaires. 15ème Congrès de la SFA. Lyon. 2020
6. **Gaveau V.**, Coudert A. SPHERE. Spatialisation auditive : nouvelle approche par la Réalité Virtuelle. Séminaire Interne INSERM 1093. Dijon, 2020

7. Coudert A., **Gaveau V.** and Truy E. Rôle du mouvement de la tête dans la perception spatiale du son. Journée de l'Association Française d'Otoneurologie. Dijon. 2019
8. **Gaveau V.** Exploration de la perception spatiale auditive en réalité virtuelle, vers une nouvelle approche multisensorielle en rééducation. 18ème journée nationale de rééducation de l'AHREK. Valence. 2017
9. **Gaveau V.**, et al. Rééducation de l'hémiplégie par un protocole stimulant "le pilote automatique". Colloque interne Hôpital H. Gabrielle. Lyon. 2009.
10. **Gaveau V.**, et al. Stimulation de la régulation automatique de la motricité. Colloque 'Rééducation et motricité automatique', Hôpital H. Gabrielle, Lyon. 2007
11. Blangero A., **Gaveau V.**, et al. Déficiences visuo-spatiales et visuo-manuelles dans l'ataxie optique révélées par les performances de contrôle moteur en ligne. Rencontres Scientifiques de l'IFNL, Lyon. 2006
12. **Gaveau V.**, et al. Le contrôle moteur : approche psychophysique et comportementale chez l'Homme et le singe. 'Neurophysiologie intégrative et computationnelle', Collège de France, (Concarneau). 2004

☞ **Note scientifique (en français)**

1. **Gaveau V.**, et al. Rôle des régulations automatiques du contrôle moteur : mise en place d'un protocole de rééducation des hémiplégies. *Kinésithérapie Scientifique*. 541 : 59. 2013
2. **Gaveau V.**, et al. Mise en évidence de la cellule Ba insuline chez le poisson zebra brachydanio rerio. *Annales d'endocrinologie – SPIFF Masson Service*, 1994

Financements de mes projets de recherche

Total = 1 128 870 € obtenus en 8 ans dont 565 492 € en tant qu'investigatrice principale.

- | | |
|-----------|--|
| 2016-2018 | Appel d'offre : ITMO NNP (Neurosciences, Neurologie et Psychiatrie). « La neuro-immersion : une nouvelle approche à l'exploration du cerveau et à la rééducation des déficits neurologiques »
Montant forfaitaire reçu = 15 K€
Position : Co-investigatrice. |
| 2016-2020 | Appel d'offre : ANR-16-CE17-0016. « VirtualHearing 3D. Auditory space and multisensory attention in cochlear implant patients »
Montant reçu = 414 971 €
Position : Co-investigatrice |
| 2017 | Appel d'offre : LabEx CORTEX SATT (ANR-11-LABX-0042). « Spatial hearing rehabilitation through action in 3D multisensory space, SPHERE: SPatial HEaring Rehabilitation »
Montant forfaitaire reçu = 20 K€
Position : Investigatrice principale |
| 2017 | Appel d'offre : PICS CNRS. « INSB PICS Italie »
Montant forfaitaire reçu = 8 K€
Position : Co-Investigatrice |
| 2018 | Appel d'offre : LabEx CORTEX SATT (ANR-11-LABX-0042). « SPHERIC: SPatial HEaring Rehabilitation In Children »
Montant forfaitaire reçu = 20 K€
Position : Investigatrice principale |
| 2020 | Appel d'offre : Pack Ambition Recherche région AURA. « Rééducation de l'espace auditif par l'action dans un espace multisensoriel »
Montant reçu = 0 K€ (projet sélectionné au 1er tour, notes finales de 20/20, 18/20 et 12/20) |

- 2020 Position : Investigatrice principale
Appel à Manifestation d'intérêt (AMI) via PULSAlys (SATT Lyon). « SPHEREApp : Evaluation, rééducation et accompagnement technologique de l'espace auditif, accompagnement au développement d'un démonstrateur numérique »
Montant reçu = 60 K€
- 2021 Position : Investigatrice principale
Appel à projet : Institut rhônalpin des Systèmes complexes (IXX). « Spatialisation auditive et rééducation orthophonique : évaluer et entraîner l'audition spatiale pour améliorer les capacités linguistiques des enfants sourds appareillés »
Montant reçu = 5 K€
- 2018-2021 Position : Investigatrice principale
Partenariat industriel : CRNL- HCL & Cochlear ®. « SPHERIC IIR-1948 »
Montant reçu = 113 400 €
- 2023-2026 Position : Co-investigatrice
Partenariat industriel : CRNL- HCL & Advanced-Bionics. « Kid-Train CT69HCL 22_0599 »
Montant reçu = 107 110 €
- 2024-2027 Position : Co-investigatrice
Appel à projet du Centre Hospitalier Le Vinatier. « LOCIZ : Spatialisation auditive en réalité virtuelle chez des patients avec un premier épisode psychotique »
Montant reçu = 17 730 €
- 2024-2028 Position : Co-investigatrice
Doctoral Networks Call : HORIZON-MSCA-2022-DN-01-01 « Cochlear implants and spatial hearing: Enabling access to the next dimension of hearing- CherISH ».
Montant reçu = 565 387 €
Position : Leader France

Activités de valorisation scientifique

- « SPHERE: Method and apparatus for acquiring a spatial map of auditory perception of a subject ». Brevet européen WO2017203028A1; Publication EP 3463084 A1 20190410 (EN)
Inventeurs : Belle A., Farnè A., **Gaveau V.**, Koun E., Pavani F., Salemme R.
Demandeur : CNRS, INSERM, UCBL, Université de Saint Etienne.
- « SENSE-I : Système Evolutif pour les NeuroSciences Expérimentales – Immervives »
Demande en cours à l'APP (Agence pour la Protection des Programmes, Europe)
Inventeurs : Verdelet G., Salemme R., **Gaveau V.**, Coudert A., Pavani F.
Demandeur : PULSAlys (SATT de Lyon), INSERM, UCBL, Université de Saint Etienne, Université de Trento.
- « SPHEREapp : usage autonome du dispositif SPHERE par les professionnels de l'audition »
Investissement en maturation (60K). 2022
Partenaires : PULSAlys (SATT de Lyon), INSERM, UCBL (**Gaveau V.**), Univ. Saint Etienne.

4. « HERVé- Handicap & Expérience en Réalité Virtuelle pour les Étudiants »
Dépôt de subvention auprès de la Région Auvergne Rhône Alpes ‘Innover pour compenser le handicap’
pour le développement d’une application (79 K€). 2023
Partenaire privé : UniVR Studio (SAS, Villeurbanne).
Position : Co-investigatrice

Direction de recherche (Master 2 & doctorat)

On notera les étudiants en Master 2 (Master Santé, Orthophonie, Kinésithérapie, etc.) que j’ai encadrés possèdent un diplôme d’état en rééducation ; ils ont fait le choix, pour une très grande majorité, de ne pas continuer en thèse et d’entrer dans la vie active en tant que professionnels de rééducation avec un grade de Master.

- 2015-2016 Master 2 Santé Handicaps Sensoriels et Cognitif - ISTR-UCBL
« Perception tridimensionnelle de l’espace acoustique chez les patients implantés cochléaires »
Stagiaire : J-L Eeckhout (Psychomotricienne DE)
Encadrement : 100%
- 2015-2016 Master 2 Santé Handicaps Sensoriels et Cognitif - ISTR-UCBL
« Perception tridimensionnelle de l’espace acoustique chez les sujets normo-entendants : rôle des déplacements céphaliques »
Stagiaire : T. Salteur de la Serraz (Kinésithérapeute DE)
Encadrement : 100%
- 2016-2017 Master 2 Santé Handicaps Sensoriels et Cognitif - ISTR UCBL
« Effets de la perception de stimulations sensorielles préférées sur la cognition des sujets sains »
Stagiaire : C. Guironnet (Ergothérapeute DE)
Encadrement : 100%
→ Publication : Ann Phys Rehabil Med. 2018 Nov; 61(6):401-406.
- 2017-2018 Master 2 Neurosciences Neurosciences Fondamentales et Clinique - UCBL
« Sound Perception in 3D space: strategies and deficits in adults with cochlear implant »
Stagiaire : Coudert A (Formation médicale)
Encadrement : 100%
- 2018-2019 Master 2 Santé Handicaps Sensoriels et Cognitif - ISTR-UCBL
« Recherche de marqueurs de la conscience chez des patients en éveil de coma »
Stagiaire : A-N. Heizmann
Encadrement : 50%
- 2019-2020 Master 2 Orthophonie - ISTR UCBL
« KID SSQ : questionnaire de localisation spatiale pour enfants implantés cochléaires »
Stagiaire : L. Fombertasse

Encadrement : 50%
→ Publication BU Lyon 1 - Mémoire N° 2035

- 2019-2020 Master 2 Orthophonie - ISTR UCBL
« Rééducation en réalité virtuelle de l'espace auditif des patients bi-implantés cochléaires adultes : étude de faisabilité »
Stagiaire : L. Bonnamy
Encadrement : 50%
→ Publication BU Lyon 1 - Mémoire N° 2036
- 2020-2021 Master 2 Santé Handicaps Sensoriels et Cognitif - ISTR UCBL
« Évaluation de l'effet d'un protocole d'Entraînement Visio-Auditif sur l'audition spatiale chez les sujets ayant une perte auditive, cas des patients implantés cochléaires unilatéraux »
Stagiaire : S. Bouzaid (Audioprothésiste)
Encadrement : 100%
→ Publication : Eur Arch Otorhinolaryngol. Mar 11 2023.
→ Prix recherche
- 2020-2021 Master 2 Handicap neurologique- Université Paris Sarclay
« EVA : Évaluation de l'effet d'un protocole d'Entraînement Visio-Auditif sur l'audition spatiale chez les sujets ayant une perte auditive, cas des patients implantés cochléaires bilatéraux »
Stagiaire : S. Grenouillet (Kinésithérapeute)
Encadrement : 100%
→ Publication : Eur Arch Otorhinolaryngol. 2023
- 2019-2021 Doctorat de Neurosciences. 2^{ème} année - UCBL
« Évaluation fondamentale des performances de localisation sonore de patients implantés cochléaires et mise en place d'une rééducation ciblée de la localisation dans cette population »
Doctorante : A. Coudert
Encadrement : 50%
→ Publication : Trends in Hearing 2023; Eur Arch Otorhinolaryngol. 2023; Ear Hearing 2023; Exp.Br. res 2022; Ear Hearing 2022
→ Brevet: SENSE-i
→ Conférences: 8
- 2021-2022 Master 2 kinésithérapie - ISTR UCBL
« Les effets fonctionnels des différentes techniques de rééducation de la négligence spatiale unilatérale évalués par l'échelle Catherine Bergego, une revue systématique de littérature »
Stagiaire : R. Heraud
Encadrement : 100%
→ Publication BU Lyon 1

- 2022-2023 Master 2 Santé Handicaps Sensoriels et Cognitif - ISTR UCBL
 « Perception tridimensionnelle de l'espace acoustique chez les adultes traumatisés sonores des Armées. SPHEAR »
 Stagiaire : C. Rolle
 Encadrement : 100%
 → Publication BU Lyon 1
- 2022-2023 Master 2 Orthophonie - ISTR UCBL
 « Remédiation des habiletés sociales des adolescents de 12 à 20 ans présentant un Trouble du Spectre Autistique et technologies immersives (telles que la Réalité Virtuelle) »
 Stagiaire : L. Moron-Grand-Dufay
 Encadrement : 50%
 → Publication BU Lyon 1 - Mémoire N°2372
- 2022-2023 Master 2 kinésithérapie - ISTR UCBL
 « Apport de protocoles de rééducation basés sur la réalité virtuelle non immersive associée à la rééducation conventionnelle dans la prise en charge de l'équilibre des enfants paralysés cérébraux, une revue systématique de littérature »
 Stagiaire : N. Giorgis
 Encadrement : 100%
 → Publication BU Lyon 1
- 2022-2023 Master 2 kinésithérapie - ISTR UCBL
 « Effets d'une rééducation par les nouvelles technologies de la négligence spatiale unilatérale chez les patients post-AVC, une revue systématique de littérature »
 Stagiaire : M. Lopez
 Encadrement : 100%
 → Publication BU Lyon 1
- 2023-2024 Master 2 kinésithérapie déficient visuel - Les PEP69 UCBL
 « Évolution de la prise en soin kinésithérapique des troubles de la sensibilité de la main: une revue de la littérature »
 Stagiaire : D. Mbaye
 Encadrement : 100%
- 2023-2024 Master 2 Santé Handicaps Sensoriels et Cognitif - ISTR UCBL
 « Perception 3D de l'espace acoustique chez les adultes traumatisés sonores des Armées. SPHEAR »
 Stagiaire : A. Bernaudat
 Encadrement : 100%

Activités pédagogique - Enseignements

2023-	Co-Responsable. Cours de Spatialisation auditive. ISTR - UCBL. Licence*
2023-	Co-Responsable. Cours de Réalité virtuelle en innovation en Santé. ISTR - UCBL. Master 2*
2022-	Co-Responsable. Cours de Réalité Virtuelle - orthophonie. ISTR - UCBL. Master 1*
2022-	Co-Responsable. Cours de Oculométrie fondamentale et clinique. ISTR - UCBL. Licence
2021-	Responsable. TP Handicaps et Réalité Virtuelle-Augmentée. ISTR - UCBL. Master 1**
2017-	Responsable. Cours de Plasticité cérébrale (mutualisé). ISTR - UCBL. Licence**
2016-	Responsable. Cours de Exploration et Imagerie du vivant. ISTR - UCBL. Licence
2015-	Co-Responsable. Cours de Neuropsychologie. ISTR – UCBL. Licence
2015-	Responsable. Cours de Neuropsychologie (mutualisé). ISTR – UCBL. Licence
2015-	Responsable. Cours de Psychologie Générale (mutualisé). ISTR – UCBL. Licence
2005	½ ATER. Cours de Neurosciences, UFR de Biologie -UCBL. Licence et Master 1
2004	½ ATER. Cours de Sciences cognitives. UFR STAPS UJF- Grenoble. Licence
2014	Vacataire. Physiologie. Kinésithérapie Ergothérapie, ISTR- UCBL. Licence
2014	Vacataire Neurophysiologie, ISTR- UCBL. Master 2
1999-2001	Vacataire Psychophysiologie, UFR STAPS UCBL. Licence
1996-2001	Vacataire Neuropsychologie. Université Lumière Lyon II. Licence
1996	Vacataire Neuropsychologie. Université Lumière Lyon II. Licence & Master 1
1996	Vacataire Biologie animale. Université Lumière Lyon II. Licence
2008-2014	Formatrice. Physiologie & Neurosciences. Kiné Déficiants de la vue Lyon. PAS-Licence
1997-2014	Formatrice. Biologie, Instituts de Formation en Soins Infirmiers, et Manipulateurs Radio, HCL,
1998-2002	Formatrice. Biologie pour concours paramédicaux, Lycée JB de la Salle, Lyon

* Innovation pédagogique en spatialisation auditive 'SPHERE pour les étudiants'

** innovation pédagogique sur le handicap visuel en Réalité Virtuelle et Réalité Augmentée

Formations pédagogiques

2018	« Pédagogie innovante : formation à distance, formation hybride, classe inversée »
2018	« Atelier Claroline Connect »
2021	« Intégrer la ludopédagogie dans l'animation de formation à distance »
2021	« Adapter son TP/TD à distance avec la vidéo asynchrone »
2021	« Virtual Lunch Break »
2021	« Virtual Lunch Break : La réalité virtuelle au service de la réadaptation ». Participation
2022	« Enseigner avec la réalité virtuelle »
2022	« Virtual Lunch Break »
2023	« Atelier Moddle »

Innovations pédagogiques et Prix

- 2021 « HandVisSim-RV, la réalité du handicap à travers la réalité virtuelle ». Collaboration avec la plateforme Neuroimmersion du CRNL pour la réalisation d'une application en réalité virtuelle.
- 2022 « HERVé, Handicap et Expérience en Réalité Virtuelle pour les étudiants en rééducation ». 1^{er} Prix de la Région Auvergne-Rhône Alpes pour ce projet collaboratif lors du HackingHealth Lyon # 6.
- 2022 « Quand l'élève est aussi le professeur, une histoire en prisma-vision ». Collaboration avec le service Innovation Conception et Accompagnement pour la Pédagogie ICAP VR pour la réalisation d'un film sur l'adaptation prismatique.
- 2022 « HandVisSim-RA, la réalité du handicap à travers la réalité augmentée ». Collaboration avec le service Innovation Conception et Accompagnement pour la Pédagogie ICAP VR pour la réalisation d'une application en réalité augmentée.
- 2023 « SPHEREapp : usage pédagogique du dispositif SPHERE pour les étudiants en audioprothèse et orthophonie »

Jurys de diplôme et membre de comité

- Jurys d'Ergothérapie, Masso-kinésithérapie et Orthophonie pour l'obtention du DE
- Jurys de Masso-kinésithérapie pour Déficiants visuels pour l'obtention du DE
- Jurys de M2 du Master Santé de l'ISTR.
- Jury de M1 du Master Neurosciences de l'UFR BioScience de Lyon
- Jury du baccalauréat Professionnel
- Membre de Comités de Suivis de Thèse : C. Bonnet (Univ. Bourgogne- Franche Comté), L. Seropian et M. Ferschneider (UCBL), V. Ardonceau (Univ. Bourgogne- Franche Comté) , C. Hoarau (UCBL)

Activités collectives

a) Champ administratif

- Conseil du Centre de Recherche en Neurosciences de Lyon. 2011-2013
- Conseil du Centre de Recherche en Neurosciences de Lyon. 2023-2026
- Comité Technique d'Orientation des Travaux de Recherche, UCBL. 2016-2018
- Conseil d'administration de l'ISTR. 2021-2026.
- Groupe de Travail Structures et Conventions de l'UCBL. 2020-2025
- Comité de renouvellement quinquennal de l'équipe IMPACT du CRNL. 2021-2025
- Organisatrice et présidente du comité de la 1^{ère} Journée Scientifique de l'ISTR. 2023
- Comité d'organisation NeuroFrance Lyon. 2023

b) Champ scientifique

- Éditrice pour Frontiers In Neurosciences. 2022

c) Champ pédagogique

- Membre du Conseil pédagogique de la formation en Masso-kinésithérapie de l'ISTR
- Membre du Conseil pédagogique de la Licence Pro Codeur en Langue Parlée Complétée de l'ISTR
- Membre du Conseil de Sélection (COSEL) des formations en Orthoptie et Psychomotricité de l'ISTR
- Encadrement de stagiaires de collège, lycée, étudiants en 2ème année d'IUT de Biologie A, PCM2 médecine, Master 2 de l'université de Bologne (Italie).

d) Champ d'événements pour le grand public

- « Sciences en Fête », Institut Fédératif des Neurosciences de Lyon, 2004 et 2005
- « L'Homme en mouvement », Village des Sciences, Paris, INSERM, 2005.
- « Motricité et Neuroplasticité : Neurosciences, cerveau et système nerveux ». IFE- ENS Lyon. 2012
- « A la découverte des laboratoires », INSERM Rhône-Alpes. 2012
- « L'emprise des Sens », Bibliothèques Municipales Lyon & INSERM. 2012 et 2013
- « Semaine du cerveau », CRNL, 2011, 2012 et 2013
- « C'est où ? La spatialisation auditive ». 26ème Fête de la Science. Musée des Confluences de Lyon. 2017.
- « Virtuellement tout est possible ! ». 11ème Festival Sciences et Mangas. La Doua Lyon. 2021.
- « Les possibles de la Réalité Virtuelle ». 30ème Fête de la Science INSERM Lyon. 2021
- « Le handicap et la RV ». INSERM et Zone d'éducation prioritaire de Vénissieux. 2021.
- « Lab'INSERM Saison 2- Ep.2 : NeuroImmersion ». Participation à la série vidéo YouTube. 2022
- « Conférences embarquées' », CNRS, Parc de la Tête d'Or Lyon. 2022.
- « La réalité du handicap à travers la RV ». 31-32ème Fête de la Science, Hôp. Sp. Vinatier. 2021- 2022.
- « Des fourmis dans les jambes », 24ème Ed. Semaine du Cerveau. Musée C. Bernard, 2022.
- « La réalité du handicap à travers la réalité virtuelle ». ENS de Lyon-Science. 2022.
- « Journée de sensibilisation au handicap visuel ». IFMK pour Déficients visuels de Lyon. 2022.
- « Rêv'Elles : Parcours de Femmes en Recherche ». CRNL Lyon. 2024.

e) Champ d'événements pour le grand spécialisé

- « Physiologie de la jonction musculo-tendineuse ». Journée Société Française de Radiologie Rhône-Alpes, Traumatologie du muscle et du tendon en pratique sportive, imagerie, conduite à tenir », Institut des Sciences Cognitives, Lyon 2012.
- « Motricité et Neuroplasticité ». Institut Français de l'Education IFE « FormaVie : Neurosciences, cerveau et système nerveux ». ENS Lyon 2012.

1. Activités de Recherche passées

Mon parcours en recherche a débuté il y a plus de 20 ans, en 1995, lorsque j'ai rejoint le laboratoire U94 sous la direction du Professeur M. Jeannerod. À l'époque, en raison du manque de financement pour entreprendre une thèse en électrophysiologie chez le primate humain, j'ai choisi de partager mes connaissances à travers diverses formations académiques (e.g. en STAPS) et professionnelles (e.g. dans des Instituts de Formation en Soins Infirmiers), tant dans le secteur de la formation privée que publique. Ces expériences de formation m'a rapidement fait prendre conscience du rôle crucial de l'enseignement dans ma carrière professionnelle.

En parallèle, Claude Prablanc, chercheur au sein de l'U94, a éveillé en moi l'intérêt pour la psychophysique et m'a fait prendre conscience de la complexité du contrôle moteur et de l'intérêt de son étude. En 1998, j'ai pris la décision de me consacrer pleinement à l'enseignement, tout en maintenant l'idée de persévérer dans mes activités de recherche.

Forte de cette nouvelle expertise en enseignement et désireuse de renouer avec la recherche en neurosciences, je suis devenue enseignante-chercheuse "indépendante", travaillant sur contrats. Pendant cinq ans, j'ai jonglé entre des contrats d'enseignement et de recherche, ainsi qu'avec des services tant publics que privés (CNRS, INSERM, Hospices civils de Lyon, STAPS de l'Université de Lyon, lycée privé, Rectorat, etc.). Cette période m'a permis de collaborer sur différents projets de recherche menés chez l'animal (Ravel et al., 2003) et l'Homme (Karoumi et al., 2001).

En 2001, j'ai saisi l'opportunité de réaliser une thèse en neurosciences à l'Université C. Bernard de Lyon, au sein de la nouvelle équipe U864, portant sur le contrôle moteur et le rôle des afférences visuelles. En 2005, j'ai obtenu un doctorat en neurosciences me permettant d'être sélectionnée dans des Conseils Nationaux des Universités (en neurosciences - CNU 69, et en psychologie -CNU 74). Ces sélections m'ont permis de postuler la même année pour des postes universitaires (Maître de Conférences MCU). Depuis l'obtention de ma thèse, j'ai continué à participer activement à la recherche et à l'enseignement universitaire, ainsi qu'à la vulgarisation scientifique auprès de différents publics. Cette période m'a permis de développer mes compétences dans le domaine de la plasticité sensorimotrice en utilisant diverses approches psychophysiques chez des individus sains et des patients (e.g. patients hémiparétiques, héminégligents, etc.), en collaborant avec des chercheurs de l'équipe U864 et ensuite de l'équipe IMPACT du Centre de Recherche en Neurosciences de Lyon (CRNL, créé en 2011).

En 2012, une nouvelle réglementation de la recherche biomédicale a été introduite avec la loi Jardé. Cette loi visait à simplifier le cadre juridique en adaptant les réglementations en fonction des risques encourus par les participants aux recherches. Bien que sa mise en application ait été retardée, elle a profondément impacté la recherche fondamentale impliquant des humains. En 2013, j'ai entrepris une formation universitaire d'un an pour

devenir investigatrice d'essais cliniques, suivie de stages hospitaliers dans le cadre d'essais cliniques académiques et industriels. Bien que cette formation m'ait permis de mieux comprendre les enjeux de la recherche impliquant des humains et de maîtriser la conduite de protocoles de recherche (notamment les aspects de collaboration entre les acteurs hospitaliers, académiques et juridiques), j'étais insatisfaite de ne pas être davantage impliquée dans la conception des projets que je menais.

Enfin, en 2015, j'ai obtenu l'opportunité d'intégrer le corps des MCU, et j'ai été nommée enseignante-chercheuse à l'Université Claude Bernard de Lyon, au sein de l'Institut en Sciences et Techniques de la Réadaptation (ISTR). Mon projet, développé en collaboration avec les membres de l'équipe IMPACT, portait sur "L'intégration multisensorielle chez les patients déficients auditifs". Cette opportunité m'a permis de mettre à profit mes expertises acquises d'une part sur la plasticité sensori-motrice et d'autre part sur la conduite de protocoles d'études chez l'Homme. De nouvelles collaborations ont émergé avec des ingénieurs de la plateforme Neuroimmersion du CRNL, tels que R. Salemme et A. Belle, ainsi qu'avec des chercheurs d'exception comme le professeur F. Pavani de l'Université de Trento et A. Farnè, directeur de recherche à IMPACT. Une nouvelle ère de recherche s'est ouverte pour moi en tant qu'experte des systèmes sensorimoteurs et spécialiste des essais cliniques, en vue de développer des travaux sur la spatialisation auditive.

Parallèlement à mon intégration au sein de l'équipe IMPACT, j'ai rejoint l'ISTR pour accomplir mes missions d'enseignement. L'ISTR regroupe des formations paramédicales au sein de l'Université Lyon1, menant à l'obtention de diplômes universitaires (licence ou master) et de diplômes d'État (sur une durée de trois ou cinq ans). Bien que ces formations se concentrent principalement sur la professionnalisation des étudiants dans les métiers paramédicaux (exception faite du Master Santé, délivrant un diplôme académique), j'ai pu enrichir mes compétences grâce à mes expertises en neurosciences et dans la conduite d'études cliniques en réhabilitation, notamment pour la rééducation des déficits auditifs. En tant que maîtresse de conférences non issues du corps paramédical, je ne peux pas assumer la responsabilité d'une formation ou d'une année de diplôme, mais je contribue aux programmes d'enseignement et assure des enseignements adaptés aux formations. La mutualisation de certains de mes enseignements complexifie le format (groupes nombreux aux profils variés, avec des objectifs de professionnalisation et des parcours de formation différents). Par conséquent, je dois régulièrement rappeler ma position en tant que MCU au sein des formations paramédicales, tout en proposant des innovations pédagogiques pour les métiers de la rééducation et en accompagnant des étudiants en stage et en thèse en neurosciences.

Comme vous l'avez compris, mon parcours s'est enrichi de diverses approches et thématiques scientifiques. Bien que je divise arbitrairement mon parcours de recherche en deux parties, avec l'obtention du statut d'enseignant-chercheur en 2015 comme repère temporel, deux grandes thématiques émergent : (1) la flexibilité

et l'adaptation de la précision des mouvements oculaires et du bras vers des cibles visuelles (21 articles publiés) et (2) l'impact de l'intégration multisensorielle et de l'exploration active de l'environnement sur la spatialisation auditive (11 articles publiés).

Ces deux volets partagent un dénominateur commun qui forme mon expertise depuis plus de 20 ans dans la conduite de projets de recherche et la mise en œuvre de mes enseignements : les interactions perception-action et leurs implications en rééducation. Bien que ces deux aspects soient liés, pour des raisons de concision et de clarté, j'ai choisi de ne présenter dans ce manuscrit que mes travaux les plus récents. Ces derniers se concentrent sur le thème de la spatialisation auditive et appuient ma candidature au diplôme d'Habilitation à Diriger les Recherches.

Ainsi, ce manuscrit se structure en plusieurs parties qui permettront de contextualiser :

- La problématique de l'étude de la spatialisation auditive. A travers différents chapitres, cette partie présentera les différentes problématiques liées à la spatialisation auditive ainsi que mes contributions scientifiques dans ce domaine ;
- Une présentation de mes projets de recherche en spatialisation auditive ;
- Une conclusion ainsi qu'une réflexion personnelle sur l'évolution des études de spatialisation auditive ;
- Pour compléter ce document, plusieurs articles que j'ai coécrits avec des étudiants ont été sélectionnés et inclus, afin de détailler les méthodologies des études qui n'auraient pas été couvertes dans le document principal.

2. Activités de Recherche présentes

Généralités sur la SA

La spatialisation auditive (SA), également appelée localisation auditive ou audition spatiale, est un processus perceptif essentiel conférant la capacité de localiser des sources sonores dans notre environnement et contribuant à la représentation de notre espace sensoriel. Cruciale pour la survie des animaux, elle les informe des éléments acoustiques environnants, facilitant ainsi leur positionnement en tant que proies ou prédateurs (Blauert, 1996; Rayleigh, 1907). Bien que la résolution spatiale de la SA soit inférieure à celle de la vision (Blauert, 1996; R. S. Heffner & Heffner, 1992), il n'en demeure pas moins que la SA reste nettement supérieure à la vision pour localiser des sources sonores. En effet, non seulement la SA fonctionne dans l'obscurité, mais elle fournit également des informations sur les sources situées en dehors du champ de vision. De plus, elle offre un temps de réaction plus court lorsque l'éloignement des sources n'entraîne pas de retard de propagation de l'information acoustique par rapport à l'information visuelle (voir la revue d' Alais et al., 2010). Toutes ces caractéristiques font que certains environnements sont plus "transparents" au son qu'à la lumière.

Localiser une source sonore dans l'espace 3D (localisation absolue) implique de déterminer sa position en azimut (plan horizontal), en élévation (plan vertical) et en profondeur. Les humains localisent les sources sonores en utilisant des indices captés par les oreilles ((Blauert, 1996; Middlebrooks, 2015; Yost, 1998). Ces signaux, qu'ils soient monauraux (captés à partir d'une seule oreille, Angell & Fite, 1901) ou binauraux (en combinant les signaux des deux oreilles, Rayleigh, 1907), résultent des différences de trajets et de transformations subies par les ondes sonores lors de leur déplacement autour de la tête

Lorsque le son contourne la tête, il subit des altérations dues aux caractéristiques anatomiques de celle-ci, créant un retard (différence de phase entre les oreilles, appelée ITD ou Interaural Time Difference) et une atténuation (différence de niveau entre les oreilles, appelée ILD ou Interaural Level Difference), qui sont essentiels pour la localisation horizontale. De plus, les pavillons des oreilles, la tête et le torse modifient également les ondes sonores, et ces transformations sont regroupées sous le terme fonctions de transfert (HRTFs, voir les revues de Blauert, 1996; Brungart & Rabinowitz, 1999; Hofman & Van Opstal, 2003; Middlebrooks, 2015) cruciales pour la localisation en élévation. La localisation en profondeur fait appel à divers indices. Ainsi, les indices d'intensité binauraux (ILD) sont utilisés pour percevoir la distance sonore en champ proche (Brungart & Rabinowitz, 1999; Kolarik et al., 2016), tandis que des facteurs tels que la réverbération et les caractéristiques spectrales du son sont pris en compte en champ lointain (Blauert, 1996; Brungart et al., 1999; Middlebrooks &

Green, 1991). À mesure qu'une source se rapproche, l'intensité sonore perçue par les deux oreilles augmente (Zahorik, 2002), et cette variation fournit un indice relatif de distance.

Ainsi, dans des conditions d'écoute normale avec une audition normale, différents indices sont pondérés et combinés pour permettre la représentation optimale des positions spatiales des événements auditifs dans l'espace tridimensionnel, assurant ainsi une perception spatiale des sons.

L'intégration des indices auditifs spatiaux s'effectue le long du névraxe et implique, en fin de compte, un réseau cérébral qui va au-delà des structures d'intégration auditive. En effet, des études ont démontré que des structures sous-corticales (Antunes & Malmierca, 2021) et des structures corticales non auditives (comme le cortex frontal et pariétal, Degerman et al., 2006) sont également modulées en relation avec la localisation sonore. Plusieurs études en neuroimagerie tendent à suggérer l'existence de réseaux corticaux dédiés au traitement des indices acoustiques et impliquant les cortex auditif, frontal et pariétal dans le processus de localisation auditive (e.g. Alain et al., 2001; Arnott et al., 2004; Brunetti et al., 2005; van der Zwaag et al., 2011). Ces réseaux seraient organisés en deux voies spécialisées : une voie "quoi" pour l'identification de l'objet sonore et une voie "où" pour le traitement spatial et l'orientation de l'attention (analogue au modèle pour le système visuel de Goodale & Milner, 1992).

Si récemment des études ont montré que la localisation auditive est un traitement intégré qui est déjà présent au sein du cortex auditif primaire (Wood et al., 2019), le réseau fonctionnel cérébral auquel appartient ce cortex (van der Heijden et al., 2019) tend à montrer que la représentation de l'espace acoustique est une construction sollicitant d'autres ressources cérébrales (van der Heijden et al., 2019). Au-delà de cette reconnaissance fonctionnelle attribuée au cortex auditif primaire (qui devient de moins en moins "primaire" si l'on considère la localisation auditive comme un attribut sonore d'ordre supérieur), c'est surtout l'organisation du réseau cortical auditif dorsal qui mérite d'être explorée pour identifier les facteurs impliqués dans la localisation des sons. Ainsi, dans son article, Rauschecker (Rauschecker, 2018) suggère que la voie dorsale auditive devrait être conçue comme étant davantage liée à "comment" nous interagissons avec les objets, c'est-à-dire comment nous combinons des éléments sensorimoteurs pour localiser et interpréter les sons dans notre environnement, et ainsi pouvons organiser et produire des actions dirigées vers les objets. L'hypothèse fonctionnelle de cette voie auditive dorsale laisse entrevoir que la SA est un processus complexe mettant en jeu des circuits sensori-moteurs.

En effet, les sons perçus ne contribuent pas seulement à la construction de la représentation auditive de l'espace, mais ils influencent également les stratégies d'exploration (Kolarik et al., 2016). Par exemple, localiser une source sonore permet d'orienter le regard (Brungart et al., 1999) afin d'obtenir des informations plus précises sur la source sonore, telles que des informations visuelles. Cette complémentarité de représentations auditive et visuelle s'avère essentielle pour construire une représentation cohérente de l'espace (Bregman, 1990), notamment quand l'audition est altérée (Bishop et al., 2012; H. E. Heffner & Heffner, 2014; R. S. Heffner & Heffner, 1992;

Knudsen & Knudsen, 1985). Cette complémentarité se manifeste également dans la fonction d'équilibre, où la localisation des sources auditives améliore la stabilité posturale (Gandemer et al., 2017; Zhong & Yost, 2013).

Enfin, la spatialisation auditive est étroitement liée à d'autres fonctions cognitives. Une des fonctions étroitement liées à la SA est l'attention. En effet, la spatialisation des sons peut involontairement ou sélectivement diriger l'attention vers des indices spatiaux importants (Fritz et al., 2007; B. Shinn-Cunningham et al., 2017) ; le travail de concert entre les cortex auditif et préfrontaux explique, entre autres, pourquoi l'attention portée à des emplacements spécifiques permet un traitement plus rapide des cibles auditives (Mondor & Zatorre, 1995), accélère le traitement de la localisation sonore (Johnen et al., 2001), en particulier dans des environnements acoustiques complexes (Kolarik et al., 2016). De plus, l'attention facilite l'intégration d'autres indices sensoriels (pour revue Spence & Santangelo, 2009) et favorise les interactions du sujet avec son environnement sonore (Ahveninen et al., 2006; Kolarik et al., 2016; Wightman & Kistler, 1999).

Au-delà des interactions réciproques entre la SA et l'attention, la SA contribue également à la communication. En fournissant des indices sur le nombre relatif et la position des sources, elle permet une ségrégation auditive de l'espace, facilitant ainsi le suivi d'une source sonore spécifique parmi d'autres sources concurrentes. Cette ségrégation contribue à l'effet de "cocktail party", facilitant ainsi la compréhension de la parole dans le bruit (Akeroyd, 2008; Oberfeld & Klöckner-Nowotny, 2016; B. Shinn-Cunningham et al., 2017).

Cette capacité à ségréguer les informations auditives de l'environnement, et plus généralement l'audition, participe également à la préservation de l'intégrité cognitive. En cas de détérioration des performances de SA, souvent observée chez les personnes âgées ou celles atteintes de troubles auditifs, on observe une instabilité posturale accrue et un risque de chute (pour une revue, voir Seiwert, 2023), ainsi que des difficultés à suivre des conversations en milieu bruyant, des altérations de la performance cognitive globale, y compris de l'attention (B. G. Shinn-Cunningham & Best, 2008) et de la mémoire (Shende & Mudar, 2023). Ces effets néfastes peuvent entraîner une diminution significative de la qualité de vie, conduisant à l'isolement social, à la stigmatisation et à la dépression (Shield, 2019), ainsi qu'à un déclin cognitif (Loughrey et al., 2018). La détérioration de la capacité de localisation des sons a des répercussions importantes sur la situation économique et l'indépendance des individus touchés (Nordvik et al., 2018).

Parallèlement à ces constats négatifs, il est reconnu que la perte auditive multiplie jusqu'à cinq fois le risque de maladies neurodégénératives et constitue un facteur de risque de démence (Lad et al., 2022). Cependant ces risques peuvent être limités par une prise en soin adaptée (Livingston et al., 2017). Prendre en charge les déficits de SA est un enjeu majeur en matière de santé, étant donné que près de 466 millions de personnes dans le monde souffrent de pertes auditives importantes nécessitant une prise en soin médicale (World Health Organization, 2021). Jusqu'à présent, les aspects de la SA ont été largement négligés dans le diagnostic et la prise

en soin des patients atteints de pertes auditives profondes, mais aussi dans les stratégies visant à élaborer des solutions technologiques pour compenser ces déficits.

L'ensemble de ces concepts généraux montre que l'étude de la spatialisation auditive va au-delà de la simple investigation de l'intégration des signaux acoustiques. L'implication d'un réseau auditif dorsal dans ce processus suggère : (1) l'implication de processus d'intégration multisensorielle, en particulier le rôle des informations visuelles dans la localisation précise des sons ; (2) le rôle des mouvements intentionnels des yeux et de la tête dans l'alignement des informations visuelles et auditives pour améliorer la perception de la localisation des sons ; (3) l'impact de l'intégration de nouvelles informations (par exemple, en cas d'altération des indices acoustiques) dans la dynamique de ce réseau ; et (4) pourquoi une modification de la représentation auditive spatiale impacte d'autres fonctions cognitives (par exemple, la compréhension de la parole dans le bruit, l'attention)

Si le processus de SA est un processus complexe, son étude doit tenir compte des différents facteurs qui contribuent à son élaboration. Ainsi, elle nécessite le développement de nouvelles méthodologies de recherche permettant d'examiner comment l'intégration multisensorielle et les actions dirigées dans l'espace acoustique affectent la SA. De plus, cette stratégie de recherche offre l'opportunité de comprendre les éléments qui contribuent à la réorganisation de la représentation spatiale auditive, un concept clé dans la rééducation des déficits de SA. Comprendre quels facteurs contribuent à la construction de l'espace acoustique est d'autant plus pertinent lors de l'évaluation et de la rééducation des patients présentant des déficits de SA (comme les porteurs d'implants cochléaires), pour lesquels il n'existe ni évaluation ni prise en soin spécifique de leurs capacités à localiser les sons.

Les chapitres suivants exposeront mes diverses contributions scientifiques à l'étude de la SA :

- Le chapitre 1 présentera comment, à travers une nouvelle approche combinant différents prérequis méthodologiques pour étudier la SA, j'ai pu (1) démontrer qu'il est possible d'évaluer les capacités de SA en tenant compte des trois dimensions de l'espace de stimulation et de réponse, (2) démontrer l'importance de contrôler le regard et les informations visuelles lors de l'évaluation des capacités de SA, (3) identifier le rôle majeur des stratégies d'exploration dans la construction de son espace auditif et (4) montrer la dynamique adaptative de cet espace.
- Le chapitre 2 abordera l'évaluation et la rééducation des déficits de SA. Dans ce chapitre, je montrerai qu'en étudiant différents modèles d'altération auditive, nous avons pu identifier des éléments cruciaux qui participent à la plasticité de la SA, éléments qui sont le pivot de futures prises en soin des déficits de SA.

CHAPITRE 1. Étudier les performances de localisation auditive : prérequis méthodologiques.

Supervisions d'étudiants

- Eeckhout J-L (Psychomotricienne DE). Master 2 Santé. Handicaps Sensoriels et Cognitif
- Salteur de la Serraz T. (Kinésithérapeute DE). Master 2 Santé. Handicaps Sensoriels et Cognitif
- Coudert A. (Formation médicale). Master 2 Neurosciences et Doctorat de Neurosciences

Articles liés au projet

- **Gaveau V.**, **Coudert A.**, Salemme R, Koun E, Desoche C, Truy E, Farné A, Pavani F. Benefits of active listening during 3D sound localization. *Exp Brain Res.* 2022 Nov;240(11):2817-2833.
- **Coudert A.**, **Gaveau V.**, Gatel J., Verdelet G., Salemme R., Farné A., Pavani F., Truy E. Spatial hearing difficulties in bilateral cochlear implant children improve with head movements. *Ear and Hearing.* Ear Hear. 2022 Jan/Feb;43(1):192-205
- Valzolgher C., Alzhaler M., Gessa E., Todeschini M., Nieto P., Verdelet G., Salemme S, **Gaveau V.**, Marx M., Truy E., Barone P., Farné A. and Pavani F. The Impact of a Visual Spatial Frame on Real Sound-Source Localization in Virtual Reality. *Current Research in Behavioral Sciences* 1. Nov. 2020
- Valzolgher C, Verdelet G, Salemme R, Lombardi L, **Gaveau V.**, Farné A, Pavani F. Reaching to sounds in virtual reality: A multisensory-motor approach to promote adaptation to altered auditory cues. *Neuropsychologia.* 2020 Dec; 149:107665.
- Verdelet, G., Salemme R., Desoche C., Volland F., Farné A., **Coudert A.**, Hermann R., Truy E., **Gaveau V.** and Pavani F. Assessing Spatial and Temporal Reliability of the Vive System as a Tool for Naturalistic Behavioural Research. In 2019 International Conference on 3D Immersion (IC3D)
- P. Barone, F. Pavani, A. Farné, E. Truy, V. **Gaveau**, O. Deguine. Intégration multisensorielle chez le patient déficient auditif. In *Surdit , actualit s, innovations et espoirs- Perception auditive.* 2018, ELSEVIER-MASSON, de Truy E., Lescanne E., Loundon N., Roman S. 2018.
- Pavani F., Farn  A., Truy E. and **Gaveau V.** Perception auditive spatiale. In *Surdit , actualit s, innovations et espoirs- Perception auditive.* 2018, ELSEVIER-MASSON, de Truy E., Lescanne E., Loundon N., Roman S. 2018.

Brevet li  au projet

« SPHERE: Method and apparatus for acquiring a spatial map of auditory perception of a subject ». Brevet europ en WO2017203028A.

Financements des  tudes

- PICS CNRS
- Appel   Manifestation d'int r t par Pulsalys (SATT Lyon)
- ITMO NNP
- ANR-16-CE17-0016
- LabEx CORTEX SATT (ANR-11-LABX-0042).
- Institut rhonalpin des Syst mes complexes (IXXI)
- SPHERIC IIR-1948

Collaborations

- Services d'ORL des Hospices Civils de Lyon (Lyon, France)
- Plateforme Neuroimmersion du Centre de Recherche en Neurosciences de Lyon (Lyon, France)
- Center for Mind/Brain Sciences (CIMEC), University of Trento, (Rovereto, Italie)
- Centre de Recherche Cerveau & Cognition, (Toulouse, France)

Rationnel

1. Prérequis n°1 : Étudier la spatialisation auditive dans les 3 dimensions de l'espace

Dans les paradigmes conventionnels d'étude de la spatialisation auditive, les stimulations auditives sont généralement réalisées en champ libre, c'est-à-dire à l'aide de dispositifs de haut-parleurs disposés autour du sujet (pour une revue Akeroyd & Whitmer, 2016). Cependant, en raison de la complexité de ces dispositifs (voir la figure 1 extraite de Tabry et al., 2013), ils ne permettent pas d'échantillonner l'intégralité de l'espace acoustique, qui est par essence tridimensionnel. Ainsi, bien que ces dispositifs puissent explorer l'espace frontal (haut-parleurs disposés en demi-cercle) voire l'espace arrière (haut-parleurs disposés en cercle), aucun dispositif n'existe pour explorer simultanément toutes les dimensions de l'espace, notamment la profondeur. Par conséquent, les dispositifs actuels ne permettent d'étudier les performances 3D de la SA.

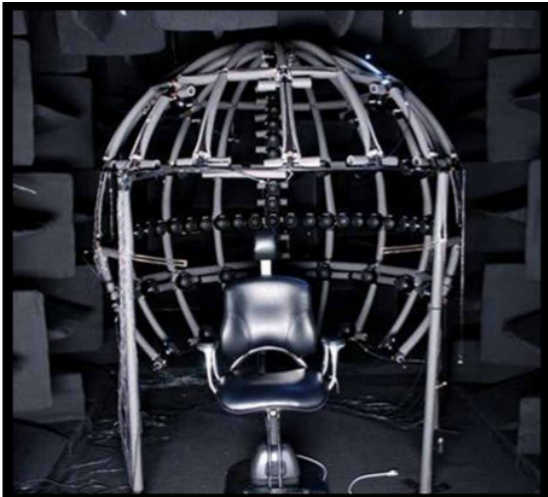


Figure 1. Système de localisation issue de l'expérience de Tabry et al. 2013

Or, exploiter les performances de la SA revient à les caractériser dans toutes les dimensions de l'espace (azimute, élévation et profondeur) tout en offrant la possibilité d'échantillonner l'espace sonore dans sa globalité. Un premier défi à relever consiste à développer une approche méthodologique qui offre la liberté d'étudier tout l'espace 3D sonore autour du sujet (c'est-à-dire de placer une source sonore à une position 3D désirée), sans contrainte technique, et tout en contrôlant la position de la source, condition sine qua non pour assurer une stimulation reproductible au niveau des oreilles entre les essais, les sessions et les participants. Pour s'assurer d'une stimulation acoustique reproductible, il est courant que la tête du sujet soit immobilisée (Brungart et al., 1999; Litovsky et al., 2009; Pavani et al., 2008; Távora-Vieira et al., 2015), ou bien que la durée de la stimulation auditive soit réduite de manière à limiter les mouvements de tête des participants pendant l'émission des sons (par exemple Ahrens et al., 2019).

De plus, bien que les sons émis en champ libre soient supposément perçus par le sujet dans son espace tridimensionnel, la méthodologie de mesure des performances de la SA varie considérablement d'une étude à l'autre et ne permet pas d'accéder aux performances en 3D de manière uniforme. Ainsi, certaines études ont utilisé des supports de réponse en 2D tels qu'une tablette tactile tandis que d'autres études ont restreint les réponses des participants à une seule dimension (i.e. la performance étant évaluée par la différence d'angle entre leur réponse et celle de la source sonore) ou deux dimensions (i.e. différence entre la direction de la réponse et de la source) (e.g. Andéol & Simpson, 2016; Haber et al., 1993; Oldfield & Parker, 1984; Wightman & Kistler, 1999). Pourtant, plusieurs études ont démontré que l'interaction avec l'environnement permettait d'améliorer les performances de localisation auditive (Aggius-Vella et al., 2017; Brungart et al., 2000; Etchemendy et al., 2017; Haber et al., 1993). Ainsi, un autre défi à relever est de permettre aux sujets d'interagir avec leur environnement afin d'estimer leur capacité de SA en 3D, sans qu'ils soient limités par le dispositif expérimental, et tout en garantissant la reproductibilité des mesures entre les essais, les sessions et les participants.

Afin d'échantillonner l'espace acoustique tridimensionnel tout en mesurant les performances de spatialisation 3D des sujets, nous avons développé une solution simple consistant à appliquer un référentiel commun à la fois pour décrire l'espace de stimulation et l'espace de réponse du sujet. Cette solution propose de définir le centre de la tête du participant comme référentiel commun. Cette approche unifie l'espace de stimulation et de réponse, et présente plusieurs avantages : (1) elle évite d'adapter le dispositif expérimental aux considérations physiques des participants (i.e. modifier le dispositif de stimulation en fonction de la position du participant dans la pièce expérimentale), (2) elle ne contraint pas le participant à occuper une position spécifique dans l'espace de stimulation, et (3) elle permet de décrire les performances de SA des participants dans un même cadre de référence. La solution 3D que nous avons choisie pour répondre à ces exigences méthodologiques repose sur l'utilisation d'un système de réalité virtuelle (RV).

Ma contribution dans ce champ méthodologique s'est concentrée sur la validation interne de cette nouvelle approche pour l'étude de la spatialisation auditive (Gaveau et al., 2022). Elle s'est appuyée sur l'utilisation d'outils validés dans les études en neurosciences pour leur robustesse dans l'acquisition spatiale et temporelle des mesures. Cependant, ces outils scientifiques sont coûteux et restent le plus souvent cantonnés à la sphère de la recherche fondamentale ce qui en limite leur déploiement en dehors de ce champ d'expertise, et notamment en direction du champ clinique. De plus, fort et de constater que la technologie des systèmes de RV évolue très vite, et qu'il est maintenant possible d'intégrer des systèmes de RV commerciaux dans les approches méthodologiques en neurosciences, et en conséquence de proposer ces approches à un plus vaste panel d'utilisateurs. Ainsi, j'ai contribué à la validation de l'utilisation d'un système de RV grand public pour répondre aux exigences méthodologiques des études de la SA (Verdelet et al., 2020).

☞ Contribution= une nouvelle approche pour l'étude de la SA (Gaveau et al. 2020)

Pour échantillonner l'espace acoustique de manière non contraignante à la fois techniquement pour l'expérimentateur et pour le participant, l'approche la plus simple consiste à utiliser une source sonore unique (haut-parleur) et à la déplacer à la position désirée autour du sujet. Cela requiert à la fois la capacité de contrôler en temps réel sa position tridimensionnelle et de garantir la reproductibilité de cette position d'un essai à l'autre, d'une session à l'autre et d'un participant à l'autre.

Pour ce faire, nous avons tiré parti des avancées de la technologie de réalité virtuelle (RV). En effet, un système de RV permet nativement de suivre en continu la tête du sujet (et de mettre à jour les informations visuelles affichées dans le casque de RV en suivant le référentiel du centre de la tête), ce qui permet de déterminer la position du centre de la tête. En 2015, nous avons choisi d'utiliser le système RV Oculus® DK2 (qui disposait également d'un système de suivi de la position des yeux). Cependant, ce dispositif commercial n'était pas conçu comme un outil de recherche, nous avons donc associé ce système de RV à un système de capture de mouvement 3D en temps réel (Vicon®) dont les données acquises correspondaient aux exigences spatiales et temporelles pour mener des travaux de recherche (les détails techniques sont disponibles dans l'article en annexe).

En couplant ainsi le système de RV avec le système de capture de mouvement, notre approche permettait de guider et de contrôler la position des sources sonores en champ libre vers des coordonnées prédéfinies (voir figure ci-dessous), tout en enregistrant en temps réel la position tridimensionnelle de la tête et de la main du sujet (détails méthodologiques disponibles dans l'article en annexe). Ainsi, le dispositif de stimulation et celui de mesure en 3D étaient calculés à partir de la position du centre de la tête du participant, cette dernière devenant alors le centre du dispositif expérimental.

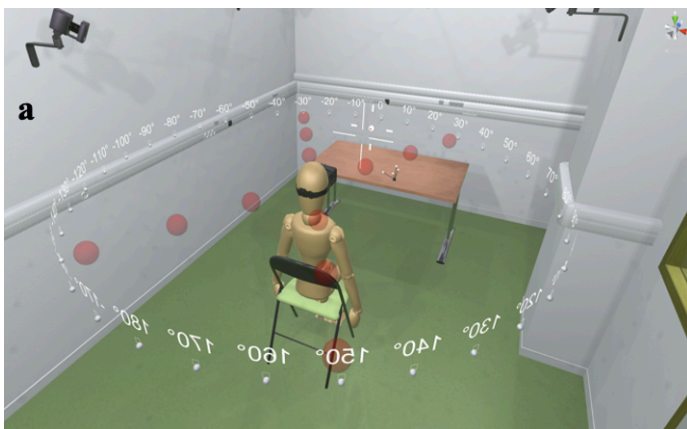


Figure 2. Représentation des différentes positions 3D désirées de la source sonore, exprimées selon le référentiel centre tête du sujet.

Remarque : Avec cette approche méthodologique du "référentiel centre tête", il est possible d'aligner dans un même espace les informations auditives et visuelles. Les informations visuelles, disponibles dans le casque de RV, sont naturellement alignées sur le centre de la tête. Cette approche sera détaillée dans les études du chapitre 2.

Résultat. Notre dispositif ne contraint pas la posture du participant tout en assurant une reproduction satisfaisante de la position 3D d'une source sonore placée manuellement autour du sujet. L'erreur tridimensionnelle du placement manuel de la source sonore (position réelle) par rapport à la position souhaitée était inférieure à 1 cm (voir figure 3). Par cette étude, j'ai pu démontrer que notre nouvelle approche méthodologique, dénommée SPHERE, permettait de placer une source sonore à des positions 3D contrôlées, et donc accéder à tout l'espace acoustique autour du participant.

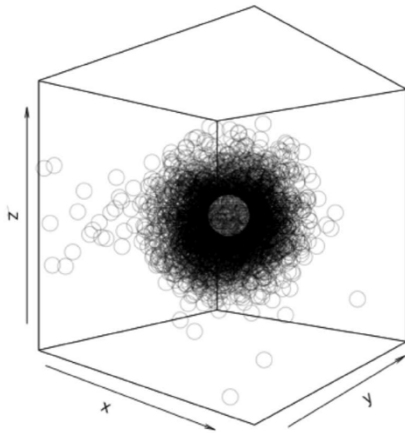


Figure 3. Erreurs cumulées des positions 3D réelles de la source sonore par rapport aux positions 3D désirées (alignement de toutes les positions, soit plus de 3900 essais cumulés).

Dans notre première approche méthodologique (Gaveau et al., 2022), nous avons interfacé deux systèmes : un système de RV et un système de capture du mouvement 3D. L'évolution des systèmes de réalité virtuelle offre désormais la possibilité d'utiliser un seul système qui assure à la fois la gestion de la RV et le contrôle des données de position 3D d'objets : ainsi la position de la tête est suivie par l'enregistrement de la position du casque de RV, la position 3D d'objets ('capteur') est intégrée de manière native dans le système de RV (par exemple, un 'capteur' positionné sur un haut-parleur permet de suivre la position 3D de ce dernier). Pour exploiter ces avancées technologiques, il a été nécessaire de valider l'usage de nouveau dispositif de RV 'tout en un' en fonction de nos exigences en matière de recherche sur la SA.

Remarque : comme ces nouveaux dispositifs de RV ont aussi l'avantage d'être commercialisés auprès du grand public, ils sont plus à même d'être intégrés dans des dispositifs commerciaux dédiés aux professionnels de santé (e.g. audioprothésistes, ORL, etc.) ; c'est cette stratégie évolutive que nous avons mise en avant en réalisant un partenariat avec la SATT PulsaLys, partenariat qui a visé la création d'un démonstrateur ('SPHERE application') en vue de la commercialisation du brevet SPHERE aux industriels de la santé.

☞ **Contribution= Validation de l'utilisation d'un système de RV pour les études neuroscientifiques (Verdelet et al. 2020)**

Le choix du système Vicon ® comme système de capture du mouvement dans l'étude précédente s'est basé sur sa validation en tant qu'outil scientifique pour la capture de signaux de position 3D. De nouveaux systèmes de

RV tels que le système HTC VIVE® (Pro) qui est commercialisé auprès du grand public pour le gaming, offrent la possibilité de capturer la position 3D réelle d'objets, tout en immergeant le participant dans des scènes visuelles virtuelles « écologiques » où les objets réels se retrouvent alignés avec les scènes. Le passage d'un système technique validé en recherche à un nouveau système conçu à des fins récréatives, a nécessité l'évaluation des capacités et limites spatiales et temporelles de ce nouveau système. Notre objectif a été de déterminer si ce nouveau système de RV commercial peut être utilisé à des fins de recherche. À cet égard, en collaboration avec les ingénieurs de la plateforme NeuroImmersion de Lyon, nous avons évalué les performances spatiales et temporelles de l'acquisition des données de position du système HTC VIVE® (Pro) en les comparant à celles du système de capture du mouvement en temps réel 3D que nous avons utilisé lors de la validation de SPHERE (Gaveau et al., 2022).

Résultat. Les résultats montrent que le système commercial HTC VIVE® (Pro) offre des mesures spatio-temporelles suffisamment précises et fiables, ce qui en fait une option convenable pour les études portant sur le comportement humain, notamment celles liées à la spatialisation auditive (SA).

2. Prérequis n°2 : importance du contrôle du regard et du contrôle des informations visuelles

Localiser les sons dans l'espace uniquement à l'aide d'indices auditifs constitue un défi pour le système auditif en raison de la nature complexe des scènes sonores du quotidien et des nombreux facteurs qui influent sur la propagation et la déformation des sons. C'est ici que les informations visuelles entrent en jeu. En effet, la représentation de l'espace est un phénomène résultant de l'intégration d'indices multisensoriels, qui en se combinant assure une perception cohérente de l'environnement (pour revoir voir le chapitre de (Gutfreund & King, 2012). En fusionnant les indices auditifs et visuels, les animaux peuvent améliorer leur capacité à localiser avec précision des objets ou des événements.

De nombreuses études ont mis en évidence la complémentarité des deux systèmes visuel et auditif dans les performances de spatialisation auditive (e.g. (Alais & Burr, 2004; Tonelli et al., 2015). Lorsque les indices visuels sont peu fiables (e.g. vision floue), le biais perceptuel est davantage influencé par les informations auditives (Burr & Alais, 2006). En revanche, en cas d'altération des entrées auditives (Knudsen & Knudsen, 1985) ou dans des conditions d'écoute difficiles (Bishop et al., 2012; H. E. Heffner & Heffner, 2014; R. S. Heffner & Heffner, 1992), ce sont les indices visuels qui sont privilégiés pour calibrer l'espace. Cette prédominance de la calibration de l'espace par les informations visuelles en cas de diminution des capacités auditives revêt une importance cruciale pour la localisation des sons, comme le démontreront mes études présentées dans le chapitre 4.

S'il est admis que les informations visuelles participent activement à la représentation acoustique de l'espace, de nombreux protocoles visant à étudier cette représentation se sont avérées peu fiables quant au contrôle des indices spatiaux visuels : certaines études ont permis une vision complète des sources sonores (van Hoesel & Tyler, 2003), tandis que dans d'autres, les sources sonores ont été dissimulées derrière des barrières visibles tout en offrant la possibilité de voir l'ensemble de la pièce (e.g., (Freigang et al., 2015; Nava et al., 2009; Pavani et al., 2001, 2003). Dans d'autres études, les participants ont effectué la tâche de spatialisation auditive dans le noir (Goossens & van Opstal, 1999; Van Barneveld et al., 2011; Van Grootel et al., 2011), les yeux fermés (Brungart et al., 1999), ou bandés (Ahrens et al., 2019; Bahu et al., 2016) sans qu'il soit possible de contrôler la direction du regard. Or, la position des yeux influence les performances de localisation auditive (Lewald & Ehrenstein, 1996) ainsi que l'intégration des indices acoustiques binauraux (Groh & Sparks, 1992; Pavani et al., 2008). Ce n'est que très récemment que les protocoles d'étude de la spatialisation auditive se sont attachés à contrôler les informations visuelles (Ahrens et al., 2019; Majdak et al., 2010), et ont montré qu'en manipulant ces informations, les performances de SA étaient impactées (Ahrens et al., 2019). Par conséquent, contrôler les informations visuelles accessibles au sujet tout en contrôlant son regard (i.e. posture oculo-céphalique) revêt une importance capitale dans la mise en place de protocoles d'étude de la SA.

Connaissant ces deux exigences méthodologiques, nous avons implémenté dans notre stratégie de développement méthodologique SPHERE un contrôle du regard tout en offrant la possibilité de contrôler aussi les indices visuels. Cette stratégie offre deux avantages : (1) non seulement elle permet au participant d'ajuster naturellement sa posture oculo-céphalique avant toute stimulation sensorielle, et ceci sans contrainte physique et en se servant de retours visuels le renseignant sur son état postural (Gaveau et al. 2022), mais (2) elle permet aussi de manipuler la structure visuelle de l'espace et ainsi étudier sa contribution sur les performances de SA (Valzolgher et al., 2020).

☞ **Contribution= Contrôle de la posture d'écoute (Gaveau et al., 2022)**

L'orientation du regard (direction combinée des yeux et de la tête) influant l'intégration des indices acoustiques, il a fallu contrôler cette orientation avant toute stimulation acoustique, et garantir que cette posture initiale d'écoute soit reproductible entre essais, sessions et participants. De plus, aucune contrainte matérielle ne devait être utilisée pour y parvenir afin de laisser la participant libre de bouger la tête lorsque de la stimulation sonore était délivrée. Pour répondre à cette exigence méthodologique, il a fallu s'appuyer sur la technologie des casques de RV qui disposent à la fois d'un système de suivi oculaire (tel que le HTC VIVE ® Pro, ou l'Oculus ® DK2) et d'un écran de stimulation permet de réguler les informations visuelles accessibles au participant.

Pour ce faire, nous avons défini une posture initiale de référence (droit devant postural) à laquelle le participant devait se conformer en début d'essai en utilisant des indices visuels sur la posture à atteindre et des retours visuels sur l'état de sa posture. La figure 5 illustre la chronologie des éléments visuels mis en jeu en fonction du comportement oculo-céphalique du participant : en début d'essai, le participant dispose d'indices visuels fixes dans son environnement matérialisant la posture initiale d'écoute à atteindre, et d'indices visuels dynamiques reflétant la position réelle de son regard (orientation des yeux et de sa tête) calculé en temps réel ; en se servant des indices dynamiques, le participant peut alors aligner sa posture avec celle de référence. Une fois ce contrôle de posture validée, la stimulation auditive peut alors commencer si le contrôle de la position de la source sonore est lui-même validé (voir chapitre précédent).

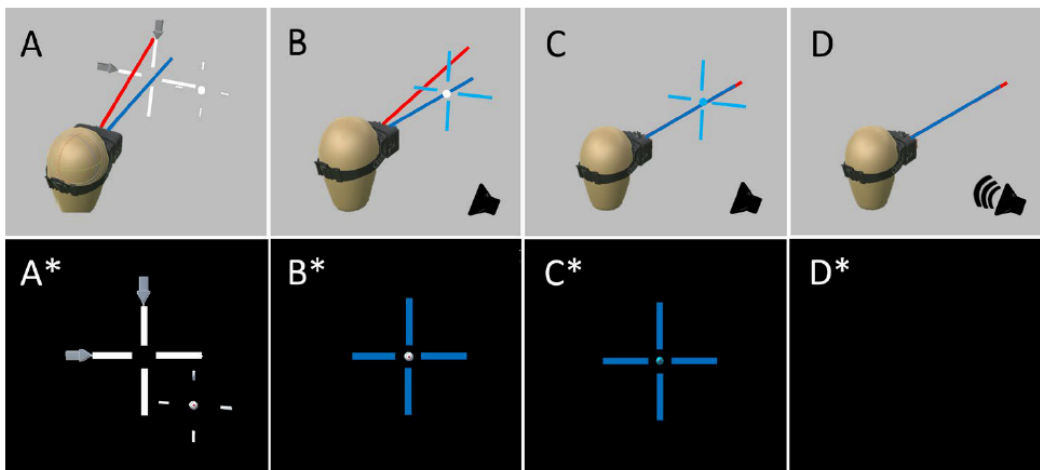


Figure 5. Alignement de la tête et des yeux avant la stimulation auditive. A : au début de chaque essai, le participant est libre de bouger sa tête (direction de la tête symbolisée par la ligne bleue) ainsi que ses yeux (direction de l'œil cyclopéen symbolisée par la ligne rouge). A* Affichage dans le casque de RV : une croix blanche en gras indique la position réelle de la tête, et une croix blanche en traits fins avec un point central (point de fixation) indique respectivement la position souhaitée de la tête et des yeux ; deux flèches en gris entourant la croix blanche en gras, montrent au participant dans quelle direction bouger la tête pour atteindre la position initiale souhaitée. B* Lorsque la position de la tête est alignée avec la direction souhaitée, celle-ci devient bleue. C* Lorsque les yeux se positionnent sur le point de fixation, ce point devient bleu. D Lorsque la direction du regard (tête et yeux) sont alignés et que le haut-parleur est à sa position prédéterminée, toutes les indices visuels disparaissent, D* le sujet n'a plus d'indice visuel pour réaliser sa tâche de SA.

Résultat. Les données d'orientation des yeux et de la tête étaient contrôlées en temps réel par le programme expérimental. Pour valider la qualité de ce contrôle de la posture initiale d'écoute, une vérification manuelle a été effectuée off line lors du développement de SPHERE, impliquant une analyse cinématique des données de position de la tête et des yeux. Après validation de ce contrôle ainsi que du contrôle de la position de la source sonore, le processus d'acquisition des données (positions de la tête, yeux et main) était initié. Les stratégies d'exploration oculaire que les participants ont déployés pendant l'émission du son n'ont pas été soumises à un

examen approfondi susceptible d'être publié, mais les stratégies d'exploration céphaliques ont été un élément clef sur lequel nous avons formulé nos hypothèses d'étude (voir Gaveau et al., 2022).

Donc, en se basant sur des retours visuels dynamiques de la direction du regard, on peut aisément contrôler la posture initiale d'écoute des participants sans les contraindre par un dispositif de contention ; cette approche est d'autant plus 'simple' à mettre en place qu'elle a pu être déployée auprès d'enfants (voir Coudert et al., 2021).

☞ **Contribution= Impact de la scène visuelle dans les performances de SA (Valzolgher et al., 2020)**

En exploitant le potentiel de notre approche SPHERE dans le contrôle des informations visuelles, je voulais examiner l'influence des indices visuels structurant l'espace sur les performances de localisation auditive. Mon but était de montrer en quoi le contrôle de ces indices visuels structurant influent sur les performances de SA. Pour ce faire, j'ai mis l'accent sur l'effet d'un cadre spatial visuel minimal, c'est-à-dire des indices visuels sans aucun lien avec la position de la source sonore, et testé l'hypothèse que la localisation d'une source sonore s'appuie entre autres sur une représentation de l'espace élaborée à partir d'indices visuels structurant. Ainsi, notre hypothèse opérationnelle sous-jacente était que l'ajout d'un environnement visuel structuré pourrait améliorer la capacité de localisation sonore par rapport à une scène visuelle uniforme et non structurée (les deux types de scènes visuelles sont illustrés dans la figure 6). C'est donc en manipulant la qualité des indices visuels de la scène présentés au participant pendant son écoute que j'ai pu vérifier cette hypothèse.

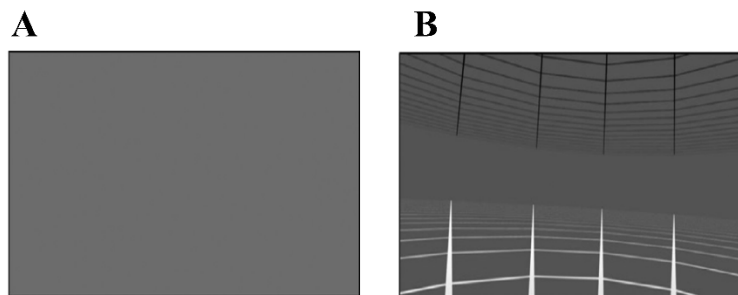


Figure 6 : Différents environnements visuels utilisés en réalité virtuelle lors d'une tâche de SA : (A) scène visuelle uniforme et non structurée (gris) ; (B) environnement visuel structuré (grille).

Résultat. La présence d'un cadre visuel structuré a conduit à une amélioration des performances de localisation auditive dans la dimension de l'élévation. De plus, ces informations visuelles structurées ont entraîné des mouvements d'orientation céphalique plus rapides par rapport à l'absence d'informations visuelles structurant l'environnement. Donc, localiser une source sonore est favorisée par la structuration de l'espace, cette dernière s'appuyant entre autres sur des informations visuelles.

3. Prérequis n°3 : Favoriser l'écoute active

La construction d'une représentation de l'espace auditif repose sur la compréhension des relations entre les indices acoustiques et la perception de cet espace. Cette correspondance se développe à travers les interactions entre le sujet et son environnement, en utilisant l'exploration active et en interagissant avec les retours sensoriels résultant de ces interactions, tels que les changements acoustiques provoqués par les mouvements de la tête (Aytekin et al., 2008).

Dans des situations naturelles, on observe que les individus tournent spontanément leur tête et leurs yeux vers les sources sonores (Fuller, 1992; Kim et al., 2013). Les mouvements de tête pendant l'émission d'un son modifient la position des oreilles par rapport aux sources sonores, créant ainsi des informations acoustiques dynamiques par changements des entrées binaurales et monaurales (Macpherson, 2015). Ces informations enrichissent les indices acoustiques et aident à résoudre les ambiguïtés perceptuelles liées à la localisation sonore.

Malgré l'importance de ces mouvements céphaliques dans la localisation sonore humaine, il est courant de restreindre les mouvements de tête et de demander aux sujets d'effectuer des tâches de localisation dans des conditions d'écoute statique. Le choix méthodologique de réduire ou limiter l'écoute active est justifié : il assure la reproductibilité des positions des sonores entre les participants et les sessions. De nombreuses études de SA ont eu recours à l'immobilisation passive de la tête des sujets à l'aide de mentonnière (Macpherson, 2015), ou à l'immobilisation active en demandant aux sujets de ne pas bouger la tête pendant l'écoute (Rabini et al., 2019; Van Wanrooij & Van Opstal, 2005). Certaines études ont même réduit la durée de la stimulation sonore pour limiter l'impact des mouvements de tête pendant l'émission sonore (Ahrens et al., 2019). Ces conditions d'écoute statique sont devenues la norme en matière de localisation sonore, et les mesures obtenues dans ces conditions expérimentales sont implémentées dans des réseaux neuronaux profonds pour les entraîner à la localisation des sons (Francl & McDermott, 2022).

Pourtant, des études ont montré que l'exploration active par des mouvements de tête peut réduire les erreurs de localisation auditive (Andéol & Simpson, 2016; Brimijoin et al., 2013; Honda et al., 2013; McAnally & Martin, 2014; W. G. Noble, 1981; Perrett & Noble, 1997a, 1997b; Pollack & Rose, 1967; Thurlow & Runge, 1967; Vliegen & Van Opstal, 2004; Wightman & Kistler, 1999). Ainsi l'écoute active contribue à réduire les confusions de localisation de sources sonores entre les espace avant et arrière (McLachlan et al., 2023; Yost et al., 2020), et se révèle être une stratégie efficace pour localiser des sons dans des environnements bruyants (Brimijoin & Akeroyd, 2012). Malgré ce constat, l'impact des comportements d'exploration céphalique (ou écoute active) dans la SA a été relativement négligée, peut-être en raison du manque de technologie permettant de contrôler les mouvements de tête.

En tenant compte de ces observations, il est devenu évident que l'étude des stratégies d'exploration céphalique devait être considérée comme un facteur modulant les performances de localisation sonore. Ces stratégies

naturelles et spontanées semblent mieux refléter les performances réelles des individus dans des situations de la vie quotidienne où les ambiguïtés auditives sont omniprésentes (bruit d'ambiance, effets de réverbération, etc.). En outre, ces comportements d'exploration dirigés vers des positions spatiales auditives, pourraient également être un élément clef pour l'amélioration des performances de localisation auditive, en particulier en cas de déficit auditif.

Étant donné que notre dispositif SPHERE nous permet d'accéder aux comportements d'exploration céphalique, je vais présenter comment j'ai procédé à la validation de l'approche méthodologique SPHERE comme outil de choix pour l'étude des comportements d'exploration céphalique (écoute active ou stratégie céphalique) lors de tâches de localisation auditive.

☞ **Contribution= l'écoute active favorise les performances de SA (Coudert et al., 2021 ; Gaveau et al., 2022)**

Lorsque nous avons répertorié les éléments méthodologiques à développer pour étudier la SA, l'idée de permettre aux sujets d'explorer librement leur environnement 3D en utilisant des mouvements de la tête a été l'un des éléments cruciaux de notre approche développementale. En portant le casque de réalité virtuelle, les sujets avaient la possibilité de déplacer leur tête pour explorer leur espace auditif, et nous étions en mesure de quantifier cette stratégie ainsi que son impact sur les performances de localisation sonore. Notre hypothèse de départ était que l'écoute active, caractérisée par les mouvements de la tête, par opposition à une écoute statique, pourrait améliorer les performances de localisation sonore. Il restait à confirmer cette hypothèse en effectuant une validation externe de notre approche méthodologique.

Pour répondre à cette validation, deux études ont été menées : l'une avec des adultes (Gaveau et al., 2022) et l'autre dirigée vers une population pédiatrique (Coudert et al., 2021). Dans ces deux études, les participants devaient localiser une source sonore selon deux modalités d'écoute différentes : soit ils étaient autorisés à explorer librement leur environnement en déplaçant leur tête pendant l'émission sonore (écoute active), soit il leur était demandé de maintenir leur tête dans la posture initiale d'écoute (écoute statique). Les performances de SA dans ces deux conditions d'écoute ont été alors comparées afin de déterminer quelle situation d'écoute était la plus favorable pour localiser la position d'une source sonore.

Résultat. Chez les adultes, j'ai clairement démontré que l'écoute active avait un effet bénéfique sur les performances de localisation sonore. Cette écoute a conduit à une amélioration de la précision et de la cohérence des réponses dans les dimensions azimutale et d'élévation. De plus, il a été établi qu'une stratégie impliquant des mouvements spontanés de la tête était corrélée à de meilleures performances de localisation sonore (voir figure 7).

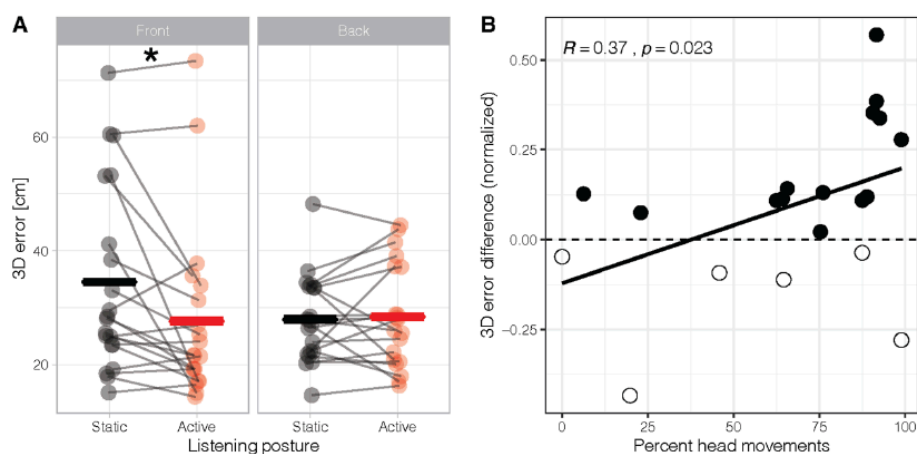


Figure 7. A. Performances 3D de SA des sujets adultes (erreur entre les positions 3D de la main et de la source sonore) dans l'espace avant (front) et arrière (back) dans les conditions d'écoute active (noir) et statique (rouge). B. Nuage de points représentant la différence dans les erreurs en 3D entre les conditions d'écoute active et statique (différence normalisée basée sur la performance d'écoute statique), en fonction du pourcentage de mouvements de la tête. Les cercles remplis indiquent les participants dont la performance de localisation sonore s'est améliorée en écoute active par rapport à statique ; les cercles vides indiquent les participants dont la performance de localisation sonore a diminué en écoute active par rapport à statique

En ce qui concerne les performances des enfants, il n'a pas été observé d'amélioration significative de leur performance de SA en situation d'écoute active par rapport à une écoute statique. Cependant, il est à noter que les performances de localisation sonore des enfants en écoute statique étaient équivalentes à celles de la population adulte en condition d'écoute active, rendant toute amélioration par une stratégie d'écoute active limitée. L'avantage que nous avons relevée chez les adultes pourrait venir de la constitution de notre population d'adultes. En effet, cette dernière se voulait être appariée en âge avec des adultes présentant des déficits auditifs profonds et porteurs d'implants cochléaires dont l'âge moyen était de 50 ans. Or il est bien établi que la capacité auditive diminue avec l'âge, affectant notamment l'identification des hautes fréquences, un phénomène connu sous le nom de presbycusis (Gates et al., 2000; Pearson et al., 1995). Cette diminution peut influencer la localisation en azimute (Rakerd et al., 1998) et en élévation (W. Noble et al., 1994). Dans notre étude chez l'adulte, bien qu'aucun déficit auditif n'ait été identifié chez nos participants normo-entendant, mais il n'est pas possible d'exclure la présence de déficits auditifs infraliminaires chez certains participants (notamment parmi ceux âgés de plus de 50 ans) qui aurait pu conduire à une réduction de leur performance de SA. En outre, il est intéressant de noter que les erreurs de localisation en trois dimensions chez les enfants étaient négativement corrélées avec l'âge au moment du test : plus l'enfant était âgé, plus les erreurs de localisation en trois dimensions étaient faibles. Cela révèle que la capacité de localisation auditive se développe tout au long de la croissance, et qu'il serait intéressant d'en étudier la mise en place chez l'enfant (voir chapitre 5- Perspectives).

4. Prérequis n°4 : Favoriser les interactions avec l'espace

Interagir avec des cibles sonores n'est que très peu étudiée en SA, la très grande majorité des études mobilise des sources sonores en dehors de l'espace d'atteinte manuelle. Or cet espace est le théâtre d'interactions multisensorielles avec les objets proches et les événements auditifs qui s'y déroulent permettent d'élaborer des réponses motrices immédiates, surtout si ces événements sont menaçants ou particulièrement intéressants (pour revue (Serino, 2019). De plus, différentes études ont montré que les stratégies d'exploration de l'environnement sonore à l'aide de mouvements simultanément de la tête et de la main peut améliorer la localisation sonore quand les indices acoustiques sont altérés (Honda et al., 2007, 2013; Parseihian et al., 2014), supportant l'idée que les mécanismes de localisation sonore sont très plastiques (Kral & Sharma, 2023).

Dans l'étude précédente, j'ai pu montrer que les interactions du sujet avec son environnement sonore (écoute active) lui permettaient d'affiner son processus de localisation (i.e. assuré alors une perception inférentielle). Le sujet se base alors sur des correspondances établies entre son espace auditif et son espace d'action (modèle interne) pour performer. La mise à jour de ce modèle interne fait appel à des retours sensori-moteurs, assurant le maintien d'un alignement efficace entre l'espace de perception auditive et l'espace d'action. En cas de déficit auditif, ce modèle est modifié et adapté aux nouvelles contraintes par un processus d'adaptation. Cependant, il reste une ambiguïté sur le type d'interaction qui favorise cette re-spatialisation des sons : l'exploration par les mouvements de tête (écoute active) est-elle suffisante pour recalibrer son espace auditif ? Si les stratégies d'exploration manuelle permettent une recalibration auditive, quel est le rôle des interactions manuelles directes du sujet avec son espace dans cette recalibration ?

☞ **Contribution= interagir dans son espace acoustique favorise la re-spatialisation auditive (Valzolgher et al., 2020)**

Dans cette étude, l'objectif était d'identifier si l'écoute active pouvait être aussi impliquée dans l'adaptation de la représentation spatiale auditive, et si cette adaptation était favorisée par des actions dirigées spatialement dans l'espace auditif. Pour induire une adaptation des performances de localisation sonore, des participants normo-entendant ont été soumis à une tâche de localisation sonore dans deux conditions d'écoute : une condition d'écoute normale (binaurale - condition de contrôle) et une condition d'écoute altérée (monaurale) induite par le port d'un bouchon d'oreille. Dans la situation d'écoute normale, l'exploration céphalique était impliquée dans la correction d'erreurs perceptives (erreurs évaluées par le modèle interne). Dans la situation d'écoute altérée, l'asymétrie auditive induite par le bouchon d'oreille avait pour but de favoriser l'adaptation et d'étudier la dynamique de l'adaptation spatiale auditive. Cette adaptation nécessite une mise à jour du modèle interne par des retours sensori-

moteurs. Afin d'identifier la contribution des stratégies d'exploration dans cette recalibration auditive, deux modalités de réponse ont été testées : soit les participants effectuaient des mouvements de la main en direction de la source sonore (action spatialisée), soit ils devaient simplement nommer la source sonore (action non spatialisée). Dans les deux conditions de réponse, les participants étaient autorisés à bouger librement la tête pendant la tâche et recevaient des retours sensoriels (auditifs et visuels) de leur performance.

Résultats. Les résultats ont révélé que dans la condition d'écoute altérée, quelle que soit la modalité de réponse (spatialisée ou non spatialisée), les performances de localisation sonore s'amélioraient au fil des essais, ce qui suggère une adaptation de la représentation auditive spatiale suite à une altération auditive. De plus, il a été constaté que cette amélioration était plus rapide et plus significative lorsque les participants réalisaient des mouvements d'atteinte en direction de la source sonore, par rapport à la simple dénomination de la position de cette dernière. En outre, relocaliser la source sonore à l'aide d'une réponse manuelle spatialisée semblait influencer la stratégie d'exploration de la tête, en induisant des mouvements plus amples en direction de la source auditive.

Remarque : Cette étude, qui exploitait les potentialités de l'approche SPHERE, a été publiée en 2020. Cependant, SPHERE, en tant que méthodologie novatrice, n'a été acceptée pour publication que deux ans plus tard (1^{ère} soumission en mars 2020). Ce n'est qu'après la publication de plusieurs articles montrant le bénéfice de cette méthodologie dans l'étude des déficits de la SA, que la communauté scientifique a reconnu les avantages de cette méthodologie pour l'étude des capacités de localisation sonore.

CONCLUSION du CHAPITRE 1

L'approche méthodologique SPHERE présentée dans ce chapitre est le fruit d'une étroite collaboration entre ingénieurs et chercheurs. Elle avait pour objectif d'offrir une nouvelle approche méthodologique de la spatialisation auditive en combinant différents éléments identifiés dans la littérature scientifique comme impliqués dans le processus de SA. En combinant ces éléments et en assurant leur contrôle, on ouvre le champ des possibles vers de nouvelles études de la spatialisation auditive qui deviennent plus 'écologiques' à travers le prisme des interactions sensori-motrices.

L'utilisation de nouvelles méthodologies en neurosciences est en plein essor et offre une approche écologique permettant de 'sortir' des environnements de laboratoire traditionnels pour sa rapprocher des événements se produisant naturellement dans la vie quotidienne (pour revue (Cornelio et al., 2021)). Toute innovation technologique en neurosciences demande à être validée et proposée au regard critique de la communauté scientifique pour que celle-ci puisse en apprécier les enjeux. Ainsi, nous avons pu montrer que notre approche offrait toutes les garanties pour réaliser des études 'écologiques' de la spatialisation auditive : elle nous permet de positionner des sources sonores de manière contrôlée et reproductible dans tout l'espace sans imposant de contrainte technique aux chercheurs tout en contrôlant la posture d'écoute des participant. Elle nous permet également de recueillir les performances de spatialisation auditive du participant (orientation de sa tête-performance 2D ; position de la main - performance 3D).

En outre, cette approche présente d'autres avantages pour l'étude de la spatialisation auditive. L'utilisation d'un système de réalité virtuelle assurant le contrôle des informations visuelles accessibles au sujet, permet de répondre d'une part à des exigences méthodologique (e.g. contrôle de la posture d'écoute) et d'autre part à des questionnements sur le rôle des intégrations visuo-auditives en SA. Ainsi, une structure visuelle simple de l'espace impacte les performances de localisation auditive, renforçant la notion que la capacité de localisation auditive doit être appréhendée dans le contexte de la multisensorialité, la vision jouant un rôle crucial au processus de localisation auditive.

De plus, mes différentes études soulignent l'importance de contrôler et de promouvoir les stratégies d'exploration oculo-manuelle. En permettant aux participants d'adopter une approche d'écoute naturelle, on peut étudier les stratégies exploratoires des sujets quand ceux-ci doivent identifier avec précision des éléments sonores dans leur espace, avantage qui devient pertinent quand on doit évaluer les capacités de localisation auditive de

sujets confrontés à des ambiguïtés perceptuelles tels que les patients souffrant de déficits auditifs (voir chapitre suivant). On peut aussi étudier comment ces stratégies participent à l'adaptation de la représentation spatiale, approche qui s'avère pertinente lorsqu'on s'intéresse à la dynamique des représentations spatiales de patients présentant des déficits de spatialisation auditive et pour lesquels il n'existe pas de stratégie de rééducation (voir chapitre suivant).

Enfin, SPHERE possède une architecture évolutive, offrant la possibilité de déployer cette approche vers d'autres champs d'expertise, notamment dans le domaine clinique de l'audioprothèse (Barone et al., 2018; Pavani et al., 2018) ainsi qu'en direction des industriels spécialistes de l'audition (e.g. projet industriel PulsaLys visant à commercialiser le brevet SPHERE).

CHAPITRE 2. Évaluation et rééducation de la SA & déficits auditifs

Supervisions d'étudiants

- Fombertasse L. (Formation orthophonie). Master 2 Orthophonie
- Bonnamy L. (Formation orthophonie). Master 2 Orthophonie
- Bouzaid S. (Audioprothésiste). Master 2 Santé - Handicaps Sensoriels et Cognitif / ISTR UCBL
- Grenouillet S. (Kinésithérapeute). Master 2 Handicap neurologique - Handicap neurologique / Paris Sarclay
- Coudert A. (Formation médicale). Doctorat de Neurosciences. UCBL

Articles liés au projet

- **Coudert A, Gaveau V, Gatel J, Verdelet G, Salemm R, Farné A., Pavani F., Truy E.** Spatial hearing difficulties in bilateral cochlear implant children improve with head movements. *Ear and Hearing*. 2022 Jan/Feb;43(1):192-205.
- Valzolgher C, Todeschini M, Verdelet G, Gatel J, Salemm R, **Gaveau V**, Truy E, Farné A, Pavani F. Adapting to altered auditory cues: Generalization from manual reaching to head pointing. *PLoS One*. 2022 Apr 14;17(4):e0263509.
- Valzolgher C, Gatel J, **Bouzaid S, Grenouillet S**, Todeschini M, Verdelet G, Salemm R, **Gaveau V**, Truy E, Farné A, Pavani F. Reaching to Sounds Improves Spatial Hearing in Bilateral Cochlear Implant Users. *Ear Hear*. 2023 Jan-Feb 01;44(1):189-198.
- **Coudert A**, Verdelet G, Reilly KT, Truy E, **Gaveau V**. Intensive Training of Spatial Hearing Promotes Auditory Abilities of Bilateral Cochlear Implant Adults: A Pilot Study. *Ear Hear*. 2023 Jan-Feb 01;44(1):61-76.
- Valzolgher C, **Bouzaid S, Grenouillet S**, Gatel J, Ratenet L, Murenu F, Verdelet G, Salemm R, **Gaveau V**, Coudert A, Hermann R, Truy E, Farné A, Pavani F. Training spatial hearing in unilateral cochlear implant users through reaching to sounds in virtual reality. *Eur Arch Otorhinolaryngol*. 2023 Mar 11.
- **Coudert A**, Reilly KT, Truy E, Moulin A, **Gaveau V**. A new tool for the assessment of speech understanding and spatial hearing difficulties in children: the Kid-SSQ questionnaire. *Eur Arch Otorhinolaryngol*. 2023 Feb 10.

Prix liés au projet

- **Coudert A.** Prix 1^{er} Poster – OTOForum (Grenoble). Rôle des mouvements de tête dans la perception spatiale des sons chez les patients implantés cochléaires.
- **Bouzaid S.** Prix chercheur - Société Française d'Audiologie 2022. Évaluation de l'effet d'un protocole d'entraînement visio-auditif sur la localisation spatiale chez des sujets bi-implantés cochléaires.

Financements du projet

- ANR-16-CE17-0016.
- LabEx CORTEX SATT (ANR-11-LABX-0042)
- Institut rhônalpin des Systèmes complexes (IXXI)
- SPHERIC IIR-1948

Collaborations

- Services d'ORL des Hospices Civils de Lyon (Lyon, France)
- Plateforme Neuroimmersion du Centre de Recherche en Neurosciences de Lyon (Lyon, France)
- Center for Mind/Brain Sciences (CIMeC), University of Trento, (Rovereto, Italie)
- Centre de Recherche Cerveau & Cognition, (Toulouse, France)
- Equipe 'Trajectoire' du CRNL (Lyon, France)
- Equipe 'Perception Attention Mémoire' du CRNL (Lyon, France)

Rationnel

L'audition spatiale est souvent absente dans le diagnostic et la prise en soin des patients souffrant de pertes auditives profondes. Les recherches sur l'intégration auditive suggèrent que la localisation précise des sons peut être une condition préalable à l'orientation de l'attention, qui à son tour, est essentielle pour d'autres fonctions cognitives fondamentales telles que la compréhension de la parole dans le bruit (pour revue (Bronkhorst, 2015; Gatehouse & Noble, 2004). Cependant, avec la croissance du nombre de personnes ayant des déficits auditifs et l'impact de ces déficits sur diverses fonctions cognitives, il devient primordial d'évaluer les capacités de localisation sonore et d'étudier les facteurs qui facilitent l'adaptation de cette capacité.

Les patients atteints de pertes auditives profondes peuvent bénéficier de dispositifs implantables tels que les implants cochléaires (IC, pour revue Wilson, 2008)) qui partiellement rétablissent la fonction auditive (Moore & Shannon, 2009), celle du langage ((Zhang et al., 2010) et améliore leur qualité de vie (Bergman et al., 2020; Sladen et al., 2017). Ces dispositifs captent les sons de l'environnement à l'aide de microphones, traitent les signaux acoustiques et les convertissent en signaux électriques qui sont ensuite transmis directement à la cochlée via des électrodes implantées chirurgicalement. Bien que ces dispositifs soient principalement axés sur la compréhension du langage (pour revue Wilson, 2008) et améliorent la compréhension de la parole, ils altèrent considérablement les indices spatiaux nécessaires à la localisation sonore (Grantham et al., 2007a; Nava et al., 2009; Van den Bogaert et al., 2006).

1. Performance de SA des patients porteurs d'implants cochléaires

Plusieurs études ont montré que les patients porteurs d'IC présentent des déficits de localisation sonore (Akeroyd & Whitmer, 2016; Grantham et al., 2007b; H. Jones et al., 2014; van Hoesel & Tyler, 2003). Ces déficits de localisation peuvent être expliqués en partie par les stratégies de codage sonore des prothèses auditives. En effet, les traitements acoustiques des IC altèrent les indices binauraux tels que les différences temporelles interaurales (ITD) et les différences d'intensité interaurales (ILD) en perturbant la structure temporelle des sons, modifient la synchronisation des signaux et compriment l'intensité acoustique (via le traitement « Automatic Gain Control ») (Akeroyd & Whitmer, 2016). Ces effets délétères peuvent également être exacerbés par l'asynchronie des traitements acoustiques, particulièrement lorsque le patient possède deux implants de technologies différentes.

Si ces déficits sont connus, il n'existe pas d'études évaluant les performances de localisation sonore chez les patients IC dans les trois dimensions spatiales (beaucoup d'études se limitent à l'espace frontal ou à la dimension azimutale). De plus, peu d'études contrôlent rigoureusement les indices visuels ni ne favorisent des conditions d'écoute active (voir chapitre précédent). Pourtant, tous ces facteurs sont essentiels pour évaluer les

performances de localisation auditive de patients, et ainsi optimiser une stratégie de rééducation de la localisation sonore.

☞ **Contribution= Évaluation des capacités de localisation auditive des patients porteurs de deux prothèses implantables auditives dans les 3 dimensions de l'espace, bénéfice de l'écoute active (Coudert et al., 2021)**

L'évaluation des capacités de localisation sonore des patients IC s'est appuyée sur l'approche immersive de SPHERE (stimulation acoustique 3D, contrôle visuel et analyse des stratégies d'exploration active). Dans cette étude, nous voulions identifier et caractériser les déficits de localisation sonore chez les enfants porteurs de deux implants cochléaires (bIC), et apprécier s'ils pouvaient tirer profit de l'écoute active pour corriger les ambiguïtés auditives dues aux traitements de leur IC. Pour cela, nous avons procédé à une comparaison de leurs capacités avec des enfants normo-entendants (groupe contrôle) en utilisant la même méthodologie présentée dans le chapitre précédent : les enfants devaient localiser à l'aide de leur main une source sonore sans avoir ni d'indice sur la structuration visuelle de leur espace ni de retour sur leur performance.

Résultats. Les enfants bIC ont été capables de distinguer les sources sonores situées à gauche et à droite (grâce aux stratégies de leurs implants cochléaires qui leur permettaient d'accéder aux indices binauraux). Cependant, ils ont rencontré des difficultés importantes dans la localisation avant-arrière et ont souvent perçu les sources à proximité de l'axe interaural (coïncidant avec la position des microphones de leurs implants cochléaires). De plus, ils ont éprouvé des difficultés à localiser les sources sonores en profondeur, en raison des stratégies de codage propres à leurs implants cochléaires. Lorsque l'on considère leur performance de localisation dans les trois dimensions de l'espace, leur erreur 3D (qui tient compte de leur performance dans les 3 dimensions de l'espace) en était nettement plus grande que celle de leurs pairs normo-entendants. Cependant, grâce à l'écoute active, ces enfants bIC ont significativement réduit leur erreur 3D. De plus, nous avons pu mettre en évidence une corrélation entre l'intervalle entre les deux implantations (c'est-à-dire le délai entre l'implantation du premier et du deuxième implant) et l'erreur en 3D.

2. Contribution= Modèle d'entraînement à la spatialisation auditive suite à une altération de l'écoute binaurale et généralisation à l'espace sonore (Valzolgher et al., 2022)

Un autre facteur majeur limitant les capacités de localisation auditive des patients implantés provient de la configuration des dispositifs auditifs. En effet, un patient avec deux implants a accès à une audition binaurale et peut se reposer sur des indices binauraux pour localiser des sons. Cependant, avec un seul implant, le patient est en situation d'écoute monaurale, perdant de surcroît des indices binauraux. En conséquence ses performances

de localisation auditive sont dégradées (Grantham et al., 2008; Van den Bogaert et al., 2006, 2006). Cette asymétrie auditive peut être artificiellement créée chez des sujets normo-entendants pour créer un modèle de déficience auditive binaurale (voir chapitre précédent Valzolgher et al., 2020) et ainsi offre l'opportunité d'étudier les éléments-clefs nécessaire à une re-spatialisation auditive en cas d'écoute monaurale.

Il est bien établi que les mécanismes de localisation sonore sont plastiques (Kral & Sharma, 2023). Cette plasticité peut être renforcée par des informations multisensorielles convergentes tels que les indices audio-visuels (Kumpik et al., 2019; Strelnikov et al., 2011) mais aussi par l'exploration active de l'environnement (voir Valzolgher et al., 2020). Afin de proposer une approche de re-spatialisation auditive à des patients implantés cochléaires, il faut avant tout déterminer les facteurs les plus à même d'induire cette plasticité auditive spatiale, et identifier si cette plasticité peut se généraliser à d'autres fonctions cognitives.

Dans l'étude présentée dans le chapitre précédent (Valzolgher et al. 2020), nous avons démontré qu'en utilisant un modèle de déficience auditive (sujets normo-entendants en situation d'écoute monaurale), les mouvements d'atteinte favorisaient la re-spatialisation des sons. Dans cette étude, nous avons souhaité vérifier si cet effet d'entraînement pouvait s'étendre à l'espace non entraîné et avoir un impact sur d'autres fonctions cognitives basées sur l'identification des sources sonores, comme l'attention audio-visuelle (ou capture automatique de l'attention visuelle par un son spatialisé).

Ainsi, des sujets normo-entendants dont l'audition binaurale a été altérée (voir Valzolgher et al. 2020) ont été soumis à deux types d'entraînement différents et séparés dans le temps : soit (1) ils devaient localiser une source sonore spatialisée autour d'eux en effectuant un mouvement d'atteinte vers cette source (entraînement spatial), avec un retour audio-visuel spatialisé de leur performance, soit (2) ils devaient identifier la fréquence du son émis par cette source sans la localiser (discrimination de fréquence) en effectuant un mouvement d'atteinte vers une position frontale (entraînement non spatial), avec un retour audio-visuel non spatialisé de leur performance. Ainsi, les participants se focalisaient soit sur les caractéristiques spatiales, soit sur les caractéristiques non spatiales du son. Notre hypothèse était qu'un entraînement spatial serait le plus à même d'induire une adaptation des capacités de localisation sonore, et que cette adaptation pourrait se propager à d'autres dimensions spatiales et à d'autres fonctions cognitives liées à l'audition spatiale.

Pour tester notre hypothèse, différents tests ont été réalisés avant l'altération de l'audition (en condition d'écoute binaurale), immédiatement après altération de l'audition (en condition d'écoute monaurale) et après chaque session d'entraînement. Ces tests nous ont permis (1) d'évaluer la généralisation des effets de l'entraînement à l'espace auditif non entraîné, et (2) de mesurer l'effet de l'entraînement sur une autre fonction cognitive utilisant la localisation sonore (attention audio-visuelle).

Résultat. L'entraînement spatial, basé sur une exploration manuelle, a entraîné une amélioration des performances de localisation sonore qui s'est généralisée à l'espace non entraîné. De plus, il a conduit à des changements dans

la stratégie d'exploration active, notamment des mouvements céphaliques plus rapides et plus précis en direction de la source sonore. Cependant, aucun effet sur l'attention n'a été observé, quelle que soit la nature de l'entraînement. Cette étude démontre que l'entraînement spatial de la localisation sonore a un effet généralisable et impacte la stratégie d'exploration céphalique, un mécanisme crucial pour enrichir les indices auditifs.

3. Contribution= Entraînement à la spatialisation auditive de patients implantés cochléaires (Valzolgher et al. 2023a, 2023b)

Les patients implantés cochléaires bilatéraux (porteurs de deux implants cochléaires - bIC) ont accès à une audition spatiale binaurale (dans la mesure où la qualité des indices binauraux générés par les ICs le permet). Les patients uni-implantés (porteurs d'un seul implant -uIC) et ont en conséquence une audition spatiale binaurale fortement altérée et ceci même s'ils portent une aide auditive (cette dernière ne résout pas les asymétries auditives). Nous avons cherché dans un premier temps à identifier les performances de SA de ces deux groupes de patients, puis dans un second temps à connaître dans quelle mesure un entraînement spatial pouvait améliorer leurs performances de localisation sonore et moduler leurs attention audio-visuelle.

Pour réaliser cette étude, nous avons appliqué le même protocole d'entraînement à la localisation sonore que celui développé pour les sujets normo-entendants (Valzolgher et al. 2022) à des adultes implantés bi- ou monocochléaires.

Résultats. Grâce à un entraînement spatial, les patients bIC et uIC ont réussi à améliorer leurs performances de localisation sonore, et cet effet s'est étendu à des positions auditives qui n'avaient pas été directement entraînées. En outre, nous avons constaté une modification de leur stratégie d'exploration céphalique, caractérisée par des mouvements d'orientation de la tête plus précis. Cependant, aucun bénéfice en termes d'entraînement spatial n'a été observé dans la tâche impliquant un traitement attentionnel audio-visuel, que ce soit pour les patients bIC ou uIC. De manière inattendue, un avantage a été observé après un entraînement non spatial pour les patients bIC. Une hypothèse avancée est que les patients bIC se seraient davantage engagés dans le traitement d'attributs non spatiaux, un engagement favorisé par l'entraînement non spatial.

4. Contribution= Entraînement intensif à la spatialisation auditive de patients implantés cochléaires, et généralisation à d'autres fonctions cognitives (Coudert et al. 2022)

Si l'étude de re-spatialisation auditive menée chez les patients bIC a démontré qu'il était possible de modifier les performances de localisation sonore ainsi que les stratégies d'exploration en une seule séance d'entraînement (Valzolgher et al. 2023), cet effet n'a toutefois pas pu être généralisé à une autre fonction cognitive (en l'occurrence, l'attention audio-visuelle) et les bénéfices n'ont pas pu être évalués sur le long terme. Dans une

nouvelle étude, nous voulions déterminer si un entraînement spatial intensif (composé de 8 séances, avec une fréquence de 2 séances par semaine) pouvait engendrer des modifications au niveau d'une fonction cognitive déficiente chez ces patients, à savoir la compréhension de la parole en environnement bruyant. De plus, nous souhaitons évaluer si cet entraînement pouvait également améliorer leur qualité d'écoute, mesurée à l'aide de l'échelle de qualité d'écoute SSQ 15 (Moulin et al., 2019). Par ailleurs, nous nous sommes intéressés à la durabilité de cet effet au-delà de la période d'entraînement, en réalisant des mesures un mois après la fin de l'entraînement.

Résultat. Nos résultats ont révélé une amélioration significative des performances de localisation sonore chez les patients bIC, ainsi qu'un bénéfice observé dans la compréhension de la parole en situation de bruit ainsi qu'une amélioration de leur qualité d'écoute. Ces effets bénéfiques se sont maintenus de manière significative un mois après la fin de la période d'entraînement.

5. Contribution= Évaluation des capacités spatiale auditive de patients implantés cochléaires par un questionnaire à passation rapide (Coudert et al. 2023)

Étant donné que nous étions en mesure de mesurer objectivement les performances de localisation sonore chez les adultes et les enfants porteurs d'implants cochléaires (voir contributions précédentes), nous avons considéré comme essentiel de replacer ces résultats dans un contexte clinique plus large en tenant compte des évaluations subjectives des patients concernant leurs capacités de localisation sonore dans leur vie quotidienne. Alors que ces performances pouvaient être rapidement évaluées chez les adultes à l'aide du questionnaire SSQ 15 (Speech, Spatial and Qualities of hearing, (Moulin et al., 2019), aucune version adaptée n'était disponible pour les enfants.

Par conséquent, l'objectif principal de cette étude était de développer une version abrégée en français du questionnaire SSQ en se basant sur des questionnaires existants (Galvin & Noble, 2013; Moulin et al., 2019) et de valider cette nouvelle version auprès d'une population d'enfants porteurs d'un ou deux implants cochléaires.

Nous avons sélectionné 12 items provenant de trois sous-échelles (perception de la parole, localisation spatiale et qualités de l'audition) et les avons soumis à des passations auprès de 154 enfants porteurs d'implants cochléaires (100 bIC et 54 uIC). La validité de construction a été évaluée en examinant les relations entre les scores obtenus à ce questionnaire (nommé Kid-SSQ) et des indicateurs cliniques objectifs (par exemple, l'âge lors du test, la durée d'expérience en audition binaurale, etc.).

Résultats. Notre questionnaire sous forme courte pouvait être complété en moins de 10 minutes par les enfants, et présentait des corrélations significatives entre les scores obtenus et des mesures cliniques (par exemple, une corrélation positive a été observée entre l'expérience binaurale et le score de la sous-échelle spatiale). En conséquence, notre questionnaire Kid-SSQ se révèle être un outil solide et cliniquement pertinent pour l'auto-évaluation des difficultés de localisation sonore chez les enfants.

CONCLUSION du CHAPITRE 2

L'ensemble des études exposées dans ce chapitre met en avant plusieurs conclusions significatives : (1) il est possible d'objectiver les déficits de localisation auditive 3D de patients (adultes et enfants), (2) l'amélioration des ambiguïtés auditives de spatialisation par l'exploration active de l'espace acoustique est un avantage certain pour les patients possédant des prothèses auditives, (3) cette stratégie exploratoire est cruciale dans la mise en place de rééducation des déficits de spatialisation, (4) une rééducation de la spatialisation auditive a des effets bénéfiques sur d'autres fonctions cognitives, et (5) il est possible de réaliser une évaluation rapide des déficits spatialisation auditive en routine clinique.

Ainsi, forte de mes expertises développées au cours des 8 dernières années, je souhaite maintenant poursuivre mes activités de recherche propres autour de la spatialisation auditive avec comme objectif principal d'identifier les déficits et les ressources-clefs des patients quand ils localisent des sons dans l'espace, et leur proposer des pistes de rééducation adaptées. Pour y parvenir, je vais articuler ma démarche autour de plusieurs axes :

* Identifier les stratégies d'appareillage prothétique les moins délétères pour localiser les sons chez les patients qui portent des prothèses auditives implantables (implants cochléaires). Dans le chapitre suivant, ce projet sera présenté au regard des stratégies actuellement proposés aux patients implantés cochléaires.

* Identifier objectivement des déficits de localisation auditive auprès de patients reportant des plaintes subjectives (personnes ayant subi un traumatisme auditif). Dans le chapitre suivant, je présenterai comment je vais réaliser ce projet et montrer que cette évaluation permettra à court terme de proposer une prise en soin adaptée à ces patients. A long terme, il sera possible d'envisager de proposer un protocole de rééducation adapté à ces patients en se basant sur leurs déficits et ressources.

* Identifier objectivement des déficits de localisation auditive auprès de patients présentant des altérations cognitives (schizophrènes avec un premier épisode psychotique). Dans le chapitre suivant, je montrerai que ce projet apportera des informations précieuses sur les altérations sensorielles et cognitives associées à cette maladie. A long terme, on pourra utiliser ces nouveaux éléments de connaissances pour proposer un protocole de rééducation adapté (qui restera à définir).

* Proposer aux patients les plus vulnérables (les enfants) un programme de rééducation « écologique » qui correspond à leurs attentes et leurs capacités. Ce programme fera partie d'un projet présenté dans le chapitre suivant et portant sur le développement d'une solution de rééducation ludique.

3. Perspectives de recherche

Mes études précédentes m'ont permis de tisser un réseau scientifique qui nourrit à la fois dans mes activités de recherche et d'enseignement, me permettant ainsi de nourrir plusieurs projets de recherche (soit en tant que collaboratrice, soit en tant qu'investigatrice principale). Dans ce chapitre, j'ai choisi de ne présenter que les projets qui me semblent les plus aboutis dans leur réflexion. Ces projets sont centrés sur la thématique de la spatialisation auditive et reposent sur l'approche SPHERE en tant qu'outil d'évaluation et de rééducation à la localisation sonore. Certains projets (MicLo, Kid Train) prolongent des collaborations existantes afin de répondre à des questions encore en suspens concernant l'impact des implants cochléaires sur la localisation sonore, ainsi que la plasticité développementale des capacités de spatialisation auditive dans la population pédiatrique. D'autres projets (SPHEAR, Lociz) résultent de nouvelles collaborations émergentes, renforçant ma place d'experte scientifique dans le domaine de l'audition spatiale, tout en m'incitant à acquérir de nouvelles connaissances dans d'autres champs d'expertise. Quelque soit le projet, sa réussite va dépendre étroitement de la qualité des collaborations avec d'autres chercheurs, des ingénieurs et des cliniciens, ainsi que par l'encadrement d'étudiants en Master et en thèse.

1. Impact du positionnement des microphones des implants cochléaires dans les performances de SA. Projet MicLo.

De nombreuses études ont montré que les appareils auditifs dégradaient les différents indices acoustiques indispensables à une bonne localisation sonore (Carette et al., 2014; Denk et al., 2019; Nava et al., 2009). Cette perte de résolution du signal auditif n'est pas fondamentalement délétère pour l'accès au langage oral, mais elle s'avère largement insuffisante pour la bonne perception de la structure tridimensionnelle des sons donc la spatialisation auditive (Keidser et al., 2009). Le traitement du signal fait par les prothèses n'est pas le seul facteur à impacter les performances de localisation. En effet, l'étude de Denk (2019) a mis en évidence que la position du microphone des prothèses auditives (c'est-à-dire l'endroit où se fait la prise de son) était un facteur déterminant dans les performances de localisation spatiale, et affectait les indices spatiaux (Mayo & Goupell, 2020). Dans cette optique, plusieurs stratégies de placement ont été évaluées dans les prothèses auditives, mais principalement au service d'une bonne compréhension de la parole dans le bruit. Or, ces choix stratégiques sont souvent proposés au détriment de la localisation (Kolberg et al., 2015), relayée au second plan par manque de sensibilisation des professionnels de santé aux déficits de localisation auditive. Ces constats amènent à se questionner sur l'impact du positionnement du microphone chez les patients porteurs de deux IC (patients implantés bIC) pour localiser des sons. Cette question prend tout son sens lorsque l'on sait que la position du microphone peut varier selon 3

modalités sur le système d'implant cochléaire (voir figure 8 ci-dessous) : (1) sur le processeur de son, le processeur étant porté en contour d'oreille, (2) sur l'antenne qui est positionné sur le crâne au regard du récepteur sous-cutané, (3) devant le conduit auditif externe, le un microphone est alors déporté grâce à un câble fin.

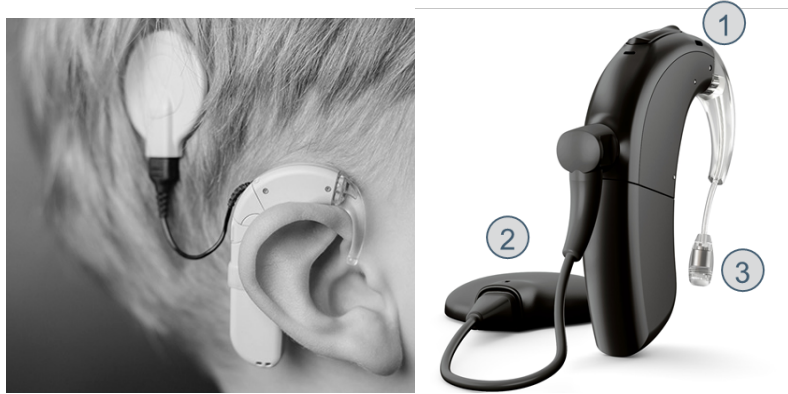


Figure 8: Partie externe d'un implant cochléaire (figure de gauche), positions potentielles du microphone (1 : sur le processeur ; 2 : sur l'antenne ; 3 : devant le conduit auditif)

La plupart des microphones d'implants cochléaires sont portés par la prothèse positionnée en contour d'oreille. Ce positionnement à l'extérieur de l'oreille supprime l'effet pavillonnaire qui, en conséquence, dégrade les indices monauraux. Pour pallier à cet effet, plusieurs solutions technologiques ont été développées par les industriels telles que le positionnement des microphones à l'entrée du conduit auditif externe et l'utilisation d'algorithmes de traitement de signal dont le but est de reproduire les caractéristiques du pavillon, sans accessoire additionnel (ex. Real Ear Sound ou RES).

Suite à cette stratégie de positionnement du microphone et d'utilisation d'algorithmes, certaines observations ont été faites : 1) En ce qui concerne la localisation spatiale, le microphone placé devant le conduit auditif externe permet une meilleure localisation en azimut et en élévation (Denk et al., 2019; H. G. Jones et al., 2016; Van den Bogaert et al., 2011), ainsi qu'une meilleure discrimination spatiale des sons dans le bruit (H. G. Jones et al., 2016; Kolberg et al., 2015; Mantokoudis et al., 2011). Cependant ces résultats ne sont pas en accord avec ceux d'autres études montrant qu'aucune position de microphone ne semble favoriser une tâche auditive par rapport à une autre (e.g. localisation spatiale par rapport à la compréhension de la parole ; (Bayri & Çiprut, 2020; Davis & Gifford, 2018; Dwyer et al., 2020). 2) En termes de compréhension de la parole, aucune différence significative n'a été observée entre le microphone placé devant le conduit auditif externe et celui positionné sur la prothèse (Bayri & Çiprut, 2020; Davis & Gifford, 2018; Dwyer et al., 2020). 3) L'algorithme RES simulant les caractéristiques pavillon permet d'obtenir des performances similaires à celles obtenues lorsque le microphone est placé devant le conduit auditif externe, en termes de localisation spatiale et de compréhension de la parole dans le bruit (source industrielle de microSavia Digital Bionics).

Ces études ne semblent pas faire consensus sur le bénéfice d'une position de microphone par rapport aux autres. En fait, les méthodologies employées dans ces études pour estimer les bénéfices de positionnement sur la SA des patients souffrent de biais méthodologiques qui ne permettent pas d'étudier la SA en conditions 'écologique' d'écoute : ainsi des indices visuels sont fournis (e.g. figure 9, Denk et al., 2019), les participants ne bénéficient pas d'écoute active (e.g. Davis & Gifford, 2018), ou réalisent la tâche dans des ambiances sonores appauvrie (e.g. anéchoïque). On peut donc émettre l'hypothèse que les données actuelles sur le bénéfice d'une position de microphone manquent encore des données objectives quant aux capacités réelles des patients à exploiter les signaux acoustiques des prothèses.

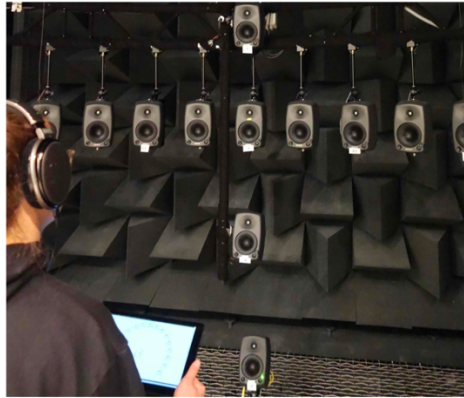


Figure 9: disposition des sources sonores (extrait de Denk et al. 2019) : seules les haut-parleurs possédant un indice blanc sont des sources sonores potentiels, donnant un indice visuel au patient pour identifier la source sonore.

Évaluer les capacités de localisation en tenant comptes de prérequis méthodologiques (i.e. mesures dans les trois dimensions de l'espace, en contrôlant les indices visuels et en favorisant l'écoute active) permettra de déterminer si ces stratégies technologiques améliorent effectivement la perception de la localisation sonore chez les patients équipés d'implants cochléaires.

Dans ce projet MicLo, nous faisons l'hypothèse que, en faisant varier la position des microphones des systèmes d'implant cochléaire, les performances de localisation seront optimisées dans la condition « microphone devant le conduit auditif externe » ou « avec l'algorithme RES », sans altération des capacités de compréhension de la parole.

Objectif. Notre étude vise à évaluer comment le positionnement des microphones des implants cochléaires influence les performances auditives (localisation spatiale, compréhension de la parole et qualité d'écoute spatiale) des patients porteurs de deux implants cochléaires (bIC). Cette étude examinera également si l'utilisation d'un algorithme comme le Real Ear Sound (RES, www.phonakpro.com/) peut améliorer la spatialisation des sons. Pour ce faire, nous effectuerons un changement symétrique de positions de microphones sur les prothèses, et à

chaque changement, nous réaliserons des tests pour évaluer l'impact sur les performances auditives subjectives et objectives des patients. Le dispositif utilisé dans cette étude sera basé sur l'approche SPHERE, impliquant des tâches de localisation spatiale en environnement tridimensionnel, contrôlé (sans indice visuel), nécessitant une écoute active, et réalisé avec des environnements sonores complexes (i.e bruyant) afin d'être au plus près des conditions d'écoute naturelle du quotidien des patients.

Protocole. Pour évaluer l'effet du positionnement des microphones, nous utiliserons le système d'implant cochléaire de la marque Advanced Bionics (AB) de type Marvel ®. En effet de ce dispositif offre trois options de position pour le microphone : position 1 sur le processeur de son (position standard), position 2 sur l'antenne (déporté du processeur et situé sur le crâne) qu'il est courant de voir afin d'alléger ou déporter le contour d'oreille, position 3 devant le conduit auditif externe. Cette position est la plus proche du fonctionnement physiologique et offre un confort d'écoute supérieur à la position 1 (Frohne-Büchner et al, 2004). Mais elle reste peu répandue car celle-ci est à l'heure actuelle considérée comme un accessoire à la charge des patients. Le positionnement du microphone à la position 1, contrairement aux deux autres positions, offre également la possibilité d'utiliser l'algorithme RES pour simuler les caractéristiques du pavillon. Cet algorithme RES est intéressant pour notre étude car il permet de simuler les caractéristiques pavillon (mimer une prise de son microphone en position 3 tout en conservant la position standard sur le processeur, et contrairement au microphone positionné devant le conduit auditif externe (position 3), il ne nécessite pas de matériel additionnel.

Donc, nous proposons de tester différentes conditions de position (position / position *algorithme) : condition A- microphone activé sur le processeur, condition B- microphone activé sur le processeur avec l'algorithme Real Ear Sound (RES) actif, condition C- microphone placé devant le conduit auditif externe, condition D- microphone activé sur l'antenne.

Les patients bIC testeront successivement les quatre conditions de réglage ci-dessus lors de visites espacées d'au moins 14 jours pour permettre une adaptation à la nouvelle position et à l'utilisation de l'algorithme du microphone dans leur vie quotidienne. A chaque visite, les changements de position seront faits en même temps et en symétrique pour chaque implant (par exemple, lors de la visite 2, nous activerons les microphones de l'implant gauche et de l'implant droit sur l'antenne).

Étant donné que le changement de position des microphones des implants cochléaires peut avoir un impact direct sur la localisation spatiale et un impact indirect sur la compréhension de la parole dans le bruit, les tests effectués dans ce protocole cibleront ces différentes fonctionnalités auditives : une évaluation objective de la localisation 3D en milieu bruyant (test SPHERE avec scène auditive complexe), un test objectif de compréhension de la parole (Framatrix, évaluant les capacités de compréhension de la parole dans le bruit, Jansen et al., 2012) et un test subjectif des capacités de localisation sonore (SSQ15 ou Speech, Spatial and Qualities of Hearing Scale short-

form (Moulin et al., 2019). Pour le test objectif SPHERE, il sera demandé aux participants d'évaluer subjectivement leur difficulté et leur confiance ressenties lors de ce test en utilisant une échelle de Likert ; cette échelle nous permettra d'apprécier l'effort cognitif déployé par le patient pour réaliser ce test de localisation. Enfin, un questionnaire de satisfaction quant à ces différentes positions sera réalisé en fin de protocole afin d'obtenir l'avis des patients sur la configuration qu'ils estiment la plus favorable au quotidien.

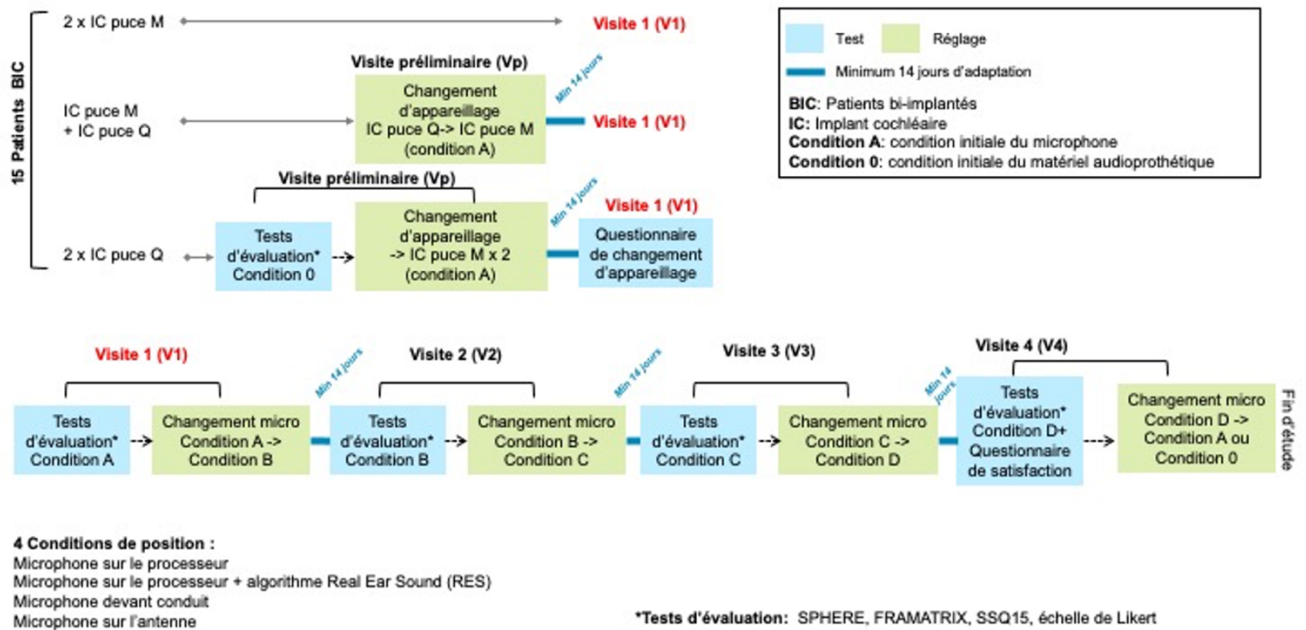


Figure 10 : Flowchart de l'étude MicLo. 15 patients bi-implantés cochléaires seront recrutés pour ce projet visant à étudier l'impact de la condition de positions de microphones sur l'écoute spatiale. Pour bénéficier de ces changements de conditions (A à D), les patients doivent être équipés de deux prothèses AB-Marvel (2xIC puce M) qui est défini par une condition de position initiale (condition A). Les patients qui désirent participer à l'étude et qui disposent d'une prothèse AB-Marvel et d'une prothèses ancienne génération (IC puce M+ IC puce Q) ou deux prothèses ancienne génération (2xIC puce Q) avec une condition de position initiale (condition 0), se verront proposer un changement d'appareillage pour deux prothèses à puce M lors d'une visite préliminaire (Vp). Puis tous les patients réaliseront des tests d'évaluation aux visites V1 à V4 pour les 4 conditions de positions de microphones (condition A ou initiale à condition D). Après les tests d'évaluation, la condition de position testée lors de la visite sera changée. Lors de la dernière visite (V4), les patients donneront leur avis sur les différentes stratégies de positionnement qu'ils auront pu expérimenter dans leur quotidien. A la suite de cette visite, les patients reviendront à leur condition de position initiale (condition A) ou leur appareillage initiale (condition 0)

Attendus et perspectives. Cette étude permettra de quantifier les déficits de perception 3D chez les adultes porteurs d'implants cochléaires en fonction du positionnement des microphones et de l'utilisation de l'algorithme RES. Les changements de position des microphones et l'utilisation de RES nous permettront d'évaluer l'impact de ces choix à la fois sur la spatialisation auditive et les performances langagières de ces patients. Actuellement, les microphones sont automatiquement activés au niveau du processeur pour une meilleure compréhension de la parole, sans considérer l'optimisation des capacités de localisation spatiale des patients et notamment en élévation

où les indices monauraux d'origine pavillonnaire sont cruciaux. De plus, ce projet permettra de mieux saisir l'impact de l'exploration céphalique dans l'acquisition des signaux auditifs dynamiques essentiels à la SA en milieu bruyant, et ainsi identifier si cette stratégie d'écoute active est favorisée par une ou des conditions de position de microphone. Enfin, ce projet permettra aux audioprothésistes (diplômés d'état et étudiants) de collaborer activement à un projet de recherche mêlant des expertises d'audioprothèse et de recherche.

2. Re-spatialiser les sons chez les enfants implantés cochléaires avec une approche ludique. Projet Kid Train.

Après une implantation cochléaire, la rééducation orthophonique joue un rôle essentiel dans le développement auditif et langagier des enfants, en particulier dans des environnements auditifs variés. Cependant, il manque actuellement un référentiel national pour guider la rééducation de la spatialisation auditive chez ces enfants. En l'absence d'outils appropriés, la localisation spatiale n'est que rarement abordée pendant les séances de rééducation. Or, les enfants implantés cochléaires présentent des déficits en spatialisation auditive, et ces difficultés peuvent être atténuées par une écoute active (comme évoqué dans la chapitre précédent). De plus, des questions persistent concernant les capacités de localisation spatiale des enfants en milieu bruyant, que ces derniers présentent ou non des déficits auditifs. Mes études précédentes ont également souligné l'efficacité croissante de la réalité virtuelle (RV) pour évaluer et rééduquer la localisation spatiale. Ainsi par un protocole intensif de rééducation de la localisation spatiale en réalité virtuelle avec des indices spatiaux (retours sensoriels de performance), des adultes implantés cochléaires ont pu améliorer significativement leurs compétences de localisation et leur compréhension de la parole dans le bruit en seulement un mois. De plus, ces améliorations ont persisté un mois après la fin de la rééducation (Coudert et al., 2022).

En prenant en compte ces constatations qui mettent en évidence le manque de prise en charge rééducative de la localisation spatiale après une implantation cochléaire, ainsi que l'importance de mettre en place cette rééducation dès que possible chez les enfants bi-implantés cochléaires, et en considérant l'efficacité démontrée chez les adultes implantés cochléaires, nous souhaitons élaborer un nouveau protocole d'entraînement de la localisation spatiale en réalité virtuelle pour les enfants bi-implantés cochléaires âgés de 7 à 17 ans. Ce protocole sera basé sur les résultats de mes études précédentes : il s'appuiera sur une approche immersive encore plus écologique en immergeant les participants dans des environnements reflétant des situations de la vie quotidienne des enfants, environnement visuel supportant la localisation auditive (Valzolgher et al., 2020, 2023a et 2023b ; Coudert et al., 2022) et surtout dans des environnements sonores complexes (milieux bruyant). Ainsi, en combinant des sons et ambiances sonores variés avec des scènes visuelles en RV 'écologiques' (e.g. salle de classe), on sera à même d'apporter un regard plus intégré à la problématique de la re-spatialisation auditive. Simultanément, un groupe témoin d'enfants normo-entendants sera recruté pour évaluer les capacités de

localisation spatiale en milieu bruyant en fonction de l'âge. Cela nous permettra d'obtenir de nouvelles perspectives sur la maturation auditive de la population pédiatrique.

Objectif. L'objectif de cette étude est d'évaluer les effets d'une rééducation spatiale sur les compétences auditives (localisation spatiale, compréhension de la parole dans le bruit et qualité d'écoute) des enfants bi-implantés cochléaires. Cet objectif sera complété par l'évaluation de la maturation des ces compétences auditives chez les enfants normo-entendants. Ainsi, chez l'enfants bi-implantés cochléaires, nous faisons l'hypothèse que le protocole de rééducation à l'aide d'un indice spatial permettra de réduire significativement les déficits de spatialisation auditive du fait d'une plasticité cérébrale importante à cet âge. Concernant l'enfant normo-entendant, nous faisons l'hypothèse qu'en milieu bruyant, les capacités de localisation spatiale s'affineront avec l'âge du fait d'une plasticité cérébrale liée à la maturation en âge et seront fonction de l'espace auditif exploré (par exemple, avant ou arrière).

Protocole.

Pour étayer notre hypothèse sur les effets d'un protocole d'entraînement, nous faisons le choix de cibler des enfants BIC âgés de 7 à 17 ans. En effet, les patients BIC, contrairement aux patients porteurs d'un seul implant cochléaire, ont accès à la binauralité grâce à leurs deux implants cochléaires, et sont capables d'utiliser les indices binauraux pour améliorer leurs performances de localisation spatiale lors de notre protocole de rééducation. Ces patients devront être porteurs de leur deuxième implant cochléaire depuis plus d'un an car il a été montré que la première année suivant l'implantation cochléaire, il existe un pic d'apprentissage spontané important avec une restructuration de la localisation des sons (Chang 2010). Le groupe contrôle d'enfants NE permettra d'évaluer les effets de la plasticité cérébrale liée à l'accroissement en âge sur les performances de localisation spatiale. Deux groupes d'enfants (7-12 ans et 13-17 ans) seront analysés en sous-groupes, Coudert et al. (2022) ayant montré que les performances de localisation auditive se stabilisent vers 13 ans.

Ainsi, deux groupes d'enfants seront constitués : un groupe d'enfants *normo-entendants* (NE, n=40) appariés en termes d'âge avec un groupe d'enfants *bi-implantés cochléaires* (bIC, n=16). En début de protocole, chaque groupe réalisera des tests d'évaluation : évaluation objective de localisation spatiale (protocole d'évaluation de la spatialisation auditive - SPHERE), évaluation objective d'écoute dans le bruit (test audiométrique standardisé - Frasimat) et répondra à un questionnaire de qualité d'écoute spatiale (évaluation subjective - KID-SSQ). L'inclusion d'enfants normo-entendant permettant ainsi d'obtenir des mesures de contrôle de la spatialisation auditive et d'évaluer l'impact du développement sur la maturation sur ces compétences. Ces tests seront répétés

lors de deux visites successives (V1 et V2) afin de s'assurer la stabilité des mesures (voir flowchart figure 11). Les enfants BIC participeront à d'autres visites d'évaluation tout au long du protocole : après 4 séances d'entraînement (V7), après 8 séances d'entraînement (V12) et 1 mois après la fin des entraînements (V13).

Lors des séances de rééducation, les enfants BIC seront immergés dans un environnement virtuel et une ambiance sonore bruyante (les sources sonores constitutifs des bruits ne seront pas visibles). Dans cet environnement audiovisuel, les enfants devront prêter attention à une voix qui dictera le nom d'un objet (e.g. « renard »). Pour rendre l'environnement aussi réaliste et ludique que possible, nous avons choisi une salle de classe virtuelle simulée en réalité virtuelle plongée dans des bruits de conversation provenant de l'extérieur (couloir et cour, provenant de hauts parleurs). Dans la salle, des éléments visuels imaginaires seront présentés (de style Pokémon) dont les positions peuvent varier dans toutes les dimensions de l'espace ; un seul élément imaginaire sera positionné au même endroit que le haut-parleur qui émet le son de la voix, les autres éléments étant des distracteurs spatiaux. De plus, les enfants pourront voir un tableau situé en face d'eux sur lequel seront présentées des images, une seule image représentera l'objet énoncé par la voix, les autres images étant des distracteurs sémantiques (voir figure 11).



Figure 11 : Projet en RV de l'environnement immersif de rééducation de KidTrain. La salle de classe comportera des éléments visuels structurant (murs, plafond, tables, etc.) et des sources sonores potentielles (symbolisées par des dragons) dont une seule sera à l'origine de la voix à localiser (image de gauche). Cette source sonore réelle sera positionnée manuellement par l'expérimentateur en suivant la procédure SPHERE, d'autres haut-parleurs assureront de délivrer un bruit de fond (photo de droite) autour de l'enfant portant un casque de réalité virtuelle.

Afin d'éliminer des facteurs confondants du protocole de rééducation tels que ceux liés à un entraînement à l'écoute en milieu bruyant (qui pourraient impacter les capacités d'écoute dans le bruit), deux types d'entraînement rééducatifs seront testés. Ainsi, tout en écoutant la voix, ils devront réaliser soit :

(1) un entraînement qualifié de *spatial* où les enfants devront identifier dans l'espace la position de la voix sans qu'il soit besoin de comprendre le mot. Pour ceci ils pointeront leur tête en direction de la source sonore. Ils recevront alors un retour visuel de leur performance : seul élément imaginaire correspondant à la position exacte du son s'anamera et ils devront alors corriger leur orientation ;

(2) un entraînement qualifié de *non spatialisé* (ou sémantique) où les enfants devront comprendre le mot dicté par la voix sans qu'il soit besoin d'identifier la position de la voix dans l'espace. Pour ceci, ils pointeront leur tête en direction de l'image disposée sur un tableau virtuel situé en face d'eux. Ils recevront alors un retour visuel de leur performance : une seule image correspondant à la nature du mot dictée s'anamera et ils devront alors corriger leur orientation.

Les enfants BIC seront randomisés entre ces deux groupes : la moitié des enfants BIC (groupe BIC^A) effectuera 8 séances d'entraînement spatialisé, tandis que l'autre moitié (groupe BIC^B) débutera avec 4 séances d'entraînement non spatial suivi de 4 séances d'entraînement spatialisé (voir figure 12).

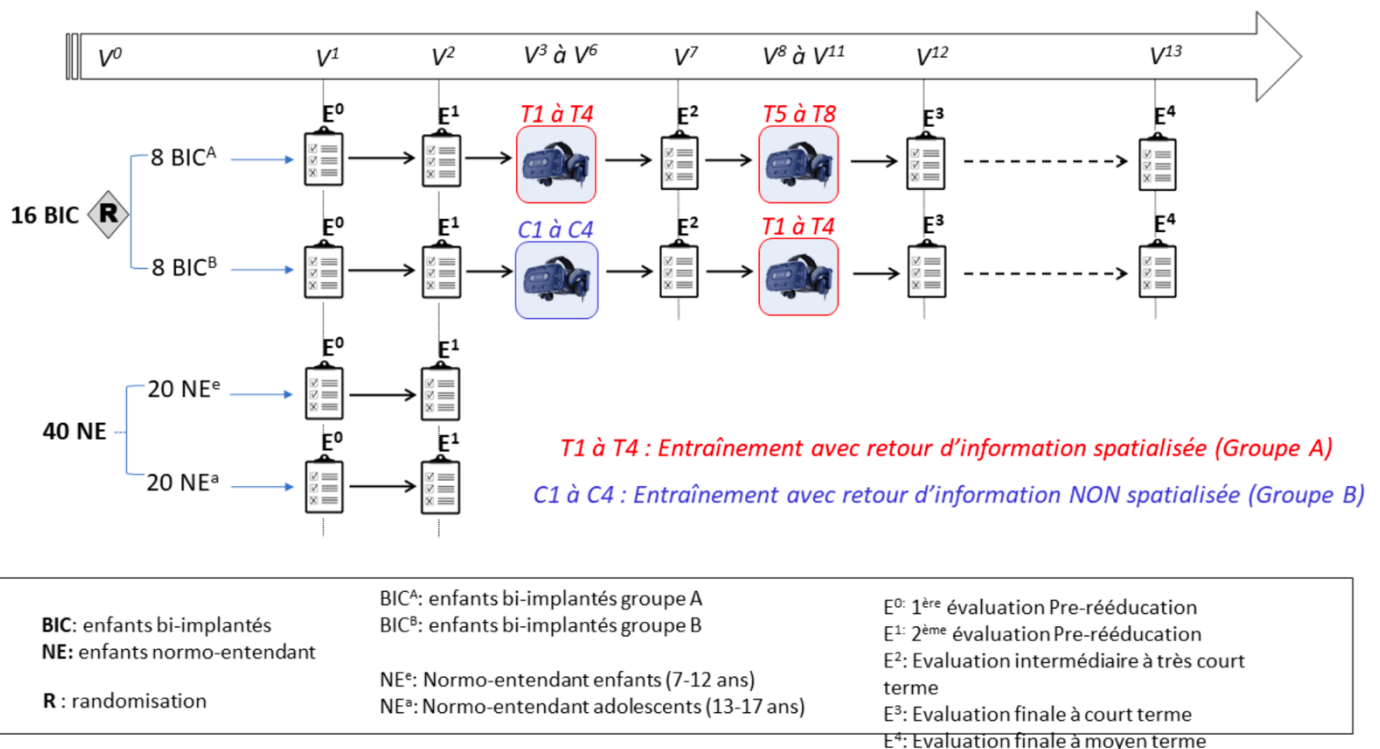


Figure 12 : Flowchart de l'étude Kid-Train. Tous les enfants réaliseront deux séances d'évaluation initiale (E0 et E1) permettant d'identifier leurs capacités auditives. Les enfants BIC^A réaliseront 4 séances d'entraînement avec retour d'information spatialisée (T1 à T4), les enfants BIC^B auront 4 séances d'entraînement sans retour d'information spatialisée (C1 à C4). Une séance d'évaluation (E2) permettra de tester leurs capacités auditives après 4 séances de rééducation spatialisée versus 4 non spatialisées. Ensuite, tous les enfants BIC (A et B) réaliseront les 4 séances d'entraînement suivantes avec retour d'information spatialisée (BIC^A: T5 à T8 et BIC^B: T1 à T4). Une séance d'évaluation à court terme sera réalisée dès la fin de l'entraînement (E3), et une autre séance d'évaluation à moyen terme sera réalisée 1 mois après l'entraînement (E4).

Attendus et perspective. Nous nous attendons à ce que les enfants bIC tirent parti d'un entraînement spatialisé après 4 séances d'entraînement (e.g. réduction des confusions de localisation sonore entre l'espace avant et arrière ; augmentation de leurs capacités d'écoute dans le bruit). De plus, nous voulons démontrer qu'après 8 séances de ce type d'entraînement, les enfants reportent des bénéfices sur leur qualité d'écoute au quotidien et que les bénéfices de l'entraînement spatial se maintient dans le temps. Enfin, en incluant des enfants normo-entendants dans les tests d'évaluation uniquement, nous allons montrer que les capacités d'audition spatiale mûrissent tout au long du développement. Ainsi, l'ensemble de ces résultats permettra de confirmer la faisabilité d'une telle rééducation en milieu pédiatrique, d'en évaluer les bénéfices sur les performances auditives avant d'envisager la mise en place d'une étude de plus grande ampleur. Cette rééducation spatiale pourrait être par la suite proposée : 1) à toute personne implantée bimodale (un implant cochléaire et une aide auditive) ou bilatérale (deux implants cochléaires), manifestant une gêne quotidienne de localisation du son dans l'espace, et 2) de façon systématique en post-opératoire, le plus précocement possible, une fois les réglages audioprothétiques stabilisés. Ce protocole d'entraînement permettra également de proposer de nouveaux axes de rééducation orthophonique en libéral mais aussi sur la plateforme Orthophonie & Surdit , en combinant la localisation spatiale des sons   la compr hension de la parole dans le bruit.

Enfin, ce projet permettra aux orthophonistes (dipl m s d' tat et  tudiants) de collaborer activement   un projet de recherche.

Financement. Hospices Civiles de Lyon « Kid-Train CT69HCL 22_0599 »

3. Impact du traumatisme sonore sur les capacit s de SA chez les militaires traumatis s sonores. Projet SPHEAR.

Les pertes auditives li es au bruit constituent la deuxi me cause la plus courante de d ficiance auditive apr s les alt rations li es   l' ge. Elles touchent environ 5% de la population mondiale (Natarajan et al., 2023). Le d pistage et le diagnostic de ces pertes auditives r sultent de l'examen des ant c dents d'exposition au bruit du patient, ainsi que d' valuations cliniques telles que les audiogrammes et les tests de compr hension de la parole dans le bruit. Bien que certaines professions soient plus susceptibles de causer des pertes auditives, tels que pour les musiciens (Pouryaghoub et al., 2017) ou les pompiers (Millet et al., 2023; Snapp et al., 2022), l'environnement des militaires est particuli rement propice   la survenue de d ficits auditifs (Medina-Garin et al., 2016). En effet, ces professionnels sont expos s   des niveaux sonores tr s  lev s lors d'exercices d'entra nement et d'op rations : les bruits impulsifs des armes   feu atteignent des intensit s sonores allant de 150   165 dB, tandis que les bruits continus des  quipements militaires varient entre 100 et 120 dBA. Ces niveaux sonores, bien sup rieurs aux recommandations d'exposition s re (le seuil de danger est   80 dB et celui de douleur   130 dB), peuvent entra ner des pertes auditives consid rables (Yehudai et al., 2017), d'autant plus si les oreilles ne sont pas correctement

isolées des sons traumatiques (Kwak & Han, 2021) . Il est également important de noter que les effets du bruit peuvent être cumulatifs, ce qui signifie que des expositions répétées au bruit, comme celles liées aux tirs d'armes à feu, peuvent causer des dommages irréversibles (Natarajan et al., 2023).

En cas de traumatisme auditif suite à une exposition sonore excessive, des problèmes de surdité surviennent. Lors d'un incident ou d'un traumatisme sonore aigu, les militaires sont immédiatement pris en charge pour une évaluation ORL, un diagnostic et des soins. Si les tests révèlent des seuils auditifs conformes aux critères d'aptitude militaire, les militaires sont maintenus dans leurs fonctions et conservent leur aptitude au service, y compris aux opérations extérieures et au tir. Cependant, malgré leur capacité à remplir leurs fonctions professionnelles, de nombreux patients signalent des plaintes récurrentes de localisation des sons, en particulier lorsque leur environnement de travail est bruyant (comme dans le cas des chars). Actuellement, il n'existe aucun test objectif permettant d'identifier et de quantifier ces problèmes de spatialisation auditive.

En raison de mon expertise dans le domaine de la spatialisation auditive et de l'opportunité de travailler avec le Centre de Santé des Armées de Lyon (qui prend en soin les militaires traumatisés sonores - MTS), nous avons projetés à mettre en œuvre un projet de recherche visant à explorer et à mesurer la spatialisation auditive chez ces militaires.

Objectif. Notre objectif principal est de mettre en évidence et de quantifier les déficits de spatialisation auditive chez les militaires ayant subi un traumatisme sonore (MTS), en comparaison avec des adultes normo-entendants (NE), lors d'une tâche de localisation spatiale. De plus, nous évaluerons leur qualité auditive spatiale au moyen d'un questionnaire (SSQ 15, (Moulin et al., 2019), leur capacité de compréhension de la parole en situation de bruit (Framatrix, Jansen et al., 2012), ainsi que leurs stratégies d'écoute active (implication des mouvements exploratoires de la tête) lors de la tâche de spatialisation auditive.

Protocole. La spatialisation auditive sera réalisée à l'aide du protocole SPHERE et permettra d'objectiver leurs capacités de SA. D'autres tests seront réalisés afin d'identifier leur compréhension de la parole dans le bruit (Framatrix) ainsi que leur qualité d'écoute auditive (SSQ15).

Attendus et perspective. Nous émettons l'hypothèse que les adultes militaires ayant subi un traumatisme sonore présentent des déficits en spatialisation auditive, et que ces déficits pourraient être corrélés avec leurs plaintes quotidiennes ainsi qu'avec leurs difficultés d'écoute en milieu bruyant. De plus, nous nous attendons à ce que leurs difficultés à localiser les sources sonores soient atténuées s'ils mettent en place des stratégies d'écoute active (Valzolgher et al., 2023).

Cette étude permettra de proposer pour la première fois un outil spécifiquement dédié au diagnostic des déficits de spatialisation auditive chez les militaires ayant subi un traumatisme sonore. Cette évaluation objective des déficits de spatialisation auditive nous permettra de comprendre les difficultés ressenties par les patients et d'envisager la mise en place de rééducations ciblées dont nous avons fait notre expertise (e.g. Coudert et al., 2022 ; Valzolgher et al., 2023).

Financement. Si le projet a reçu un avis favorable en Juillet 2021 pour être soutenu financièrement par leur Service des Santé des Armées de Lyon, ce financement n'a pas été déployé (l'hôpital d'Instructions de Santé des Armées ayant été fermé précipitamment). Cependant, sa promotion a été reprise par les Hospices Civils de Lyon en 2023.

4. Spatialisation auditive en réalité virtuelles chez des patients avec un premier épisode psychotique. Projet LocIZ.

La SA est cruciale au quotidien puisqu'il participe, entre autres, à la structuration de notre environnement perceptif, à nous orienter spatialement, notamment vis-à-vis d'autrui, et contribue ainsi à définir les bases de l'interaction sociale. Dans certaines pathologies psychiatriques telles que la schizophrénie, des perturbations de perceptions auditives sont reportées, dont des difficultés à analyser des scènes auditives, à localiser et à discriminer des sons dans l'espace et notamment quand les tâches impliquent une discrimination temporelle des informations sensorielles (pour revue (Gröhn et al., 2022)). Si les capacités de spatialisation auditive restent encore peu étudiées pour ces patients, des déficits de spatialisation pourraient sous-tendre les difficultés sociales et cognitives retrouvées dans la schizophrénie, comme (1) la mémoire de source audito-verbale (*source monitoring*) ou encore (2) les confusions dans la perception de sources externes avec des sources internes (*externalization*).

La mémoire de source audito-verbale, également connue sous le terme de source monitoring (Johnson et al., 1993), est un processus cognitif complexe qui concerne la capacité de se souvenir de l'origine d'une information entendue. Il s'agit de discerner la source d'une information auditive (ou verbale), en attribuant correctement cette information à sa provenance, qu'elle soit réelle ou imaginaire. Ce processus est essentiel pour déterminer si une information a été perçue directement, entendue d'une autre personne, ou générée par notre propre pensée. Les études ont montré que la mémoire de source audito-verbale peut être influencée par divers facteurs, et notamment l'attention portée à la source lors de l'encodage de l'information.

La capacité à percevoir l'origine externe des sons (externalisation, pour revue voire (Best et al., 2020) se base, entre autres sur la distinction entre les sons qui sont produits par nos propres actions ou nos pensées (sources internes) et celles provenant de sources externes (e.g. générées par autrui). Les fausses perceptions ou

hallucinations sont un symptôme de la schizophrénie, et affectent le plus souvent la perception auditive. Le patient attribue de manière erronée, des perceptions auditives internes à une source externe, et ce défaut d'intégration est un signe neurologique caractéristique fréquemment observé dès les premiers épisodes psychotiques (PEP) de la maladie (Williams et al., 2010). Différentes structures cérébrales sont impliquées dans ce processus d'externalisation, tel que le cortex auditif qui analyse les caractéristiques acoustiques des sons, et le cortex préfrontal qui est impliqué dans la prise de décision concernant l'origine des sources auditives. Ces deux structures sont elles-mêmes impliquées dans la SA des sons.

Par ailleurs, des études ont montré que la schizophrénie est associée à des altérations des capacités de spatialisation auditive telle que des déficits dans les tâches de localisation et de discrimination spatiale auditive par rapport aux individus sains (Perrin et al., 2010; Ramage et al., 2012). De plus, les patients schizophrènes présentent des altérations dans l'intégration temporelle binaurale (i.e ITD, (Matthews et al., 2013). Ces déficits se manifestent principalement dans la localisation horizontale, sans qu'on en connaisse les répercussions sur les autres dimensions spatiales comme l'élévation et la distance.

Les stratégies comportementales adoptées par les patients schizophrènes pour améliorer la localisation sonore sont peu explorées. Si l'information visuelle cohérente spatialement et temporellement facilite la localisation sonore chez ces patients, une incohérence entre les informations visuelles et auditives induit un biais de localisation vers l'information visuelle (De Gelder et al., 2003; Gröhn et al., 2022). Par ailleurs, le mouvement de la tête semble influencer leur capacité de localisation, bien que la perte d'utilisation du pavillon auditif puisse limiter cette compensation (Guterman & Klein, 1991).

Les patients schizophrènes rencontrent également des difficultés à analyser des scènes auditives complexes et à séparer différentes sources auditives, surtout en présence de bruit. Ces altérations pourraient contribuer aux difficultés sociales et cognitives associées à la schizophrénie, telles que la mémoire de source auditive et la capacité à localiser des sons. Ces déficits sont observables dès les premiers stades de la maladie (Gawęda et al., 2018; Lavallé et al., 2021).

Étant donné que les altérations cognitives et anatomiques sont présentes dès les premiers stades de la schizophrénie, étudier les capacités de spatialisation auditive dès le premier épisode psychotique pourrait aider à mieux comprendre les altérations spécifiques de la maladie et éviter les facteurs de confusion liés aux traitements et à l'expérience hospitalière (Iliadou et al., 2013; Kasai et al., 2003). Cette approche novatrice pourrait fournir des informations précieuses sur les altérations sensorielles et cognitives associées à la schizophrénie.

Objectif. L'objectif principal est donc d'étudier les capacités à localiser les sources auditives des patients avec un PEP, et ceci en comparaison à des sujets sains, et d'étudier l'impact de ces capacités sur leur capacité à

distinguer ce qui provient d'une source interne et ce qui provient d'une source externe (capacités de *source-monitoring*) ainsi que sur leur capacité à externaliser une source auditive externe par rapport à une source interne (capacités d'externalisation auditive)

Protocole. L'étude portera sur deux populations : patients avec un PEP et des volontaires sains. Chaque population réalisera un audiogramme (pour s'assurer de l'absence de troubles auditifs), et une batterie de tests psychoacoustiques mesurant les capacités de discrimination des caractéristiques auditives de bases. Puis tous sujets réaliseront une tâche de spatialisation auditives (Gaveau et al., 2022) en localisant une voix dans un environnement bruité ; ce test sera réalisé en condition d'écoute active et en condition d'écoute statique pour étudier dans quelle mesure les patients peuvent utiliser ce comportement intentionnel pour échantillonner leur espace acoustique. Enfin seuls les patients réaliseront une tâche mesurant les capacités de *source-monitoring* et une tâche mesurant les capacités d'externalisation auditive.

Attendus et perspectives. Nous nous attendons à ce que les patients avec un PEP présentent des déficits lors de la localisation de voix en milieu bruité (et notamment en condition d'écoute statique), et une corrélation positive entre les performances de spatialisation auditive et les performances de *source-monitoring* et d'externalisation auditive. De plus, on s'attend à une corrélation négative entre les performances de spatialisation auditive et la sévérité des symptômes positifs telles que les hallucinations auditives.

Ces travaux sont une première étape pour envisager des travaux de plus grandes envergures en faisant varier d'autres facteurs contextuels pour se rapprocher au mieux d'un environnement écologique des patients (e.g., environnement sensoriel à structuration simple ou ambiguë), ceci justifiant l'utilisation de la méthodologie immersive de réalité virtuelle. Ils permettront également d'investiguer par la suite les corrélats anatomiques par méthode d'imagerie ou stimulation non invasive, et d'explorer les bénéfices de traitements précoces et ciblés des dysfonctionnements cérébraux qui auront été identifiés.

Financement. Projet financé et promu par le Centre Hospitalier Le Vinatier. 2023.

CONCLUSION & réflexions personnelles

L'élaboration de ce manuscrit suit une progression chronologique de mes activités de recherche depuis mon entrée dans le corps des Maîtres de Conférences Universitaires (MCU) en 2015: depuis ma collaboration avec les ingénieurs de la plateforme Neuroimmersion pour le développement et la validation de la solution SPHERE, en passant par l'évaluation subjective (KidSSQ) et objective des performances de la spatialisation auditive, pour enfin étudier les facteurs de l'adaptation de la spatialisation auditive.

Lors de l'écriture de ce manuscrit, je me suis rendu compte que les éléments de recherche exposés offrent de nouvelles perspectives :

- ☞ Des perspectives de prise en soin. Ainsi on peut envisager à court terme que l'utilisation de questionnaires d'écoute administrés en clinique, comme le KidSSQ chez les enfants, permettra d'identifier rapidement les patients reportant des déficits de localisation sonore au quotidien. Cette évaluation pourrait déboucher sur une proposition de rééducation "écologique" basée sur l'exploration active de leur environnement avec des retours d'indices spatiaux audio-visuels (comme le KidTrain), pour in fine, leur permettre de recouvrir des fonctions cognitives qui étaient altérées par leur déficit auditif (e.g. SPHERE-éval).
- ☞ Des perspectives de réglage personnalisé des prothèses auditives. Ainsi, on peut envisager que l'évaluation des capacités de spatialisation auditive puisse permettre d'affiner la qualité des réglages audio-prothétiques des aides auditives en fonction des besoins du patients (e.g. évaluation des réglages des aides pour répondre à un besoin d'écoute en milieu bruyé). L'utilisation RV est en train de se démocratiser dans la sphère de la recherche en audiologie pour tester les patients et les entraîner (Serafin et al., 2023), cependant comme le pointent ces auteurs, ces tests manquent de validité écologique. Ce travail de recherche nécessite une étroite collaboration avec des partenaires industriels, et sera réalisé dans le cadre du projet HORIZON-MSCA-2022-DN-01-01 « Cochlear implants and spatial hearing: Enabling access to the next dimension of hearing- CherISH », projet d'envergure européenne qui implique des acteurs du monde académique et industriel.
- ☞ Des perspectives d'amélioration des technologies des prothèses auditives. Ainsi on peut imaginer que l'évaluation des capacités de spatialisation auditive puisse impacter l'orientation vers le développement de nouvelles technologies. C'est dans ce domaine de R&D que l'utilisation de l'Intelligence Artificielle (IA) est envisagée en support aux traitements acoustiques. Ainsi, l'IA est actuellement sollicitée comme un facteur d'apprentissage des réseaux neuronaux profonds (Francl & McDermott, 2022), et comme le mentionnent ces auteurs : « One natural extension of our model would be to incorporate moving sound

sources and head movements. We modelled sound localization in static conditions because most experimental data have been collected in this setting. But in real-world conditions sound sources often move relative to the listener and listeners move their heads, often to better disambiguate front from back and more accurately localize ». SPHERE est une approche qui a fait ses preuves, aux mondes de l'industrie audio prothétique de s'en emparer.

☞ Des perspectives de collaborations avec des experts d'horizons différents pour une stratégie commune visant à la compréhension des mécanismes de SA chez les patients porteurs d'aides auditives. Une réflexion personnelle m'amène à prendre conscience du peu de mutualisation des stratégies en recherche dans le domaine de la spatialisation auditive. Cette faiblesse d'échange entre professionnels de disciplines différentes (audioprothésistes < > neuroscientifiques < > industriels, etc.) s'estompent progressivement par les exigences des appels à projet qui demande l'élaboration d'une approche interdisciplinaire mêlant des expertises différentes (scientifiques, industriels, etc.). Les projets interdisciplinaires sont de plus en plus nombreux (voir le programme de la journée scientifique de l'ISTR que j'ai organisée en juin 2023 et dédiée au thème de « L'interdisciplinarité entre science et rééducation en croisant les regards de chercheurs et de cliniciens » <https://jistr23.univ-lyon1.fr/fr/pages/jistr-2023-programme>). C'est dans l'optique de renforcer l'interdisciplinarité en recherche sur l'amélioration des implants cochléaires pour la spatialisation auditive, que je vais pouvoir inscrire deux étudiants doctorants dans le projet européen HORIZON MSCA « Cochlear implants and spatial hearing: Enabling access to the next dimension of hearing », ce projet CheriSH visant à tisser des liens entre des acteurs de la recherche européenne et des industriels.

Cette section marque la conclusion de ce document de recherche. Son élaboration m'a offert l'opportunité de consolider mes contributions réalisées et de mieux situer mes projets de recherche (engagés ou à venir). De plus, c'est à travers ces travaux de recherche que je peux m'épanouir dans ma mission d'enseignante-chercheuse au sein d'un institut de rééducation, et proposer dans mes enseignements des activités pratiques et des stages de recherche sur la spatialisation auditive pour les étudiants (rééducateurs en audioprothèse, en orthophonie, en neurosciences, etc.).

J'exprime ma sincère gratitude envers ceux qui ont accepté de lire ce travail et de le juger. Enfin, pour de rendre sa lecture agréable, j'ai tenté de limiter les considérations techniques et méthodologiques, qui part ailleurs, sont accessibles dans les articles publiés fournis en annexe.

Bibliographie

- Aggius-Vella, E., Campus, C., Finocchietti, S., & Gori, M. (2017). Audio Spatial Representation Around the Body. *Frontiers in Psychology*, 8, 1932. <https://doi.org/10.3389/fpsyg.2017.01932>
- Ahrens, A., Lund, K. D., Marschall, M., & Dau, T. (2019). Sound source localization with varying amount of visual information in virtual reality. *PLOS ONE*, 14(3), e0214603. <https://doi.org/10.1371/journal.pone.0214603>
- Ahveninen, J., Jääskeläinen, I. P., Raij, T., Bonmassar, G., Devore, S., Hämäläinen, M., Levänen, S., Lin, F.-H., Sams, M., Shinn-Cunningham, B. G., Witzel, T., & Belliveau, J. W. (2006). Task-modulated “what” and “where” pathways in human auditory cortex. *Proceedings of the National Academy of Sciences*, 103(39), 14608-14613. <https://doi.org/10.1073/pnas.0510480103>
- Akeroyd, M. A. (2008). Are individual differences in speech reception related to individual differences in cognitive ability? A survey of twenty experimental studies with normal and hearing-impaired adults. *International Journal of Audiology*, 47(sup2), S53-S71. <https://doi.org/10.1080/14992020802301142>
- Akeroyd, M. A., & Whitmer, W. M. (2016). Spatial Hearing and Hearing Aids. In G. R. Popelka, B. C. J. Moore, R. R. Fay, & A. N. Popper (Éds.), *Hearing Aids* (Vol. 56, p. 181-215). Springer International Publishing. https://doi.org/10.1007/978-3-319-33036-5_7
- Alain, C., Arnott, S. R., Hevenor, S., Graham, S., & Grady, C. L. (2001). “What” and “where” in the human auditory system. *Proceedings of the National Academy of Sciences*, 98(21), 12301-12306. <https://doi.org/10.1073/pnas.211209098>
- Alais, D., & Burr, D. (2004). The Ventriloquist Effect Results from Near-Optimal Bimodal Integration. *Current Biology*, 14(3), 257-262. <https://doi.org/10.1016/j.cub.2004.01.029>
- Alais, D., Newell, F., & Mamassian, P. (2010). Multisensory Processing in Review : From Physiology to Behaviour. *Seeing and Perceiving*, 23(1), 3-38. <https://doi.org/10.1163/187847510X488603>
- Andéol, G., & Simpson, B. D. (2016). Editorial : How, and Why, Does Spatial-Hearing Ability Differ among Listeners? What is the Role of Learning and Multisensory Interactions? *Frontiers in Neuroscience*, 10. <https://doi.org/10.3389/fnins.2016.00036>
- Angell, J. R., & Fite, W. (1901). *From the Psychological Laboratory of the University of Chicago : The monaural localization of sound. Psychological Review*. 225-246.
- Antunes, F. M., & Malmierca, M. S. (2021). Corticothalamic Pathways in Auditory Processing : Recent Advances and Insights From Other Sensory Systems. *Frontiers in Neural Circuits*, 15, 721186. <https://doi.org/10.3389/fncir.2021.721186>
- Arnott, S. R., Binns, M. A., Grady, C. L., & Alain, C. (2004). Assessing the auditory dual-pathway model in humans. *NeuroImage*, 22(1), 401-408. <https://doi.org/10.1016/j.neuroimage.2004.01.014>
- Aytekin, M., Moss, C. F., & Simon, J. Z. (2008). A Sensorimotor Approach to Sound Localization. *Neural Computation*, 20(3), 603-635. <https://doi.org/10.1162/neco.2007.12-05-094>
- Bahu, H., Carpentier, T., Noisternig, M., & Warusfel, O. (2016). Comparison of Different Egocentric Pointing Methods for 3D Sound Localization Experiments. *Acta Acustica United with Acustica*, 102(1), 107-118. <https://doi.org/10.3813/AAA.918928>
- Barone, P., Pavani, F., Farne, A., Truy, E., & Gaveau, V. (2018). Intégration multisensorielle chez le patient déficient auditif. In *Surdité, actualités, innovations et espoirs- Perception auditive*. In *Surdité, actualités, innovations et espoirs- Perception auditive* (Elsevier-Masson).
- Bayri, M., & Çiprut, A. (2020). The effects of behind-the-ear and off-the-ear sound processors on speech understanding performance in cochlear implant users. *Auris, Nasus, Larynx*, 47(6), 950-957. <https://doi.org/10.1016/j.anl.2020.05.025>
- Bergman, P., Lyxell, B., Harder, H., & Mäki-Torkko, E. (2020). The Outcome of Unilateral Cochlear Implantation in Adults : Speech Recognition, Health-Related Quality of Life and Level of Anxiety and Depression: a One- and Three-Year Follow-Up Study. *International Archives of Otorhinolaryngology*, 24(03), e338-e346. <https://doi.org/10.1055/s-0039-3399540>
- Best, V., Baumgartner, R., Lavandier, M., Majdak, P., & Kopčo, N. (2020). Sound Externalization : A Review of Recent Research. *Trends in Hearing*, 24, 233121652094839. <https://doi.org/10.1177/2331216520948390>
- Bishop, C. W., London, S., & Miller, L. M. (2012). Neural time course of visually enhanced echo suppression. *Journal of Neurophysiology*, 108(7), 1869-1883. <https://doi.org/10.1152/jn.00175.2012>
- Blauert, J. (1996). *Spatial Hearing: The Psychophysics of Human Sound Localization* (2^e éd.). The MIT Press. <https://doi.org/10.7551/mitpress/6391.001.0001>
- Bregman, A. S. (1990). *Auditory Scene Analysis: The Perceptual Organization of Sound*. The MIT Press. <https://doi.org/10.7551/mitpress/1486.001.0001>

- Brimijoin, W. O., & Akeroyd, M. A. (2012). The Role of Head Movements and Signal Spectrum in an Auditory Front/Back Illusion. *I-Perception*, 3(3), 179-182. <https://doi.org/10.1068/i7173sas>
- Brimijoin, W. O., Boyd, A. W., & Akeroyd, M. A. (2013). The Contribution of Head Movement to the Externalization and Internalization of Sounds. *PLoS ONE*, 8(12), e83068. <https://doi.org/10.1371/journal.pone.0083068>
- Bronkhorst, A. W. (2015). The cocktail-party problem revisited: Early processing and selection of multi-talker speech. *Attention, Perception, & Psychophysics*, 77(5), 1465-1487. <https://doi.org/10.3758/s13414-015-0882-9>
- Brunetti, M., Belardinelli, P., Caulo, M., Del Gratta, C., Della Penna, S., Ferretti, A., Lucci, G., Moretti, A., Pizzella, V., Tartaro, A., Torquati, K., Olivetti Belardinelli, M., & Romani, G. L. (2005). Human brain activation during passive listening to sounds from different locations: An fMRI and MEG study. *Human Brain Mapping*, 26(4), 251-261. <https://doi.org/10.1002/hbm.20164>
- Brungart, D. S., Durlach, N. I., & Rabinowitz, W. M. (1999). Auditory localization of nearby sources. II. Localization of a broadband source. *The Journal of the Acoustical Society of America*, 106(4), 1956-1968. <https://doi.org/10.1121/1.427943>
- Brungart, D. S., & Rabinowitz, W. M. (1999). Auditory localization of nearby sources. Head-related transfer functions. *The Journal of the Acoustical Society of America*, 106(3), 1465-1479. <https://doi.org/10.1121/1.427180>
- Brungart, D. S., Rabinowitz, W. M., & Durlach, N. I. (2000). Evaluation of response methods for the localization of nearby objects. *Perception & Psychophysics*, 62(1), 48-65. <https://doi.org/10.3758/BF03212060>
- Burr, D., & Alais, D. (2006). Chapter 14 Combining visual and auditory information. In *Progress in Brain Research* (Vol. 155, p. 243-258). Elsevier. [https://doi.org/10.1016/S0079-6123\(06\)55014-9](https://doi.org/10.1016/S0079-6123(06)55014-9)
- Carette, E., Van den Bogaert, T., Laureyns, M., & Wouters, J. (2014). Left-Right and Front-Back Spatial Hearing with Multiple Directional Microphone Configurations in Modern Hearing Aids. *Journal of the American Academy of Audiology*, 25(09), 791-803. <https://doi.org/10.3766/jaaa.25.9.2>
- Cornelio, P., Velasco, C., & Obrist, M. (2021). Multisensory Integration as per Technological Advances : A Review. *Frontiers in Neuroscience*, 15, 652611. <https://doi.org/10.3389/fnins.2021.652611>
- Davis, T. J., & Gifford, R. H. (2018). Spatial Release From Masking in Adults With Bilateral Cochlear Implants : Effects of Distracter Azimuth and Microphone Location. *Journal of Speech, Language, and Hearing Research*, 61(3), 752-761. https://doi.org/10.1044/2017_JSLHR-H-16-0441
- De Gelder, B., Vroomen, J., Annen, L., Masthof, E., & Hodiamont, P. (2003). Audio-visual integration in schizophrenia. *Schizophrenia Research*, 59(2-3), 211-218. [https://doi.org/10.1016/S0920-9964\(01\)00344-9](https://doi.org/10.1016/S0920-9964(01)00344-9)
- Degerman, A., Rinne, T., Salmi, J., Salonen, O., & Alho, K. (2006). Selective attention to sound location or pitch studied with fMRI. *Brain Research*, 1077(1), 123-134. <https://doi.org/10.1016/j.brainres.2006.01.025>
- Denk, F., Ewert, S. D., & Kollmeier, B. (2019). On the limitations of sound localization with hearing devices. *The Journal of the Acoustical Society of America*, 146(3), 1732-1744. <https://doi.org/10.1121/1.5126521>
- Dwyer, R. T., Roberts, J., & Gifford, R. H. (2020). Effect of Microphone Configuration and Sound Source Location on Speech Recognition for Adult Cochlear Implant Users with Current-Generation Sound Processors. *Journal of the American Academy of Audiology*, 31(8), 578-589. <https://doi.org/10.1055/s-0040-1709449>
- Etchemendy, P. E., Abregú, E., Calcagno, E. R., Eguia, M. C., Vechiatti, N., Iasi, F., & Vergara, R. O. (2017). Auditory environmental context affects visual distance perception. *Scientific Reports*, 7(1), 7189. <https://doi.org/10.1038/s41598-017-06495-3>
- Francl, A., & McDermott, J. H. (2022). Deep neural network models of sound localization reveal how perception is adapted to real-world environments. *Nature Human Behaviour*, 6(1), 111-133. <https://doi.org/10.1038/s41562-021-01244-z>
- Freigang, C., Richter, N., Rübsamen, R., & Ludwig, A. (2015). Age-related changes in sound localisation ability. *Cell and Tissue Research*, 361. <https://doi.org/10.1007/s00441-015-2230-8>
- Fritz, J. B., Elhilali, M., David, S. V., & Shamma, S. A. (2007). Auditory attention—Focusing the searchlight on sound. *Current Opinion in Neurobiology*, 17(4), 437-455. <https://doi.org/10.1016/j.conb.2007.07.011>
- Fuller, JamesH. (1992). Head movement propensity. *Experimental Brain Research*, 92(1). <https://doi.org/10.1007/BF00230391>
- Galvin, K. L., & Noble, W. (2013). Adaptation of the speech, spatial, and qualities of hearing scale for use with children, parents, and teachers. *Cochlear Implants International*, 14(3), 135-141. <https://doi.org/10.1179/1754762812Y.0000000014>
- Gandemer, L., Parsehian, G., Kronland-Martinet, R., & Bourdin, C. (2017). Spatial Cues Provided by Sound Improve Postural Stabilization: Evidence of a Spatial Auditory Map? *Frontiers in Neuroscience*, 11, 357. <https://doi.org/10.3389/fnins.2017.00357>
- Gatehouse, S., & Noble, W. (2004). The Speech, Spatial and Qualities of Hearing Scale (SSQ). *International Journal of Audiology*, 43(2), 85-99. <https://doi.org/10.1080/14992020400050014>
- Gates, G. A., Schmid, P., Kujawa, S. G., Nam, B., & D'Agostino, R. (2000). Longitudinal threshold changes in older men with audiometric notches. *Hearing Research*, 141(1-2), 220-228. [https://doi.org/10.1016/S0378-5955\(99\)00223-3](https://doi.org/10.1016/S0378-5955(99)00223-3)
- Gawęda, Ł., Li, E., Lavoie, S., Whitford, T. J., Moritz, S., & Nelson, B. (2018). Impaired action self-monitoring and cognitive confidence among ultra-high risk for psychosis and first-episode psychosis patients. *European Psychiatry*, 47, 67-75. <https://doi.org/10.1016/j.eurpsy.2017.09.003>
- Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences*, 15(1),

20-25. [https://doi.org/10.1016/0166-2236\(92\)90344-8](https://doi.org/10.1016/0166-2236(92)90344-8)

- Goossens, H. H. L. M., & van Opstal, A. J. (1999). Influence of Head Position on the Spatial Representation of Acoustic Targets. *Journal of Neurophysiology*, 81(6), 2720-2736. <https://doi.org/10.1152/jn.1999.81.6.2720>
- Grantham, D. W., Ashmead, D. H., Ricketts, T. A., Labadie, R. F., & Haynes, D. S. (2007a). Horizontal-Plane Localization of Noise and Speech Signals by Postlingually Deafened Adults Fitted With Bilateral Cochlear Implants*. *Ear & Hearing*, 28(4), 524-541. <https://doi.org/10.1097/AUD.0b013e31806dc21a>
- Grantham, D. W., Ashmead, D. H., Ricketts, T. A., Labadie, R. F., & Haynes, D. S. (2007b). Horizontal-Plane Localization of Noise and Speech Signals by Postlingually Deafened Adults Fitted With Bilateral Cochlear Implants*. *Ear & Hearing*, 28(4), 524-541. <https://doi.org/10.1097/AUD.0b013e31806dc21a>
- Grantham, D. W., Ricketts, T. A., Ashmead, D. H., Labadie, R. F., & Haynes, D. S. (2008). Localization by Postlingually Deafened Adults Fitted With a Single Cochlear Implant. *The Laryngoscope*, 118(1), 145-151. <https://doi.org/10.1097/MLG.0b013e31815661f9>
- Groh, J. M., & Sparks, D. L. (1992). Two models for transforming auditory signals from head-centered to eye-centered coordinates. *Biological Cybernetics*, 67(4), 291-302. <https://doi.org/10.1007/BF02414885>
- Gröhn, C., Norgren, E., & Eriksson, L. (2022). A systematic review of the neural correlates of multisensory integration in schizophrenia. *Schizophrenia Research: Cognition*, 27, 100219. <https://doi.org/10.1016/j.scog.2021.100219>
- Guterman, Y., & Klein, E. (1991). The role of head movement and pinnae in auditory localization in schizophrenia and psychosis. *Schizophrenia Research*, 6(1), 67-73. [https://doi.org/10.1016/0920-9964\(91\)90022-J](https://doi.org/10.1016/0920-9964(91)90022-J)
- Gutfreund, Y., & King, A. J. (2012). What Is the Role of Vision in the Development of the Auditory Space Map? In B. E. Stein (Éd.), *The New Handbook of Multisensory Processing* (p. 573-588). The MIT Press. <https://doi.org/10.7551/mitpress/8466.003.0052>
- Haber, L., Haber, R. N., Penningroth, S., Novak, K., & Radgowski, H. (1993). Comparison of Nine Methods of Indicating the Direction to Objects : Data from Blind Adults. *Perception*, 22(1), 35-47. <https://doi.org/10.1068/p220035>
- Heffner, H. E., & Heffner, R. S. (2014). The Behavioral Study of Mammalian Hearing. In A. N. Popper & R. R. Fay (Éds.), *Perspectives on Auditory Research* (Vol. 50, p. 269-285). Springer New York. https://doi.org/10.1007/978-1-4614-9102-6_15
- Heffner, R. S., & Heffner, H. E. (1992). Visual factors in sound localization in mammals. *The Journal of Comparative Neurology*, 317(3), 219-232. <https://doi.org/10.1002/cne.903170302>
- Hofman, P., & Van Opstal, A. (2003). Binaural weighting of pinna cues in human sound localization. *Experimental Brain Research*, 148(4), 458-470. <https://doi.org/10.1007/s00221-002-1320-5>
- Honda, A., Shibata, H., Gyoba, J., Saitou, K., Iwaya, Y., & Suzuki, Y. (2007). Transfer effects on sound localization performances from playing a virtual three-dimensional auditory game. *Applied Acoustics*, 68(8), 885-896. <https://doi.org/10.1016/j.apacoust.2006.08.007>
- Honda, A., Shibata, H., Hidaka, S., Gyoba, J., Iwaya, Y., & Suzuki, Y. (2013). Effects of Head Movement and Proprioceptive Feedback in Training of Sound Localization. *I-Perception*, 4(4), 253-264. <https://doi.org/10.1068/i0522>
- Iliadou, V. (Vivian), Apalla, K., Kaprinis, S., Nimatoudis, I., Kaprinis, G., & Iacovides, A. (2013). Is Central Auditory Processing Disorder Present in Psychosis? *American Journal of Audiology*, 22(2), 201-208. [https://doi.org/10.1044/1059-0889\(2013/12-0073\)](https://doi.org/10.1044/1059-0889(2013/12-0073))
- Jansen, S., Luts, H., Wagener, K. C., Kollmeier, B., Del Rio, M., Dauman, R., James, C., Fraysse, B., Vormès, E., Frachet, B., Wouters, J., & Van Wieringen, A. (2012). Comparison of three types of French speech-in-noise tests : A multi-center study. *International Journal of Audiology*, 51(3), 164-173. <https://doi.org/10.3109/14992027.2011.633568>
- Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin*, 114(1), 3-28. <https://doi.org/10.1037/0033-2909.114.1.3>
- Jones, H. G., Kan, A., & Litovsky, R. Y. (2016). The Effect of Microphone Placement on Interaural Level Differences and Sound Localization Across the Horizontal Plane in Bilateral Cochlear Implant Users. *Ear & Hearing*, 37(5), e341-e345. <https://doi.org/10.1097/AUD.0000000000000297>
- Jones, H., Kan, A., & Litovsky, R. Y. (2014). Comparing Sound Localization Deficits in Bilateral Cochlear-Implant Users and Vocoder Simulations With Normal-Hearing Listeners. *Trends in Hearing*, 18, 233121651455457. <https://doi.org/10.1177/2331216514554574>
- Kasai, K., Shenton, M. E., Salisbury, D. F., Hirayasu, Y., Lee, C.-U., Ciszewski, A. A., Yurgelun-Todd, D., Kikinis, R., Jolesz, F. A., & McCarley, R. W. (2003). Progressive Decrease of Left Superior Temporal Gyrus Gray Matter Volume in Patients With First-Episode Schizophrenia. *American Journal of Psychiatry*, 160(1), 156-164. <https://doi.org/10.1176/appi.ajp.160.1.156>
- Keidser, G., O'Brien, A., Hain, J.-U., McLelland, M., & Yeend, I. (2009). The effect of frequency-dependent microphone directionality on horizontal localization performance in hearing-aid users. *International Journal of Audiology*, 48(11), 789-803. <https://doi.org/10.3109/14992020903036357>
- Kim, J., Barnett-Cowan, M., & Macpherson, E. A. (2013). *Integration of auditory input with vestibular and neck proprioceptive information in the interpretation of dynamic sound localization cues*. 050142-050142. <https://doi.org/10.1121/1.4799748>
- Knudsen, E. I., & Knudsen, P. F. (1985). Vision Guides the Adjustment of Auditory Localization in Young Barn Owls. *Science*, 230(4725), 545-548. <https://doi.org/10.1126/science.4048948>

- Kolarik, A. J., Moore, B. C. J., Zahorik, P., Cirstea, S., & Pardhan, S. (2016). Auditory distance perception in humans : A review of cues, development, neuronal bases, and effects of sensory loss. *Attention, Perception, & Psychophysics*, 78(2), 373-395. <https://doi.org/10.3758/s13414-015-1015-1>
- Kolberg, E. R., Sheffield, S. W., Davis, T. J., Sunderhaus, L. W., & Gifford, R. H. (2015). Cochlear Implant Microphone Location Affects Speech Recognition in Diffuse Noise. *Journal of the American Academy of Audiology*, 26(01), 051-058. <https://doi.org/10.3766/jaaa.26.1.6>
- Kral, A., & Sharma, A. (2023). Crossmodal plasticity in hearing loss. *Trends in Neurosciences*, 46(5), 377-393. <https://doi.org/10.1016/j.tins.2023.02.004>
- Kumpik, D. P., Campbell, C., Schnupp, J. W. H., & King, A. J. (2019). Re-weighting of Sound Localization Cues by Audiovisual Training. *Frontiers in Neuroscience*, 13, 1164. <https://doi.org/10.3389/fnins.2019.01164>
- Kumpik, D. P., Kacelnik, O., & King, A. J. (2010). Adaptive Reweighting of Auditory Localization Cues in Response to Chronic Unilateral Earplugging in Humans. *The Journal of Neuroscience*, 30(14), 4883-4894. <https://doi.org/10.1523/JNEUROSCI.5488-09.2010>
- Kwak, C., & Han, W. (2021). The Effectiveness of Hearing Protection Devices : A Systematic Review and Meta-Analysis. *International Journal of Environmental Research and Public Health*, 18(21), 11693. <https://doi.org/10.3390/ijerph182111693>
- Lad, M., Sedley, W., & Griffiths, T. D. (2022). Sensory Loss and Risk of Dementia. *The Neuroscientist*, 107385842211260. <https://doi.org/10.1177/10738584221126090>
- Lavallé, L., Dondé, C., Gawęda, L., Brunelin, J., & Mondino, M. (2021). Impaired self-recognition in individuals with no full-blown psychotic symptoms represented across the continuum of psychosis : A meta-analysis. *Psychological Medicine*, 51(16), 2864-2874. <https://doi.org/10.1017/S003329172000152X>
- Lewald, J., & Ehrenstein, WalterH. (1996). The effect of eye position on auditory lateralization. *Experimental Brain Research*, 108(3). <https://doi.org/10.1007/BF00227270>
- Litovsky, R. Y., Parkinson, A., & Arcaroli, J. (2009). Spatial Hearing and Speech Intelligibility in Bilateral Cochlear Implant Users. *Ear & Hearing*, 30(4), 419-431. <https://doi.org/10.1097/AUD.0b013e3181a165be>
- Livingston, G., Sommerlad, A., Orgeta, V., Costafreda, S. G., Huntley, J., Ames, D., Ballard, C., Banerjee, S., Burns, A., Cohen-Mansfield, J., Cooper, C., Fox, N., Gitlin, L. N., Howard, R., Kales, H. C., Larson, E. B., Ritchie, K., Rockwood, K., Sampson, E. L., ... Mukadam, N. (2017). Dementia prevention, intervention, and care. *The Lancet*, 390(10113), 2673-2734. [https://doi.org/10.1016/S0140-6736\(17\)31363-6](https://doi.org/10.1016/S0140-6736(17)31363-6)
- Loughrey, D. G., Kelly, M. E., Kelley, G. A., Brennan, S., & Lawlor, B. A. (2018). Association of Age-Related Hearing Loss With Cognitive Function, Cognitive Impairment, and Dementia : A Systematic Review and Meta-analysis. *JAMA Otolaryngology–Head & Neck Surgery*, 144(2), 115. <https://doi.org/10.1001/jamaoto.2017.2513>
- Macpherson, E. A. (2015). On the role of visual information about head motion in the interpretation of dynamic interaural cues for front/back sound localization. *Journal of the Acoustical Society of America*, 137(4_Supplement), 2203-2203. <https://doi.org/10.1121/1.4920010>
- Majdak, P., Goupell, M. J., & Laback, B. (2010). 3-D localization of virtual sound sources : Effects of visual environment, pointing method, and training. *Attention, Perception, & Psychophysics*, 72(2), 454-469. <https://doi.org/10.3758/APP.72.2.454>
- Mantokoudis, G., Kompis, M., Vischer, M., Häusler, R., Caversaccio, M., & Senn, P. (2011). In-the-Canal Versus Behind-the-Ear Microphones Improve Spatial Discrimination on the Side of the Head in Bilateral Cochlear Implant Users. *Otology & Neurotology*, 32(1), 1-6. <https://doi.org/10.1097/MAO.0b013e3182001cac>
- Matthews, N., Todd, J., Mannion, D. J., Finnigan, S., Catts, S., & Michie, P. T. (2013). Impaired processing of binaural temporal cues to auditory scene analysis in schizophrenia. *Schizophrenia Research*, 146(1-3), 344-348. <https://doi.org/10.1016/j.schres.2013.02.013>
- Mayo, P. G., & Goupell, M. J. (2020). Acoustic factors affecting interaural level differences for cochlear-implant users. *The Journal of the Acoustical Society of America*, 147(4), EL357-EL362. <https://doi.org/10.1121/10.0001088>
- McAnally, K. I., & Martin, R. L. (2014). Sound localization with head movement : Implications for 3-d audio displays. *Frontiers in Neuroscience*, 8. <https://doi.org/10.3389/fnins.2014.00210>
- McLachlan, G., Majdak, P., Reijnders, J., Mihocic, M., & Peremans, H. (2023). Dynamic spectral cues do not affect human sound localization during small head movements. *Frontiers in Neuroscience*, 17, 1027827. <https://doi.org/10.3389/fnins.2023.1027827>
- Medina-Garin, D., Dia, A., Bedubourg, G., Deparis, X., Berger, F., & Michel, R. (2016). Acute acoustic trauma in the French armed forces during 2007–2014. *Noise and Health*, 18(85), 297. <https://doi.org/10.4103/1463-1741.195802>
- Middlebrooks, J. C. (2015). Sound localization. In *Handbook of Clinical Neurology* (Vol. 129, p. 99-116). Elsevier. <https://doi.org/10.1016/B978-0-444-62630-1.00006-8>
- Middlebrooks, J. C., & Green, D. M. (1991). Sound Localization by Human Listeners. *Annual Review of Psychology*, 42(1), 135-159. <https://doi.org/10.1146/annurev.ps.42.020191.001031>
- Millet, B., Snapp, H. A., Rajguru, S. M., & Schaefer Solle, N. (2023). Prevalence of Hearing Loss and Perceptions of Hearing Health and Protection among Florida Firefighters. *International Journal of Environmental Research and Public Health*, 20(5), 3826. <https://doi.org/10.3390/ijerph20053826>
- Mondor, T. A., & Zatorre, R. J. (1995). Shifting and focusing auditory spatial attention. *Journal of Experimental Psychology*:

Human Perception and Performance, 21(2), 387-409. <https://doi.org/10.1037/0096-1523.21.2.387>

- 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>
- Moulin, A., Vergne, J., Gallego, S., & Micheyl, C. (2019). A New Speech, Spatial, and Qualities of Hearing Scale Short-Form : Factor, Cluster, and Comparative Analyses. *Ear & Hearing*, 40(4), 938-950. <https://doi.org/10.1097/AUD.0000000000000675>
- Natarajan, N., Batts, S., & Stankovic, K. M. (2023). Noise-Induced Hearing Loss. *Journal of Clinical Medicine*, 12(6), 2347. <https://doi.org/10.3390/jcm12062347>
- Nava, E., Bottari, D., Portioli, G., Bonfioli, F., Beltrame, M. A., Formigoni, P., & Pavani, F. (2009). Hearing again with two ears : Recovery of spatial hearing after bilateral cochlear implantation. *Neuropsychologia*, 47(3), 928-932. <https://doi.org/10.1016/j.neuropsychologia.2008.11.020>
- Noble, W., Byrne, D., & Lepage, B. (1994). Effects on sound localization of configuration and type of hearing impairment. *The Journal of the Acoustical Society of America*, 95(2), 992-1005. <https://doi.org/10.1121/1.408404>
- Noble, W. G. (1981). Earmuffs, exploratory head movements, and horizontal and vertical sound localization. *The Journal of Auditory Research*, 21(1), 1-12.
- Nordvik, Ø., Laugen Heggdal, P. O., Brännström, J., Vassbotn, F., Aarstad, A. K., & Aarstad, H. J. (2018). Generic quality of life in persons with hearing loss : A systematic literature review. *BMC Ear, Nose and Throat Disorders*, 18(1), 1. <https://doi.org/10.1186/s12901-018-0051-6>
- Oberfeld, D., & Klöckner-Nowotny, F. (2016). Individual differences in selective attention predict speech identification at a cocktail party. *eLife*, 5, e16747. <https://doi.org/10.7554/eLife.16747>
- Oldfield, S. R., & Parker, S. P. A. (1984). Acuity of Sound Localisation : A Topography of Auditory Space. I. Normal Hearing Conditions. *Perception*, 13(5), 581-600. <https://doi.org/10.1068/p130581>
- Parseihian, G., Jouffrais, C., & Katz, B. F. G. (2014). Reaching nearby sources : Comparison between real and virtual sound and visual targets. *Frontiers in Neuroscience*, 8. <https://doi.org/10.3389/fnins.2014.00269>
- Pavani, F., Farnè, A., & Làdavas, E. (2003). Task-dependent visual coding of sound position in visuospatial neglect patients: *NeuroReport*, 14(1), 99-103. <https://doi.org/10.1097/00001756-200301200-00019>
- Pavani, F., Farnè, A., Tru, E., & Gaveau, V. (2018). Perception auditive spatiale. In *Surdité, actualités, innovations et espoirs-Perception auditive 2018* (ELSEVIER-MASSON).
- Pavani, F., Husain, M., & Driver, J. (2008). Eye-movements intervening between two successive sounds disrupt comparisons of auditory location. *Experimental Brain Research*, 189(4), 435-449. <https://doi.org/10.1007/s00221-008-1440-7>
- Pavani, F., Meneghello, F., & Làdavas, E. (2001). Deficit of auditory space perception in patients with visuospatial neglect. *Neuropsychologia*, 39(13), 1401-1409. [https://doi.org/10.1016/S0028-3932\(01\)00060-4](https://doi.org/10.1016/S0028-3932(01)00060-4)
- Pearson, J. D., Morrell, C. H., Gordon-Salant, S., Brant, L. J., Metter, E. J., Klein, L. L., & Fozard, J. L. (1995). Gender differences in a longitudinal study of age-associated hearing loss. *The Journal of the Acoustical Society of America*, 97(2), 1196-1205. <https://doi.org/10.1121/1.412231>
- Perrett, S., & Noble, W. (1997a). The contribution of head motion cues to localization of low-pass noise. *Perception & Psychophysics*, 59(7), 1018-1026. <https://doi.org/10.3758/BF03205517>
- Perrett, S., & Noble, W. (1997b). The effect of head rotations on vertical plane sound localization. *The Journal of the Acoustical Society of America*, 102(4), 2325-2332. <https://doi.org/10.1121/1.419642>
- Perrin, M. A., Butler, P. D., DiCostanzo, J., Forchelli, G., Silipo, G., & Javitt, D. C. (2010). Spatial localization deficits and auditory cortical dysfunction in schizophrenia. *Schizophrenia Research*, 124(1-3), 161-168. <https://doi.org/10.1016/j.schres.2010.06.004>
- Pollack, I., & Rose, M. (1967). Effect of head movement on the localization of sounds in the equatorial plane. *Perception & Psychophysics*, 2(12), 591-596. <https://doi.org/10.3758/BF03210274>
- Pouryaghoub, G., Mehrdad, R., & Pourhosein, S. (2017). Noise-Induced hearing loss among professional musicians. *Journal of Occupational Health*, 59(1), 33-37. <https://doi.org/10.1539/joh.16-0217-OA>
- Rabini, G., Altobelli, E., & Pavani, F. (2019). Interactions between egocentric and allocentric spatial coding of sounds revealed by a multisensory learning paradigm. *Scientific Reports*, 9(1), 7892. <https://doi.org/10.1038/s41598-019-44267-3>
- Rakerd, B., Vander Velde, T. J., & Hartmann, W. M. (1998). Sound localization in the median sagittal plane by listeners with presbycusis. *Journal of the American Academy of Audiology*, 9(6), 466-479.
- Ramage, Erin. M., Weintraub, D. M., Allen, D. N., & Snyder, J. S. (2012). Evidence for stimulus-general impairments on auditory stream segregation tasks in schizophrenia. *Journal of Psychiatric Research*, 46(12), 1540-1545. <https://doi.org/10.1016/j.jpsychires.2012.08.028>
- Rauschecker, J. P. (2018). Where, When, and How : Are they all sensorimotor? Towards a unified view of the dorsal pathway in vision and audition. *Cortex*, 98, 262-268. <https://doi.org/10.1016/j.cortex.2017.10.020>
- Rayleigh, Lord. (1907). XII. *On our perception of sound direction*. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 13(74), 214-232. <https://doi.org/10.1080/14786440709463595>
- Seiwerth, I. (2023). Interaction of Hearing and Balance. *Laryngo-Rhino-Otologie*, 102(S 01), S35-S49. <https://doi.org/10.1055/a-1960-4641>
- Serafin, S., Adjorlu, A., & Percy-Smith, L. M. (2023). A Review of Virtual Reality for Individuals with Hearing Impairments.

Multimodal Technologies and Interaction, 7(4), 36. <https://doi.org/10.3390/mti7040036>

- Serino, A. (2019). Peripersonal space (PPS) as a multisensory interface between the individual and the environment, defining the space of the self. *Neuroscience & Biobehavioral Reviews*, 99, 138-159. <https://doi.org/10.1016/j.neubiorev.2019.01.016>
- Shende, S. A., & Mudar, R. A. (2023). Cognitive control in age-related hearing loss : A narrative review. *Hearing Research*, 436, 108814. <https://doi.org/10.1016/j.heares.2023.108814>
- Shield, B. (2019). *Evaluation of the social and economic costs of hearing impairment*. London: Brunel University.
- Shinn-Cunningham, B., Best, V., & Lee, A. K. C. (2017). Auditory Object Formation and Selection. In J. C. Middlebrooks, J. Z. Simon, A. N. Popper, & R. R. Fay (Éds.), *The Auditory System at the Cocktail Party* (Vol. 60, p. 7-40). Springer International Publishing. https://doi.org/10.1007/978-3-319-51662-2_2
- 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>
- Sladen, D. P., Peterson, A., Schmitt, M., Olund, A., Teece, K., Dowling, B., DeJong, M., Breneman, A., Beatty, C. W., Carlson, M. L., Neff, B. A., Hughes-Borst, B., & Driscoll, C. L. (2017). Health-related quality of life outcomes following adult cochlear implantation : A prospective cohort study. *Cochlear Implants International*, 18(3), 130-135. <https://doi.org/10.1080/14670100.2017.1293203>
- Snapp, H. A., Schaefer Solle, N., Millet, B., & Rajguru, S. M. (2022). Subclinical Hearing Deficits in Noise-Exposed Firefighters. *International Journal of Environmental Research and Public Health*, 19(17), 11028. <https://doi.org/10.3390/ijerph191711028>
- 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>
- Strelnikov, K., Rosito, M., & Barone, P. (2011). Effect of Audiovisual Training on Monaural Spatial Hearing in Horizontal Plane. *PLoS ONE*, 6(3), e18344. <https://doi.org/10.1371/journal.pone.0018344>
- Tabry, V., Zatorre, R. J., & Voss, P. (2013). The influence of vision on sound localization abilities in both the horizontal and vertical planes. *Frontiers in Psychology*, 4. <https://doi.org/10.3389/fpsyg.2013.00932>
- Távora-Vieira, D., De Ceulaer, G., Govaerts, P. J., & Rajan, G. P. (2015). Cochlear Implantation Improves Localization Ability in Patients With Unilateral Deafness. *Ear & Hearing*, 36(3), e93-e98. <https://doi.org/10.1097/AUD.0000000000000130>
- Thurlow, W. R., & Runge, P. S. (1967). Effect of Induced Head Movements on Localization of Direction of Sounds. *The Journal of the Acoustical Society of America*, 42(2), 480-488. <https://doi.org/10.1121/1.1910604>
- Tonelli, A., Brayda, L., & Gori, M. (2015). Task-dependent calibration of auditory spatial perception through environmental visual observation. *Frontiers in Systems Neuroscience*, 9. <https://doi.org/10.3389/fnsys.2015.00084>
- Van Barneveld, D. C. P. B. M., Binkhorst, F., & Van Opstal, A. J. (2011). Absence of compensation for vestibular-evoked passive head rotations in human sound localization : Sound localization under passive vestibular stimulation. *European Journal of Neuroscience*, 34(7), 1149-1160. <https://doi.org/10.1111/j.1460-9568.2011.07844.x>
- Van den Bogaert, T., Carette, E., & Wouters, J. (2011). Sound source localization using hearing aids with microphones placed behind-the-ear, in-the-canal, and in-the-pinna. *International Journal of Audiology*, 50(3), 164-176.
- Van den Bogaert, T., Klasen, T. J., Moonen, M., Van Deun, L., & Wouters, J. (2006). Horizontal localization with bilateral hearing aids: Without is better than with. *The Journal of the Acoustical Society of America*, 119(1), 515-526. <https://doi.org/10.1121/1.2139653>
- van der Heijden, K., Rauschecker, J. P., de Gelder, B., & Formisano, E. (2019). Cortical mechanisms of spatial hearing. *Nature Reviews Neuroscience*, 20(10), 609-623. <https://doi.org/10.1038/s41583-019-0206-5>
- van der Zwaag, W., Gentile, G., Gruetter, R., Spierer, L., & Clarke, S. (2011). Where sound position influences sound object representations : A 7-T fMRI study. *NeuroImage*, 54(3), 1803-1811. <https://doi.org/10.1016/j.neuroimage.2010.10.032>
- Van Grootel, T. J., Van Wanrooij, M. M., & Van Opstal, A. J. (2011). Influence of Static Eye and Head Position on Tone-Evoked Gaze Shifts. *Journal of Neuroscience*, 31(48), 17496-17504. <https://doi.org/10.1523/JNEUROSCI.5030-10.2011>
- Van Wanrooij, M. M., & Van Opstal, A. J. (2005). Relearning Sound Localization with a New Ear. *The Journal of Neuroscience*, 25(22), 5413-5424. <https://doi.org/10.1523/JNEUROSCI.0850-05.2005>
- van Hoesel, R. J. M., & Tyler, R. S. (2003). Speech perception, localization, and lateralization with bilateral cochlear implants. *The Journal of the Acoustical Society of America*, 113(3), 1617-1630. <https://doi.org/10.1121/1.1539520>
- Vliegen, J., & Van Opstal, A. J. (2004). The influence of duration and level on human sound localization. *The Journal of the Acoustical Society of America*, 115(4), 1705-1713. <https://doi.org/10.1121/1.1687423>
- Wightman, F. L., & Kistler, D. J. (1999). Resolution of front-back ambiguity in spatial hearing by listener and source movement. *The Journal of the Acoustical Society of America*, 105(5), 2841-2853. <https://doi.org/10.1121/1.426899>
- Williams, L. E., Light, G. A., Braff, D. L., & Ramachandran, V. S. (2010). Reduced multisensory integration in patients with schizophrenia on a target detection task. *Neuropsychologia*, 48(10), 3128-3136. <https://doi.org/10.1016/j.neuropsychologia.2010.06.028>
- Wilson, B. S. (2008). Cochlear implants : Current designs and future possibilities. *The Journal of Rehabilitation Research and Development*, 45(5), 695-730. <https://doi.org/10.1682/JRRD.2007.10.0173>
- Wood, K. C., Town, S. M., & Bizley, J. K. (2019). Neurons in primary auditory cortex represent sound source location in a cue-invariant manner. *Nature Communications*, 10(1), 3019. <https://doi.org/10.1038/s41467-019-10868-9>

- World Health Organization. (2021). *World report on hearing*. World Health Organization.
- Yehudai, N., Fink, N., Shpriz, M., & Marom, T. (2017). Acute Acoustic Trauma among Soldiers during an Intense Combat. *Journal of the American Academy of Audiology*, 28(05), 436-443. <https://doi.org/10.3766/jaaa.16043>
- Yost, W. A. (1998). Spatial Hearing : The Psychophysics of Human Sound Localization, Revised Edition: *Ear and Hearing*, 19(2), 167. <https://doi.org/10.1097/00003446-199804000-00009>
- Yost, W. A., Pastore, M. T., & Dorman, M. F. (2020). Sound source localization is a multisystem process. *Acoustical Science and Technology*, 41(1), 113-120. <https://doi.org/10.1250/ast.41.113>
- Zahorik, P. (2002). Assessing auditory distance perception using virtual acoustics. *The Journal of the Acoustical Society of America*, 111(4), 1832-1846. <https://doi.org/10.1121/1.1458027>
- Zhang, T., Dorman, M. F., & Spahr, A. J. (2010). Information From the Voice Fundamental Frequency (F0) Region Accounts for the Majority of the Benefit When Acoustic Stimulation Is Added to Electric Stimulation. *Ear & Hearing*, 31(1), 63-69. <https://doi.org/10.1097/AUD.0b013e3181b7190c>
- Zhong, X., & Yost, W. A. (2013). Relationship between Postural Stability and Spatial Hearing. *Journal of the American Academy of Audiology*, 24(09), 782-788. <https://doi.org/10.3766/jaaa.24.9.3>

ANNEXES

Sélection d'articles publiés et exposés en contribution dans ce manuscrit

Gaveau V., Coudert A, Salemme R, Koun E, Desoche C, Truy E, Farnè A, Pavani F. Benefits of active listening during 3D sound localization. *Exp Brain Res.* 2022 Nov;240(11):2817-2833. doi: 10.1007/s00221-022-06456-x. Epub 2022 Sep 7.

Valzolgher C., Alzaher M., **Gaveau V.**, Coudert A., Marx M., Truy E., Barone, P., Farne A., Pavani F. Capturing visual attention with perturbed auditory spatial cues. *Trends in Hearing.* 2023 Jan- Dec;27:23312165231182289. doi: 10.1177/23312165231182289. PMID: 37611181.

Coudert A, Reilly KT, Truy E, Moulin A, **Gaveau V.** A new tool for the assessment of speech understanding and spatial hearing difficulties in children: the Kid-SSQ questionnaire. *Eur Arch Otorhinolaryngol.* 2023 Feb 10. doi: 10.1007/s00405-023-07846-9. Epub ahead of print. PMID: 36763152.

Valzolgher C, Bouzaid S, Grenouillet S, Gatel J, Ratenet L, Murenu F, Verdelet G, Salemme R, **Gaveau V.**, Coudert A, Hermann R, Truy E, Farnè A, Pavani F. Training spatial hearing in unilateral cochlear implant users through reaching to sounds in virtual reality. *Eur Arch Otorhinolaryngol.* 2023 Mar 11. doi: 10.1007/s00405-023-07886-1

Coudert A, Verdelet G, Reilly KT, Truy E, **Gaveau V.** Intensive Training of Spatial Hearing Promotes Auditory Abilities of Bilateral Cochlear Implant Adults: A Pilot Study. *Ear Hear.* 2023 Jan-Feb 01;44(1):61-76. doi: 10.1097/AUD.0000000000001256.

Valzolgher C, Gatel J, Bouzaid S, Grenouillet S, Todeschini M, Verdelet G, Salemme R, **Gaveau V.**, Truy E, Farnè A, Pavani F. Reaching to Sounds Improves Spatial Hearing in Bilateral Cochlear Implant Users. *Ear Hear.* 2023 Jan-Feb 01;44(1):189-198. doi: 10.1097/AUD.0000000000001267. Epub 2022 Aug 19. PMID: 35982520.

Valzolgher C, Todeschini M, Verdelet G, Gatel J, Salemme R, **Gaveau V.**, Truy E, Farnè A, Pavani F. Adapting to altered auditory cues: Generalization from manual reaching to head pointing. *PLoS One.* 2022 Apr 14;17(4):e0263509. doi: 10.1371/journal.pone.0263509.

Coudert A., **Gaveau V.**, et al. Spatial hearing difficulties in bilateral cochlear implant children improve with head movements. *Ear and Hearing.* *Ear Hear.* 2022 Jan/Feb;43(1):192-205. doi: 10.1097/AUD.0000000000001090

Valzolgher C, Verdelet G, Salemme R, Lombardi L, **Gaveau V.**, et al. Reaching to sounds in virtual reality: A multisensory-motor approach to promote adaptation to altered auditory cues. *Neuropsychologia.* 2020 Dec; 149:107665.

Valzolgher C., Alzhaler M., Gessa E., Todeschini M., Nieto P., Verdelet G., Salemme S, **Gaveau V.**, et al. The Impact of a Visual Spatial Frame on Real Sound-Source Localization in Virtual Reality. *Current Research in Behavioral Sciences* 1. Nov. 2020. 100003.

Verdelet, G., Salemme R., Desoche C., Volland F., Farne A., Coudert A., Hermann R., Truy E., **Gaveau V.** and Pavani F. Assessing Spatial and Temporal Reliability of the Vive System as a Tool for Naturalistic Behavioural Research. In *2019 International Conference on 3D Immersion (IC3D)*, 1-8. Brussels, Belgium: IEEE, 2019.



Benefits of active listening during 3D sound localization

V. Gaveau^{1,2} · A. Coudert^{1,2,5} · R. Salemmé^{1,2,3} · E. Koun^{1,2} · C. Desoche^{1,2,3} · E. Truy^{1,2,5} · A. Farné^{1,2,3} · F. Pavani^{1,2,4}

Received: 5 January 2022 / Accepted: 28 August 2022
© The Author(s) 2022

Abstract

In everyday life, sound localization entails more than just the extraction and processing of auditory cues. When determining sound position in three dimensions, the brain also considers the available visual information (e.g., visual cues to sound position) and resolves perceptual ambiguities through active listening behavior (e.g., spontaneous head movements while listening). Here, we examined to what extent spontaneous head movements improve sound localization in 3D—azimuth, elevation, and depth—by comparing static vs. active listening postures. To this aim, we developed a novel approach to sound localization based on sounds delivered in the environment, brought into alignment thanks to a VR system. Our system proved effective for the delivery of sounds at predetermined and repeatable positions in 3D space, without imposing a physically constrained posture, and with minimal training. In addition, it allowed measuring participant behavior (hand, head and eye position) in real time. We report that active listening improved 3D sound localization, primarily by ameliorating accuracy and variability of responses in azimuth and elevation. The more participants made spontaneous head movements, the better was their 3D sound localization performance. Thus, we provide proof of concept of a novel approach to the study of spatial hearing, with potentials for clinical and industrial applications.

Keywords Spatial hearing · Virtual reality · Head movements · Motion tracking · Active perception

Introduction

Spatial hearing is a fundamental ability for humans and other animals. Accurate localization of sounds allows for the construction of maps of the environment beyond the limits of the visual field, guides head- and eye-orienting behavior, plays a crucial role in multisensory integration, supports auditory scene analysis and can improve discrimination of

auditory signals from noise. In everyday environments, spatial hearing is three dimensional, multisensory and active. We concurrently estimate azimuth, elevation and distance of sounds, and perceive the visual context in which they occur and often also the event that generated them (Kumpik et al. 2019). Most importantly, in real-life contexts, spatial hearing is an active process: we explore the auditory environment with head and body movements to resolve perceptual ambiguities in sound localization (Andéol & Simpson, 2016).

When spatial hearing abilities are investigated in the laboratory, several of these naturalistic aspects are overlooked. Although this choice is motivated by the aim to control the intervening variables that could contaminate the experimental design, it does not allow full appreciation of some aspects of natural behavior. For instance, although sounds are delivered in 3D and supposedly perceived in 3D space, the response is often limited to one or two dimensions at a time: e.g., participants indicate the angle of the sound source (1D) or its direction (2D) (Andéol and Simpson 2016; Haber et al. 1993; Oldfield and Parker 1984; Wightman and Kistler 1999). In addition, experimental setups often attempt to limit the use of visual cues. Some setups use visible barriers to occlude the sound sources (e.g., Nava et al. 2009;

Communicated by Bill J Yates.

✉ V. Gaveau
valerie.gaveau@inserm.fr

¹ Integrative Multisensory Perception Action & Cognition Team—ImpAct, Lyon Neuroscience Research Center, INSERM U1028, CNRS U5292, 16 Av. Doyen Lépine, BRON cedex, 69500 Lyon, France

² University of Lyon 1, Lyon, France

³ Neuro-immersion, Lyon, France

⁴ Center for Mind/Brain Sciences - CIMEC, University of Trento, Rovereto, Italy

⁵ ENT Departments, Hôpital Femme-Mère-Enfant and Edouard Herriot University Hospitals, Lyon, France

Pavani et al. 2001, 2003), others place participants inside a completely dark room (Goossens and van Opstal 1999; Van Barneveld et al. 2011; Van Grootel et al. 2011), blindfold them (e.g., (Ahrens et al. 2019; Bahu et al. 2016) or require them to keep their eyes closed (e.g., Brungart et al. 1999), thus preventing eye tracking. Yet, eye-orienting responses are relevant for behavior, as the encoding of sound position in retinocentric coordinates has an impact on sound localization (Groh and Sparks 1992; Lewald and Ehrenstein 1996; Maddox et al. 2014; Pavani et al. 2008).

One constraint frequently imposed on participants concerns head position. It is common practice to limit head movements while participants listen to sounds, to ensure reproducible stimulation at the ears across trials and participants. This can be achieved by using a chin rest (Brungart et al. 1999; Litovsky et al. 2009; Pavani et al. 2008, 2008; Távora-Vieira et al. 2015), or by limiting target sound duration to few hundred milliseconds, so that participants do not have time to plan and execute head movements during sound emission (e.g., Ahrens et al. 2019). Even when the response is collected through head movements, these often occur after the sound is finished (e.g., Bahu et al. 2016). These approaches are well motivated when the experimental aim is to achieve full control over the auditory cues reaching the ears. Humans indeed estimate the location of a sound source by combining the auditory cues derived from one ear (monaural cues), with those available at both ears (binaural cues, i.e., interaural level differences ILD and interaural time differences ITD). A rich body of work has examined the specific contributions of auditory cues to sound localization (Middlebrooks 2015). However, limiting head movements during listening brings the drawback that it excludes the contribution of head motion during sound emission in sound localization performance.

In natural listening, head motion is spontaneous and almost ubiquitous. The importance of head movements for sound localization has been remarked since the 1940s (Wallach 1940), with pioneering works in the late 1990s (Perrett and Noble 1997a; Wightman and Kistler 1999). Psychoacoustic studies demonstrate that head motion improves sound localization in humans (Brimijoin et al. 2013; Honda et al. 2013; McAnally and Martin 2014; Perrett and Noble 1997a, b; Pollack and Rose 1967; Vliegen and Van Opstal 2004; Wightman and Kistler 1999), in monkeys (Populin 2006), and cats (Tollin et al. 2005). It is generally assumed that head movements are taken into account during the computation of sound-source coordinate (Goossens and van Opstal 1999). Neck muscle stimulation (Lewald et al. 1999) or cold water in the ear canal (Lewald and Karnath 2000), which alters the perceived position of the head, produces a shift of auditory localization. The continuous integration of head-motion signals

(proprioceptive, vestibular and efferent copy signals) contributes to sound localization (Genzel et al. 2016, 2018) and provides a more stable percept of the sound source (Vliegen and Van Opstal 2004).

In the present study, we examined the role of head movements on 3D sound localization by comparing the abilities of adult listeners when their head remains static throughout sound delivery or instead it is free to move (i.e., active listening). Unlike most previous works (e.g., McAnally and Martin 2014), here we focus on spontaneous head movements during sound emission. To this aim, we dissociated possible head movements from the instructed response, which always entailed hand pointing to the sound. In addressing this experimental question, we also developed a novel methodology based on sounds delivered from a free-field sound source (i.e., a loudspeaker) that was continuously tracked and aligned with a virtual reality environment. Specifically, starting from the pioneering work of Brungart and colleagues (1999), we took advantage of current VR technology to guide precisely a loudspeaker to predetermined head-centered coordinates in each trial. Critically, we coupled this head-centered placement of the loudspeaker with criteria for sound emission, based on the participant's head and eye posture measured in real time. This allowed us to test 3D sound localization abilities without physically restraining the participant's posture and to assess the contribution of spontaneous head movements to sound localization. We predict better 3D sound localization when participants make spontaneous head movements during listening.

Methods

Participants

Twenty normal-hearing participants (range 22–75 years, mean age = 46, SD = 18; 12 females; 18 right-handed) were recruited through advertisements (e-mail or flyer). All reported normal or corrected-to-normal vision and had no history of hearing deficits. Participants were informed that they would participate in a sound localization study that would require wearing a virtual reality HMD and that their task was to localize as accurately as possible a sound delivered in the space around them using a hand-pointing response. If they agreed to participate, they were asked to sign the informed consent documents. The study was approved by the Comité Ethique d'Evaluation de l'Inserm (IRB00003888), and conducted in accordance with the ethical standards of the 1964 Declaration of Helsinki.

Apparatus and stimuli

The apparatus was set up in two rooms: a test room and a control room. The control room hosted two desktop computers. The first computer ‘Control PC’, was an HP Z820 Workstation (Windows 7 Professional, Processor Intel(R) Xeon(R) CPU E5-2609 @ 2.40 GHz 2.40 GHz), equipped with a NVIDIA Quadro K5000 graphic card (DirectX 11.0). It controlled the entire sequence of events, stimulations, response collection and data saving through a custom-made script written in Unity (Version 5.5.1f1 32-bit, Unity Technologies, San Francisco, CA). The second computer, hereafter named the ‘Vicon PC’, was an HP Z230 Tower Workstation (Windows 7 Professional, Processor Intel(R) Core (TM) i7-4771 CPU @ 3.50 GHz 3.50 GHz). It controlled the Vicon motion capture system (Vicon Tracker 2.0.1 × 64, Vicon Motion Systems LTD, Oxford, UK) and ran a custom-made script written in Unity that performed stimulus visualization. The test room comprised Vicon cameras for motion capture, three rigid bodies for real-time object tracking, the head-mounted display (HMD) incorporating an eye-tracking system, one monitor for stimulus visualization, one loudspeaker, one keyboard and one remote control. Each of these pieces of equipment is described below, with details concerning the way they were interfaced with the control and Vicon PCs.

Vicon motion capture. The Vicon motion capture system comprised seven infrared cameras (Bonita 10: frame rate 250 fps, resolution 1024 × 1024, Vicon®, Oxford, UK) mounted on the walls of the testing room. The elevation (195–205 cm) and semicircular arrangement of the cameras allowed full kinematic tracking of a wide 3D space (height: 250 cm; width: 320 cm; depth: 170 cm). The space visible to the cameras was calibrated using the Vicon Active Wand tool (www.vicon.com/products/vicon-devices/calibration), which allows a multi-plane video calibration across the entire acquisition volume. Once calibrated, object-tracking spatial precision was < 1 mm (down to 0, 5 mm in a 4 × 4-m volume). We placed the HMD on the floor in a straight-ahead position to record a straight-ahead reference direction (taking into account HMD rotations). The cameras were connected to a multiport box in the testing room, which in turn was USB connected to the Vicon PC in the control room.

Rigid bodies. The Vicon system captured the position of three distinct rigid bodies (each mounted with 4 reflective 9 mm markers), with a sampling frequency of 100 Hz. The first rigid body (rigid body 1; radius 75 mm) was fixed on top of the loudspeaker and tracked its xyz coordinates in the calibrated space; the second rigid body (rigid body 2; radius 75 mm) was fixed on top of the HMD and tracked HMD and the head-center positions; the third rigid body

(rigid body 3; radius 75 mm) was used for head-size calibration and for collecting hand-pointing responses.

Head-mounted display. The HMD was an Oculus Rift Development Kit 2 system (DK2, Oculus VR®, Menlo Park, USA, screen OLED, resolution: 1920 × 1080 (960 × 1080 per eye), maximal refresh of 75 Hz, dimensions L × W × H: 1.3 × 14.7 × 7.1 inches, and a field of view equal to 106°) running with Oculus Runtime (Version 0.6; Facebook Technologies Ireland, Dublin, Ireland). The Oculus Rift DK2 incorporated an eye-tracking system (SensoriMotoric Instruments SMI, Berlin, Germany; (Kuk et al. 2014; Tyler et al. 2010); 60 Hz frequency and 0.5° spatial precision). In our setup, the HMD served two purposes: (1) it conveyed visual instructions to the participant; and (2) allowed continuous monitoring of the participant’s eye movements.

Loudspeaker. A loudspeaker (JBL GO portable, 68.3 × 82.7 × 30.8 mm, output power 3.0 W, frequency response 180 Hz–20 kHz, signal-to-noise ratio > 80 dB) was used to deliver all target sound stimuli. Target stimuli were amplitude-modulated broadband bursts lasting 3 s (the sound was modulated at 80%, amplitude varies between 0.2 and 1). The room was a quiet 3 × 6 m place with a reverberation time RT60 of 0.32 s, not treated for being anechoic or sound-proof, and the background noise measured at the beginning of the experiment was 33.7 dB SPL. The sound duration of 3 s allows the subject to increase the possibility of moving the head (Thurlow and Mergener 1970) that is beneficial for acoustic dynamic cues (i.e., change in binaural cues ILD and ITD; review by Middlebrooks and Green 1991; Pollack and Rose 1967).

A keyboard, a remote control (Targus®, Laser Presentation Remote) and a monitor (DELL 19’’ 5:4, resolution 1280 × 1024), completed the equipment in the testing room. All devices were connected to the control PC, except the stimulus visualization monitor which streamed a copy of the screen of the VICON PC inside the testing room. The function of these four pieces of equipment is described in detail in ‘‘Procedure’’.

Procedure

Before starting the experiment, participants were introduced to the task and to the VR equipment using a visual information sheet. Participants were told that sounds could be delivered anywhere in the 3D space around them at a maximum radius corresponding to their arm-reaching distance, and they would perform the sound localization task under two conditions: ‘static listening’ in which they would have to keep their head still in the initial position throughout sound presentation, and ‘active listening’ in which they could actively search for the sound during its presentation, by freely moving their head. They were also instructed to only pay attention to the sounds, as any other noise in the

room could be misleading. In a control experiment ($N=6$), we ensured that normal-hearing participants were incapable of locating the loudspeaker based on unintentional cues (e.g., experimenter's displacement), unless the sound was actually emitted. In both conditions, they were free to move their head and body as soon as the sound finished.

The experiment began with eye and head-center calibrations: (i) eye calibration was performed using a five-point calibration grid (smart recorder of SMI Eye tracking software) which permitted control of the 3D cyclopean eye position and direction; (ii) head-center calibration was performed by collecting the 3D position of the two ears (using rigid body 3), averaging these positions to obtain the 3D head-center position. The head-center position served as the origin for the polar coordinate system that included loudspeaker, hand, head and cyclopean gaze positions. In this way, even though participants sat without a chin rest, we could carefully control the position of each sound source with respect to their head position. Twelve predetermined positions were used throughout the experiment, resulting from the combination of four different azimuths (-30° , 30° , -150° or 150°), three different depths (35 cm, 55 cm or 75 cm) and a single elevation (0° , i.e., ear level). Despite that the elevation remained constant, participants were left unaware of the 3D sound position and had no visual cue that constrained the possible sound origin. Moreover, their response could vary in all dimensions, thus rendering their response elevation also relevant to this experiment. For these reasons, we studied the participants' responses in 3D.

In each trial, two sets of instructions, generated in real time by the computer, informed the experimenter where to

position the loudspeaker in the 3D space surrounding the participant. The stimulus visualization monitor displayed in real time and in 2D (azimuth and depth) the actual position of the loudspeaker and its desired position for the upcoming trial. The precise elevation positioning of the speaker was communicated to the experimenter via an echo radar sound delivered by an in-ear headphone (non-audible by the participant). This allowed the experimenter to rapidly place the loudspeaker in a predetermined position with a margin of 3D error of 2.5 cm. Sound could only be delivered when three criteria were concurrently met: (1) the loudspeaker was in the correct 3D position; (2) the participant's head was facing straight ahead; (3) the participant's eyes were directed straight ahead. Participants actively complied with criteria 2 and 3 by aiming their head and eyes to align two crosshairs displayed in the HMD (cf Fig. 1). At the end of sound emission, the experimenter removed rapidly the loudspeaker out of the emission area so that participants did not collide with it when pointing with their hand-held rigid body. After trial completion, no feedback on performance was ever provided.

The experimental session was organized into four successive blocks. Listening conditions (static or active) changed between blocks of trials (half of the participants followed an active–static–static–active sequence, whereas the other half followed a static–active–active–static sequence). Each block consisted of 48 randomized trials, resulting in a total of 192 trials (8 trials \times position \times listening condition).

Each participant completed the whole experiment in approximately 40 min. Each trial lasted 10–15 s, with the speaker-positioning phase lasting 3–5 s, depending on the predetermined position. Multiple aspects contributed to

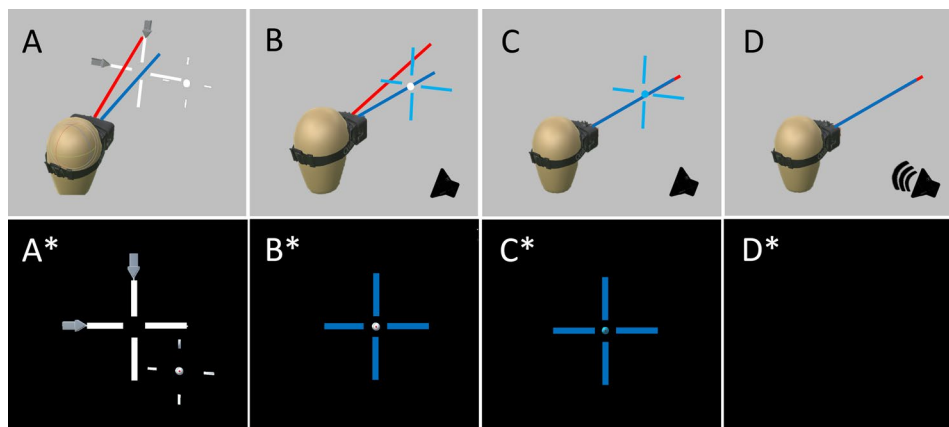


Fig. 1 Pre-stimulation alignment of head and eyes. **A** At the beginning of each trial, the participant was free to move their head (symbolized head direction: blue line) and eyes (symbolized cyclopean eye direction: red line). **A*** In the HMD display: a bold white cross indicates the actual position of the head, and a thin white cross with a central ball indicates the desired position of the head and eyes, respectively. Two gray arrows flank the bold white cross, showing the participant in which direction to move their head to achieve

the desired initial position. **B*** When the desired head position was achieved, the bold cross turned blue. **C*** When the desired eye-gaze position was reached, the central ball turned blue. **D** When all criteria were met (head and eye position straight ahead and speaker within a sphere around the predetermined position), all visual stimulations were removed, **D*** the scene became dark and the sound was delivered

trial duration: events before sound delivery (the participant actively moved head and eyes to the desired initial posture), the experimenter manually bringing the loudspeaker to the predetermined position, the sound delivery itself (3 s), and the participant's full body motor response (mean reaction time for the hand responses was $3.65 \text{ s} \pm 0.48$). No time restrictions on the response were imposed.

Data processing

The position of all tracked elements recorded in Vicon reference frame (loudspeaker, head center and direction, hand) was re-computed in head-center reference frame. Kinematic analyses of head and hand were conducted and inspected for each trial using an in-house software running on MATLAB R2013a, which allows to filter, identify and scrutinize head and hand movements in the kinematic trace (for previous description of this procedure, see (Gaveau et al. 2008)). Specifically, head and hand position signals were first filtered (50 Hz cutoff frequency, finite impulse response filter FIR) and velocities were computed from the filtered position signal using a two-point central difference derivative algorithm (Bahill and McDonald 1983). To obtain the spatio-temporal profile of head behavior and to extract relevant parameters for subsequent analyses, the beginning and the end of all movements were automatically detected using a velocity threshold procedure (80 mm/s). The number of head movements during sound emission corresponds to the number of times the speed threshold was exceeded. This procedure was also applied to hand movements. Finally, we extracted relevant kinematic parameters for subsequent analyses.

To evaluate overall sound localization performance in 3D (i.e., across azimuth, elevation and depth, and irrespective of sound position), we computed a cumulative index, called the 3D error, for each listening condition. To do this, we adapted the error calculation proposed by Rakerd and Hartmann (Rakerd and Hartmann 1986), which combines into a single measure the absolute constant error (referred to as \bar{C}) and the random error as follows: $\sqrt{\bar{C}^2 + \bar{s}^2}$. To obtain overall error in 3D (3D error), we calculated for each trial i the norm of the vector C_i . This is the distance in 3D space between the participant's response (i.e., the coordinates of the rigid body held in the participant's hand, x_h, y_h, z_h) and the speaker location at the moment sound was delivered (i.e., the coordinates of the rigid body mounted on the speaker, x_s, y_s, z_s). All values of C_i extracted for each participant were then averaged irrespective of sound position. The random error for each participant was computed as the standard deviation of the responses at each sound position, averaged across all sounds.

All statistical analyses and data visualizations were performed using R in the R-studio environment. We used 'afex' package for ANOVA designs and the 'lsmeans' package for estimating marginal means and run comparisons. We also used 'dplyr' and 'Rmisc' for preprocessing and 'ggplot2', 'ggpubr' and 'cowplot' for data visualization. Unless otherwise indicated, means \pm standard errors are reported in the text. We planned ANOVAs or t tests, and the Greenhouse–Geisser sphericity correction was applied to analyses of variance, when appropriate. The Holm–Bonferroni method was applied to adjust of the p values for multiple comparisons (Midway et al. 2020). Data for all statistics, complete analysis pipeline and scripts are available in the supplementary materials and at the following link: osf.io/8fapq.

Results

Positioning the loudspeaker at predetermined locations

In the absence of physical constraints on participant posture, the head returns to slightly different initial positions at the beginning of each trial. This poses a potential problem of reproducibility of sound source positioning across trials and participants. Participants were instructed at the beginning of each trial to align their head and eyes with respect to straight-ahead reference, using visual cues available in the HMD. Online head kinematics tracking allowed sound to be delivery only when the required eye and head posture criteria were matched. In addition, and most importantly, on each experimental trial, we guided the loudspeaker to a predetermined location in the environment defined in head-centered coordinates.

We started by assessing if the predetermined loudspeaker locations remained constant across trials and participants. We tested 12 sound locations all around the participant's head. Figure 2A and C shows initial head position and predetermined locations for all participants across all trials, in the reference frame of the motion-tracking system (i.e., Vicon coordinates). A substantial variability is observed, which reduces dramatically, however, when all positions are referenced to the center of the head, i.e., they are converted to a head-centered reference frame (Fig. 2B and D). Figure 2E summarizes the effect of head-centered referencing by comparing mean changes in standard deviation across participants in motion-tracking vs. head-centered coordinates (x : $1.08 \pm 0.24 \text{ cm}$ vs. $0.10 \pm 0.02 \text{ cm}$; y : $0.63 \pm 0.14 \text{ cm}$ vs. $0.10 \text{ cm} \pm 0.0$; z : $1.89 \pm 0.42 \text{ cm}$ vs. $0.15 \pm 0.03 \text{ cm}$);

Next, we tested if the variability of loudspeaker actual location around the predetermined position was within the established tolerance (i.e., a sphere with a 2.5 cm radius

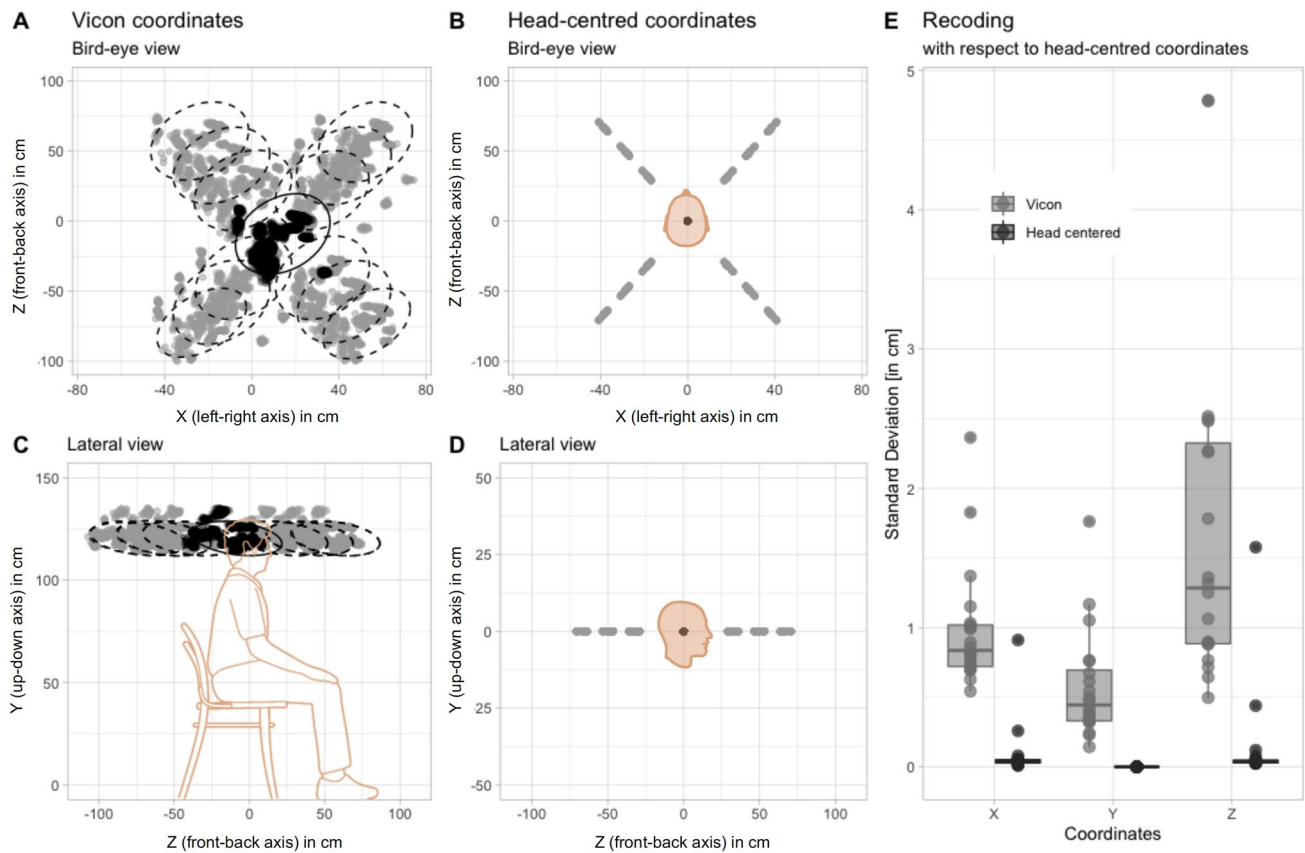


Fig. 2 Normalization to head-centered coordinates. Bird's-eye and lateral views of initial head position (in black) and 12 predetermined locations (in gray) for all participants (192 trials each), in VICON reference frame (**A–C**) and head-centered coordinates (**B–D**). Variabil-

ity of predetermined locations averaged across 12 positions for each participant in the two reference frames as a function of coordinates x , y , z (**E**)

around the predetermined position). Indeed, this was the case. Figure 3 shows all 192 stimulations for all participants when the 12 predetermined locations were re-aligned to a single coordinate, centered on the origin of the axes (mean differences with SE between predetermined and actual location: $x = 0.87 \pm 0.01$ cm, $y = 1.15 \pm 0.02$ cm, $z = 0.93 \pm 0.01$ cm; error in 3d = 1.98 ± 0.01 cm).

During sound emission, all visual references for the straight-ahead head posture in the HMD were removed. Depending on the experimental conditions, participants were either instructed to keep their head motionless (static listening condition) or they were free to move their head and explore space (active listening condition). Irrespective of the listening conditions, participants were free to move their body and indicate the 3D sound location with their hand as soon as the sound ended. Participant compliance with the instructions was examined off-line and trials in which instructions were not followed (e.g., anticipatory hand or head movements during sound delivery in static condition) were excluded from further analyses (static listening: 6.1%,

SD = 8; active listening: 6.6%, SD = 7). The main reason for trial rejection was anticipatory hand responses.

Thus, our apparatus based on head-centered positioning of a loudspeaker in the environment allowed high precision 3D control of sound position with a minimal constraint on participant posture.

Sound localization during static and dynamic listening

In this study, we wanted to examine whether active listening (i.e., free head movements during sound presentation) changed spatial hearing performance (Fig. 4 depicts hand-pointing responses in each separate dimension).

The effect of the listening condition was evaluated for the three dimensions (azimuth, elevation and depth). Figure 5 presents changes in absolute and variable localization errors as a function of listening condition; the absolute error reflects the accuracy of the performance and the variable error the precision.

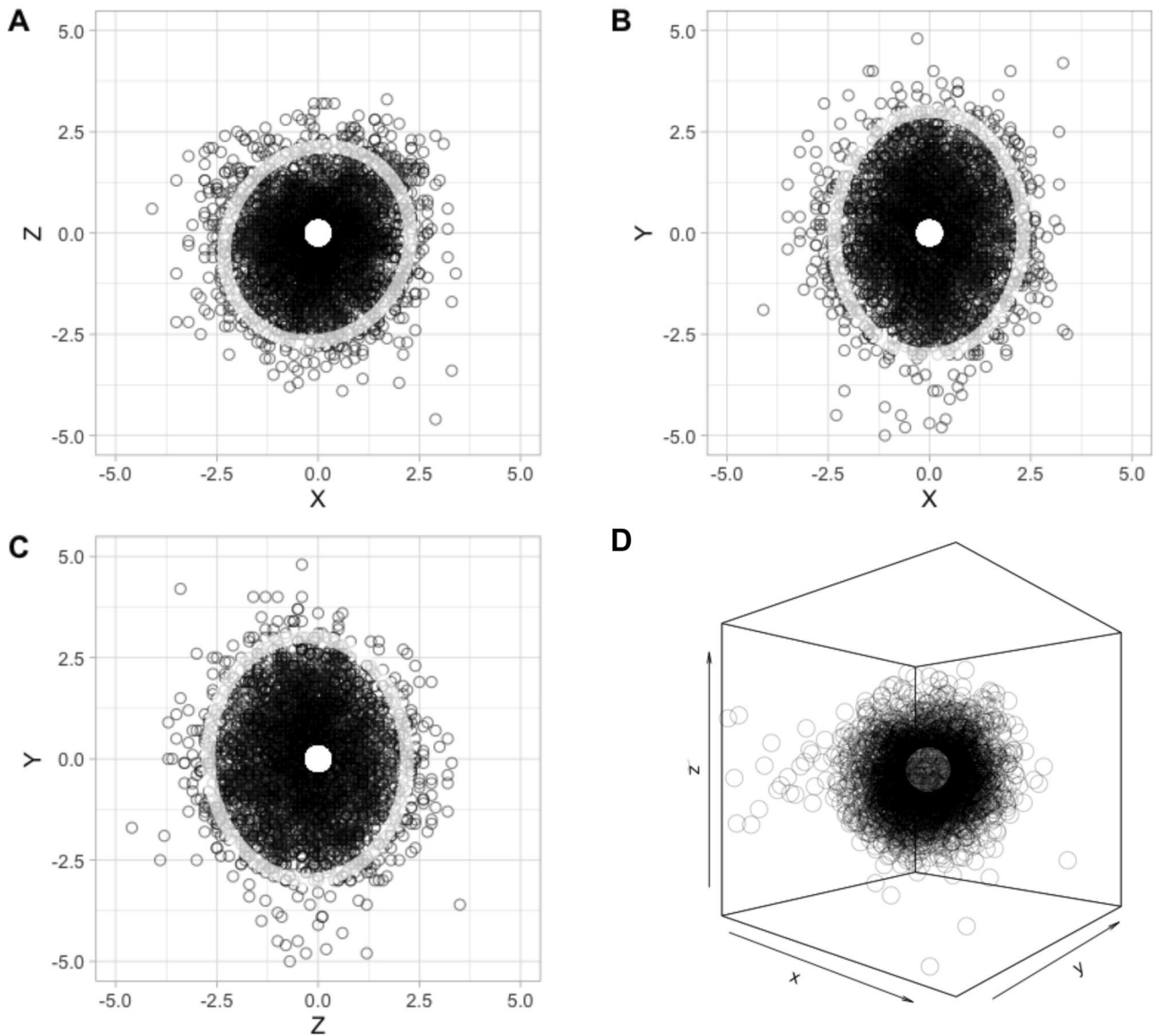


Fig. 3 Actual speaker location with respect to predetermined location, in centimeters. Stimulations delivered to all participants, when the 12 predetermined locations were re-aligned to a single coordinate,

centered on the origin of the axes. **A** top view; **B** front view; **C** lateral view; **D** 3D rendering. Ellipses in the 2D panels represent 95% confidence intervals of the distributions

To study azimuth errors along the antero-posterior axis as a function of listening condition, we entered absolute and variable errors in separate ANOVAs with antero-posterior sector (front and back) and listening condition (static and active) as within-participant factors. The analysis on azimuth absolute errors revealed an effect of listening condition ($F(1,19)=17.58, p < 0.001, \eta_g^2 = 0.04$). The error reduction occurred in front space for active ($15.9 \pm 3.8^\circ$) compared to static listening ($26.1 \pm 4.0^\circ, p = 0.0005$), whereas no such change occurred in back space (static: $17.2 \pm 2.2^\circ$; active: $16.5 \pm 2.3^\circ; p = 0.43$; Fig. 5A). The expected two-way interaction between the antero-posterior sector (front and back) and listening condition (static and active) was

significant ($F(1,19)=12.93, p = 0.002, \eta_g^2 = 0.03$). Azimuth variable error also benefited from active listening and was thus reduced in active (11.7 ± 1.4) compared to static listening (14.3 ± 1.3) ($F(1,19)=4.55, p = 0.046, \eta_g^2 = 0.03$), and in front space (10.4 ± 1.4) compared to back space (15.6 ± 1.7) ($F(1,19)=6.06, p = 0.024, \eta_g^2 = 0.10$), without interaction between these two factors (listening condition and antero-posterior sector, $F(1,19)=0.20, p = 0.660, \eta_g^2 < 0.01$) (Fig. 5B).

A convergent result emerged for elevation. When absolute and variable elevation errors were entered into an ANOVA similar to the one described above. Elevation absolute errors

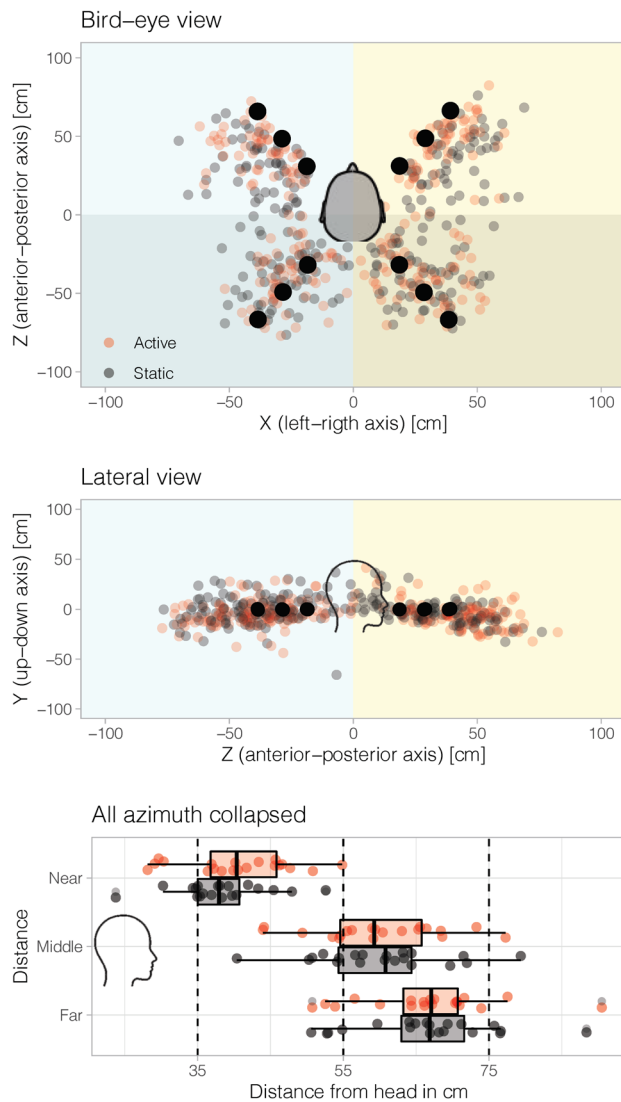


Fig. 4 Behavioral pointing and effects of static and active listening on sound localization. **A** Bird's-eye view of all target positions (black dots) and hand-pointing responses (smaller gray and red circles) for each participant, averaged across trials in a quadrant (i.e., front-left, front-right, back-left, back-right) irrespective of sound distance. Color code is a function of listening condition (black: static listening; red: active listening). **B** Lateral view of all target positions and responses. Responses for each participant are averaged across (left or right) and distance (near, middle or far). **C** Lateral view of responses in depth (black box plot: static listening; red box plot: active listening). All participants were included

in front space were reduced for active ($21.2 \pm 5.1^\circ$) compared to static listening ($31.3 \pm 6.0^\circ$, $p = 0.02$), whereas no such change occurred in back space (static: $19.0 \pm 3.1^\circ$; active: $19.3 \pm 3.0^\circ$; $p = 0.7$; Fig. 5C). The two-way interaction between antero-posterior sector and listening condition reached significance for absolute errors ($F(1,19) = 6.54$, $p = 0.019$, $n_g^2 = 0.02$). Again, elevation variable errors were reduced in active (16.2 ± 2.6) compared to static listening

(22.6 ± 3.2), irrespective of whether stimuli were in front or back space (main effect of listening condition for the antero-posterior sector, $F(1,19) = 4.59$, $p = 0.045$, $n_g^2 = 0.03$, Fig. 5D), with an interaction between listening condition and antero-posterior sector, ($F(1, 19) = 6.06$, $p = 0.024$, $n_g^2 = 0.04$).

By contrast, active listening did not affect depth estimation. When absolute errors in depth were entered into an ANOVA with distance (near, middle, far), antero-posterior sector (front and back) and listening condition (static and active) as within-participant factors, no significant main effect or interaction involving listening conditions emerged (all $F_s < 2.97$, Fig. 5E). Likewise, no main effect or interaction involving listening condition emerged for variable errors in depth (Fig. 5F), and we noticed only a main effect of the antero-posterior sector ($F(1,19) = 6.69$, $p = 0.018$, $n_g^2 = 0.02$) without interaction with the listening condition. Participants were able to perceive three distinct depth positions (hand distance from the head for near, middle and far sound position, 39.4 cm, 60.2 cm and 66.4 cm respectively; $F(1.19, 22.63) = 176.13$, $p < 0.001$, $n_g^2 = 0.62$).

Taken together, these results show that active listening (free and spontaneous head movements) improved sound localization (accuracy and precision) in azimuth and elevation.

Head movements during active listening

Continuous kinematic tracking of the HMD allowed detailed investigation of head movements during sound emission in the active listening condition (recall that participants were only told that head movements were possible, they were not explicitly instructed to move their head upon sound presentation or to orient to the sound with their head. Likewise, during the response phase it was made clear that only hand-pointing was relevant for measuring performance). Even though the active listening condition allowed free head movements during sounds, not all participants moved their head. As visible in Fig. 6A, the distribution of percent head movements revealed two outliers (i.e., points beyond 1.5 of the interquartile range, IQR): one participant who never moved his head and another who moved only in 6 out of 96 trials (6.3%). These two outliers were removed from all subsequent analyses on head movements.

The mean number of head movements during sound was 1.22 ± 0.04 , with an average onset at 1077 ± 73 ms (head movements beyond 3000 ms, i.e., after sound emission, were removed from this analysis; Fig. 6B). Head movements occurred on 73.6% of trials on average ($SD = 24.0\%$), both for targets in front and back space (74.7% and 72.5% of trials, respectively). For targets in front space, they were mostly directed to the target. On average, for sounds at $+30^\circ$,

the first head movement was directed to $23.6 \pm 1.8^\circ$, whereas for sounds at -30° it was directed to $-27.0 \pm 2.5^\circ$ (Fig. 6C). For targets in back space head movements were distributed within the entire stimulated hemisphere (Fig. 6D). They were either directed to the front quadrant on the same side as the target (e.g., left front quadrant for targets at -150°) or aimed directly at the back target (in this case involving a trunk movement). On average, for sounds at $+150^\circ$ the first head movement was directed to $107.5 \pm 6.3^\circ$, whereas for sounds located at -150° it was directed to -118.0 ± 6.6 .

3D error

As a final step, we quantified overall sound localization performance in 3D using the 3D error (see “Data processing” for details). Figure 7 shows change in 3D error in the two listening conditions. Considering all participants and trials, we ran an ANOVA with antero-posterior sector (front and back) and listening condition (static and active) as within-participant factors in 3D hearing performance. The improvement in sound localization in active ($28.0 \text{ cm} \pm 2.3$) compared to static listening ($31.2 \text{ cm} \pm 2.1$) emerged as marginally significant ($F(1,19) = 3.97$, $p = 0.061$, $\eta_g^2 = 0.02$), but was significant in front space when taking into account listening posture (active = $27.6 \pm 3.4 \text{ cm}$, static = $34.5 \pm 3.7 \text{ cm}$; $F(1,19) = 7.67$, $p = 0.012$, $\eta_g^2 = 0.02$). When the difference between active and static posture was studied as a function of the mean number of head movements, a positive correlation emerged ($r = 0.37$, $p = 0.023$, Kendall’s rank correlation tau). The higher the proportion of trials with head movements during sound emission, the greater was the performance improvement in active compared to static listening. A convergent correlation emerged also between the mean number of head movements and 3D error ($r = 0.34$, $p = 0.038$).

Discussion

In the present study, we examined to what extent spontaneous head movements improve sound localization in 3D—azimuth, elevation, and depth—by comparing static vs. active listening postures. To this aim, we developed a novel approach to sound localization based on sounds delivered in the environment brought into alignment with a VR system. Our system proved effective for the delivery of sounds at predetermined and repeatable positions in 3D space, without imposing a physically constrained posture, which required minimal training. In addition, it allowed measuring participant behavior (hand, head and eye position) in real time.

Active listening improves 3D sound localization

In the static listening posture, normal-hearing participants reliably discriminated sound sources in azimuth, elevation and distance. Absolute errors were 21.6° , 25.2° and 10.5 cm on average, respectively. Along the horizontal dimension, performance in azimuth was worse for front (26.1°) compared to back targets (17.2°), with a bias to point to more eccentric positions for frontal sources. Likewise, for elevation, inaccuracies were more evident for frontal sources compared to rear ones. It is worth mentioning that the overall angular error measured in our study is numerically greater than those obtained by other studies. For comparable elevation position, Brungart and colleagues (1999) obtained a mean angular error of 16.3° , and Wightman and Kistler (1999) a mean error of 21.1° . Of particular interest are errors in depth, which are typically much less investigated compared to those in azimuth and elevation. In the present work, participants succeeded in distinguishing the three sound sources in depth, but underestimated far targets compared to near ones. This is a well-established pattern when studying depth perception in spatial hearing (Brungart et al. 1999; Kearney et al. 2010; Kolarik et al. 2016; Middlebrooks and Green, 1991; Parseihian et al. 2014; Zahorik 2002; Zahorik et al. 2005; Zahorik and Wightman 2001). Sound distance cues, such as interaural level difference (ILD) and direct-to-reverberant energy ratio in reverberant environments are important for distance localization in the near-head acoustic field (Kopčo and Shinn-Cunningham 2011; Kolarik 2016; Brungart et al. 1999). In the near field specifically, sound distance perception also relies on low-frequency ILD (Brungart et al. 1999; Middlebrooks and Green 1991) and the closer the sound is to the listener, the more does the ILD contain low frequencies, and therefore the more does the distance accuracy increase (Brungart et al. 1999; Kolarik et al. 2016). Some other factors may increase this degree of precision, such as the presence of an echoic room (see review from Kolarik et al. 2016) and the lateral positioning of sound sources (Kopčo and Shinn-Cunningham 2011). Thus, our participants may have benefited from several effective cues to distinguish the different distances.

While performance of normal-hearing adults may appear relatively inaccurate in 3D (overall), it is important to consider three aspects of our paradigm that may have contributed to this outcome. First, at odds with most previous studies, here participants’ 3D errors combine uncertainty across all three dimensions (i.e., azimuth, elevation and distance, all unknown to participants). Second, participants had to resolve sound position in a totally dark 3D space. Concurrent estimation of the three space dimensions may have been particularly difficult in the absence of visual references, especially in front space where visual cues typically contribute to perceived sound position (Alais and Burr 2004).

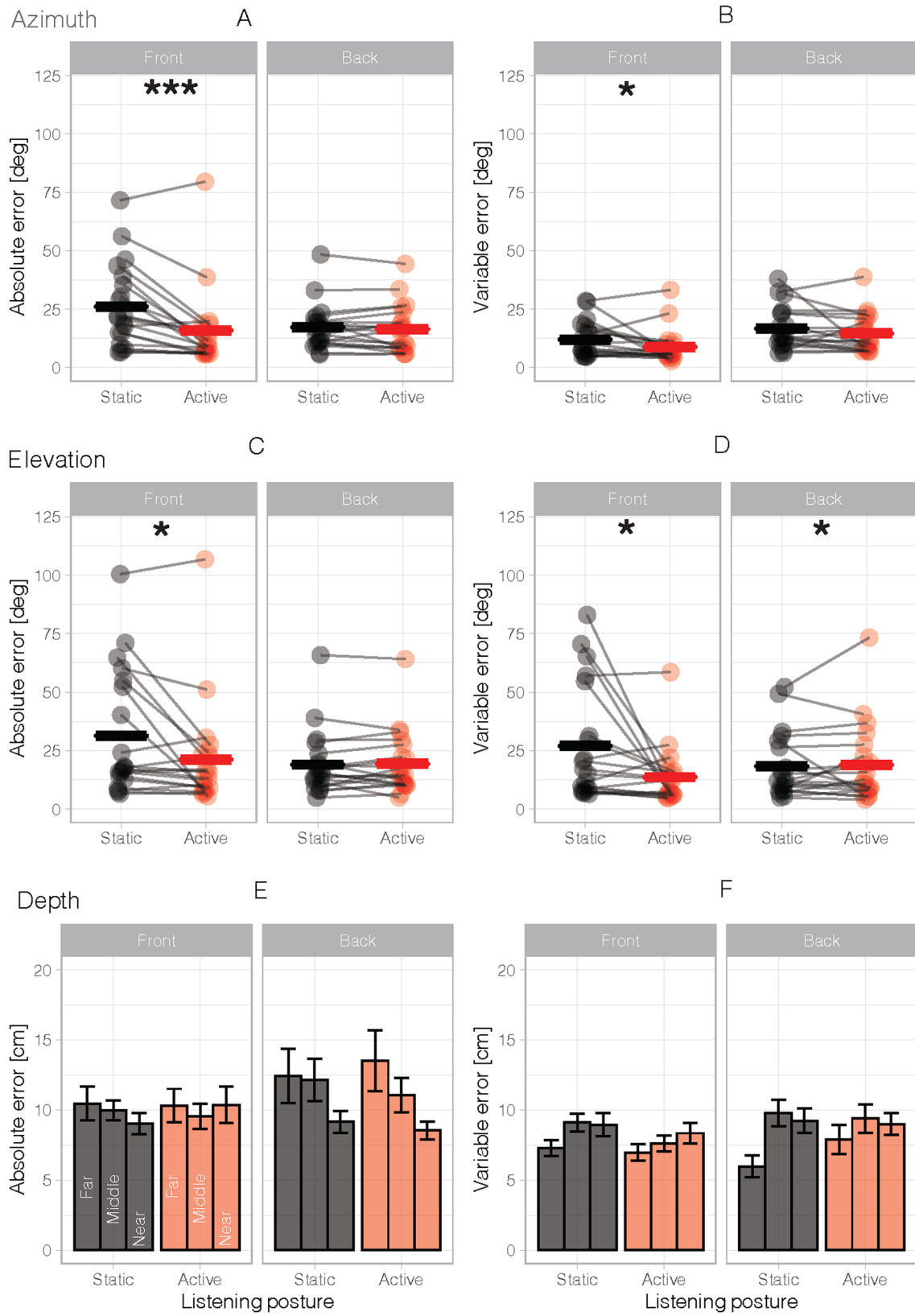


Fig. 5 Effects of static and active listening on sound localization performances. In azimuth dimension, for the hand absolute error (A) and the hand variable error (B) for each participant as a function of listening condition and antero-posterior position of target sounds. In elevation dimension, for the hand absolute error (C) and the hand variable error (D) for each participant. In depth dimension, for the hand absolute error (E) and the hand variable error (F) for each participant and for the three distance (near, middle and far sound position). Bold horizontal lines indicate the mean for all participants. Asterisks indicate significant differences (* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$)

Third, wearing an HMD may have altered sound localization cues. The HMD, which acted as a physical obstacle to sound diffusing from front near (Ahrens et al. 2019; Genovese et al. 2018; Gupta et al. 2018). For instance, Ahrens and colleagues (2019) documented larger azimuthal errors for lateral sound sources delivered in front space, when participants used the HMD compared to when they localized sounds without it).

In the active listening posture, no specific head-movement strategy was imposed to participants who were free to move their head or not during sound emission. As a matter of fact, most participants engaged in active listening (only two participants moved their head only in a few trials or not at all), which improved 3D sound localization primarily by ameliorating accuracy and variability of responses in azimuth and elevation. The more participants made spontaneous head movements, the better was their 3D sound localization performance. The benefit of active listening emerged selectively in the front space. As this is the portion of space in which participants were less accurate and less precise in the static condition, they may have leveraged a greater margin for their improvement.

Wallach (1940) was the first to report benefits of head movements for spatial hearing. Other works showed that head movements help normal-hearing listeners to distinguish between sounds coming from front and rear positions (Dunai et al. 2011; Mueller et al. 2014; Perrett and Noble 1997b; Wightman and Kistler 1999). A more recent study (Kim et al. 2013) compared azimuthal sound localization under conditions of active head movements, passive head movements, and body movements with the head fixed. The results of Kim and colleagues (2013) suggest that vestibular information associated with head movements may be both necessary and sufficient to improve sound localization. In contrast, proprioceptive information alone (available in the body movement with the head-fixed condition) does not improve localization. The impact of head movements on sound localization abilities has been recently documented also in adults and children with hearing loss who use cochlear implants. Pastore and colleagues (Pastore et al. 2018) showed that front-back confusions diminish in bilateral cochlear implant users asked to rotate their head within a range of approximately $\pm 30^\circ$,

compared to a static head posture. Similarly, Coudert and colleagues (2022) found that sound localization in 3D improves when children with bilateral cochlear implants are allowed to spontaneously move their head (as here), compared to a static head posture.

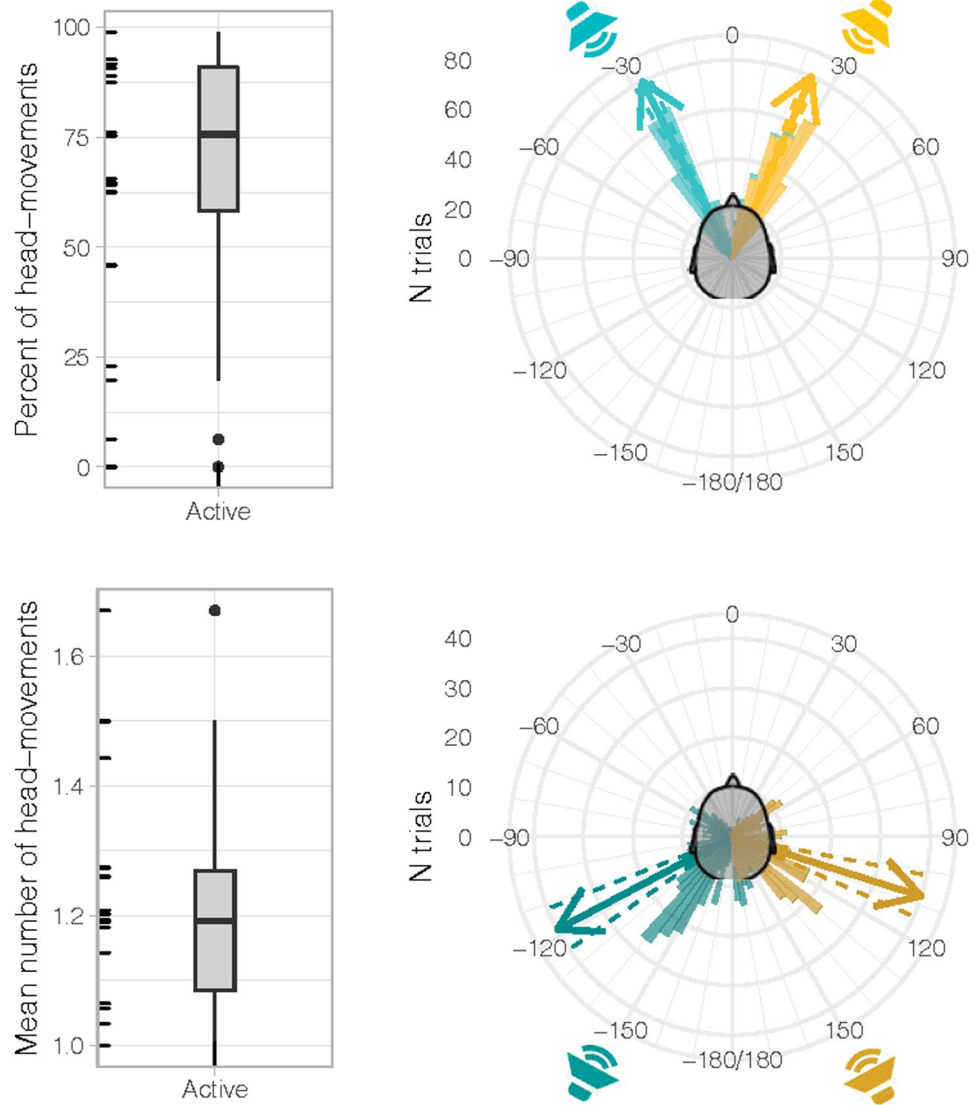
For depth perception, active listening did not change performance accuracy or variability, in front or back space. Wearing an HMD did not alter the low-frequency ILD component of sound (Ahrens et al. 2019), the position of sound sources (close to the head and lateral) and reverberation cues likely yielded enough localization cues to solve the distance discrimination task without help of head motion.

The functional mechanisms that underlie sound localization improvements by head movements remains to be ascertained. On the one hand, the active listening benefit could result from richer auditory cues at the ears (i.e., at the auditory processing periphery). Head rotations, either horizontal (left–right) or vertical (head tilted up or down) inevitably produce dynamic acoustics cues (change in binaural cues) that, in turn, could facilitate sound localization (Lambert 1974; Perrett and Noble 1997b). In addition, head movements cause the interplay of sensory and motor signals (proprioception/efference copy and vestibular), which might be better integrated with dynamic binaural cues to solve the sound localization task. On the other hand, intentional head movements make sound localization a predictive process. Participants could benefit from ‘hypothesis verification through action’: predictions about sound location are constantly updated based on the incoming error signals that result from head movements, resulting in an interactive cycle that generates a more veridical model of the auditory environment (see Yon et al. 2019) for related examples from the visual modality). Notably, this central mechanism could remain valid also when peripheral auditory information is less accurate (as in the case of people with hearing loss, or using hearing aids or cochlear implants).

Pursuing active sound localization in 3D: a methodological challenge

Pursuing active approaches to sound localization in 3D space may be a shared objective when aiming to measure and capture the complexity of this fundamental behavior in real life. Yet, it remains a methodological challenge to achieve this aim in research and clinical settings. Below, we briefly discuss alternative approaches to this problem and summarize the advantages and limitations of the novel methodology we have introduced.

Fig. 6 Head movements during sound emission in the active listening condition. **A** Box plot of percentage head movements. Note that two participants were identified as outliers (i.e., they fell outside the $1.5 \times$ interquartile range), made almost no head movements during the active listening condition and were thus excluded from subsequent analyses. **B** Box plot of mean number of head movements once outliers were removed. **C**, **D** Polar histogram showing the distribution of head-movement responses for targets in front (**C**) and back (**D**) space. Arrows indicate mean head-movement direction, dashed lines indicate ± 1 SE



Multiple loudspeakers in the physical environment

One approach to the study of sound localization in 3D space is with multiple loudspeakers placed at fixed locations in the physical environment (e.g., Ahrens et al. 2019; Bahu et al. 2016). With these experimental setups, it is practically mandatory that the participant keeps a fixed posture at the beginning of each trial, because this is the only way to ensure replicable positions of target sounds with respect to the ears (Bahu et al. 2016; Brungart et al. 1999; Oldfield and Parker 1984; Seeber et al. 2004; Wightman and Kistler 1992). However, this often implies that participants also keep a fixed posture throughout the trial. In addition, because sound sources are physically present, these approaches face the problem of controlling the contribution of visual cues to sound localization. Participants are sometimes blindfolded from the moment they enter the experimental room (Ahrens

et al. 2019; Bahu et al. 2016), or they are instructed to close their eyes at specific moments during the task (Brungart et al. 1999), or face speakers hidden behind a fabric panel of some sort (Rabini et al. 2019). Note that the first two solutions pose the problem that they prevent tracking of eye position. In natural conditions, eye-orienting responses permit encoding of sound position in retinocentric coordinates (Bulkin and Groh 2006; Pavani et al. 2008), and it has been documented that static and dynamic eye position influence sound localization (see Groh and Sparks 1992; Lewald and Ehrenstein 1996; Pavani et al. 2008). Our approach allows to control for initial eye position, while continuously monitoring the listener's head position rather than asking participants to close their eyes and stay still in a predetermined position, and allows control of visual cues.

Brungart and colleagues (1999) were the first to have the intuition of tracking the kinematic of a single loudspeaker,

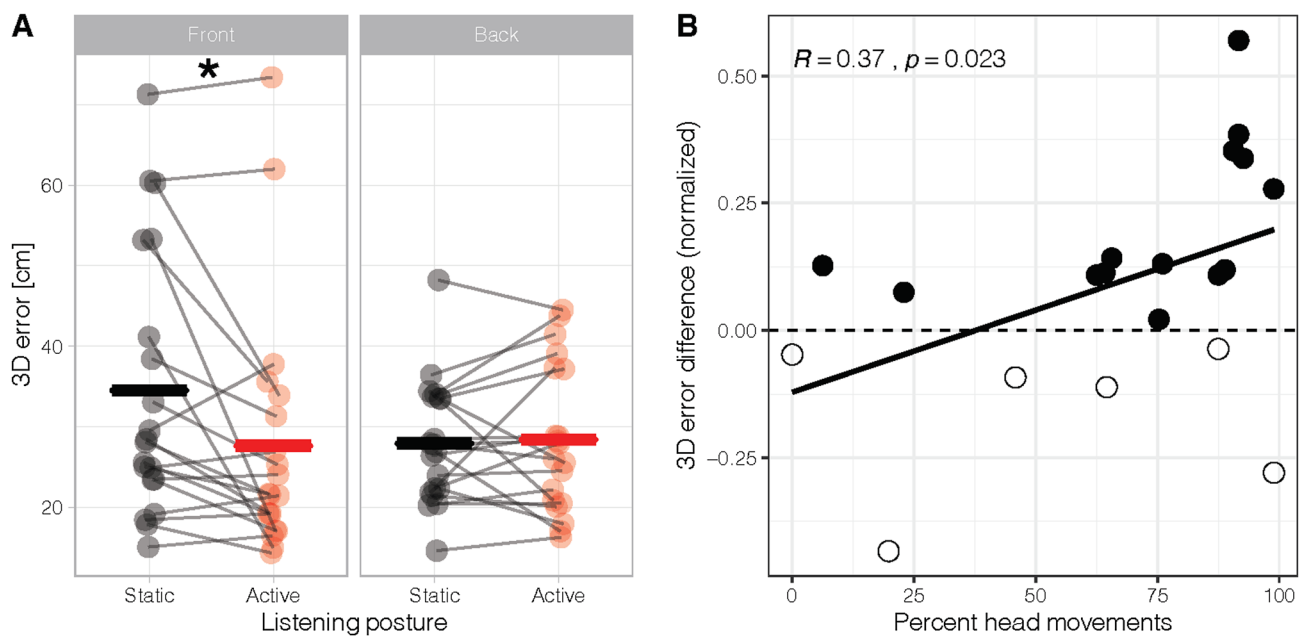


Fig. 7. 3D error. Scatterplot of the difference in 3D errors between active and static listening conditions (normalized difference based on static listening performance), as a function of percent head movements. Filled circles indicate participants who improved in active

compared to static listening; empty circles indicate participants whose sound localization performance decreased in active compared to static listening

displaced trial-by-trial at different locations around the listener head, to study spatial hearing in 3D. Although the speaker position in each trial was somewhat approximate (the experimenter received verbal instructions about the predetermined speaker location through headphones), its actual location was recorded at the end of each trial using a position-sensing system mounted on the chin rest. Using this sound method, Brungart and colleagues (1999) succeeded in placing sound sources in 3D space. While innovative, this experimental setup was complex and time consuming. Participants had to familiarize with the procedure before data collection. Moreover, the method had intrinsic limitations. First, sound source positions were variable among participants because the speaker’s coordinates were interpreted by the experimenter in each trial using a number-to-coordinate mapping. Second, participants had to close their eyes during sound positioning, thus limiting most oculomotor information that could have enhanced sound localization abilities (Maddox et al. 2014). Third—and most important—to ensure reproducibility of sound source coordinates across trials and participants, the listener’s head was immobilized with a chin rest throughout the experiment.

Virtual sound approach

When studying the impact of head movements on sound localization, one current approach is to exploit auditory virtual reality. Using head-related transfer function (HRTF),

it is possible to present sounds through headphones that appear to originate from different positions in 3D space. Virtual sounds prove useful for generating static and moving auditory sources (Dunai et al. 2011) from multiple positions around the listener. In addition, they have been exploited for studying the contribution of visual information to spatial hearing (Majdak et al. 2011). Nonetheless, it remains a challenge to track and update the 3D virtual position of sounds in real time as a function of head movements (for review, see (Lida 2019)). Furthermore, the transfer to more clinical settings remains limited because reproducing reliable virtual sounds with HRTF can be particularly difficult and time consuming when participants use hearing aids or cochlear implants (Majdak et al. 2011).

Real sound delivered in a virtual reality environment, our approach

Here, we built from the pioneering approach of Brungart and colleagues (1999), overcoming each of the previous limitations. Our approach makes it possible to position the sound source at any controlled 3D position around the subject (without given access to localization cues). The loudspeaker’s xyz coordinates were controlled online by the computer and used to place the sound source in a predetermined position in space. When referenced to the center of the head, computer-controlled placement of the speaker led to an error below 1 cm across all target positions, all participants and all

recording sessions. This proves the efficacy of our speaker positioning method, even without physical head restraints.

Our approach allows studies without time-consuming training for experimenters and participants. Notably, all experimenters achieved accurate and fast speaker 3D positioning with only a few minutes training (< 5 min). Most importantly, this approach of pointing to a sound source in its near space required no procedural training to perform the task, and less than 7% of trials were rejected for non-compliance with instructions. Indeed, in the study by Bahu and colleagues (2016), despite training the participants to familiarize with the pointing method to the sound sources, participants had difficulty performing the motor task, especially for the rear sound sources. Our ‘simple’ approach is particularly relevant for the eventual aim of applying this same methodology to developmental and clinical populations (for example, this pointing method to sound source has been used without any difficulty with children, (Coudert et al. 2022).

The use of the HMD was motivated by the fact that we wanted to control visual cues from the environment, control the position of the eyes and the head (it is also used for head-center reference frame). It has proven to be a very good way to identify the active listening strategy of participants. The direction of the head movement reveals the portion of space captured when sound is perceived: our participants faced sounds in front of them, or moved their head in back space for rear sounds. This sound space perception does not need further explicit response from the subject (e.g., hand pointing, verbal response). By recording the head direction as a tool for spatial hearing abilities, subject sound localization performance could be easily explored, and it is crucial when dealing with sound spatial abilities in case of hearing impairment, or dealing with sound targets in the far space. The HMD is an object placed on the face, which constitutes a physical obstacle to sound diffusion close to the ears, modifies the HRTF of the head and impacts sound localization for front sound sources. But in active listening condition, the HMD’s effect is attenuated and it no longer alters the auditory spatial processing. Maybe in the future, HMD will be smaller and therefore have less impact on the HRTF.

Limitations and perspectives

Age range of participants. In this study, we deliberately spread out the age range of the participants. Our objective was to assess the feasibility of our approach in both young and older adults, as this may prove useful when assessing spatial localization ability in hearing impaired populations. We reached our goal as all the participants included in this protocol followed the instructions and attended the 40-min experiment. However, performance variability was likely introduced into the group of participants. The ability

to process auditory spatial information changes over the lifetime and auditory localization accuracy deteriorates in older adults (see Freigang et al. 2015).

When participants responded to the emitted sound by holding their hand at the perceived sound position, no visual feedback of the hand position was given. We limited visual cues from the environment as we wanted to avoid visuo-motor training effect that could be used by the subject to modify his head strategy and/or auditory perception of sound localization. As discussed by Ahrens and colleagues (2019), providing visual information might help to learn possible source locations, which can improve localization accuracy. Moreover, visual cues could influence sound localization abilities: minimal visual spatial frame benefits sound localization task (Valzolgher et al. 2020a, b), a reaching to sound localization task coupled with visual feedback modified head-movement behavior and improves sound localization performance (Valzolgher, et al. 2020a, b). Future studies could manipulate the visual scene and/or the vision of the hand to answer the multifactorial nature of spatial audition.

Finally, we wanted to examine the 3D sound localization performance in the reaching space. This space is actually quite relevant for humans: the near-field portion of space is particularly relevant for social interactions, where fast motor responses are needed in case of an approaching auditory object (e.g., a mosquito), when reaching toward a sound source (e.g., our phone ringing) or when orienting toward a nearby talker. Noticeably, a recent study (Valzolgher et al. 2020a, b) has shown that the ability to interact with a sound in the reaching space improves localization performance and promotes head movements, and this interaction also benefits spatial hearing rehabilitation (Valzolgher et al. 2022). As the participants localize sound sources manually, far-field stimulation is not feasible with our actual setup. However, by adapting the response method (i.e., use of a virtual pointer instead of hand-reaching, or by measuring head-direction as in Valzolgher et al. 2020a, b), this limitation could be addressed.

Conclusion

Researchers agree on the general notion that spatial hearing is an active and multisensory task. However, this awareness led to little adjustments to the methodological approach typically used when studying this fundamental perceptual ability. For instance, the study of head movements in sound localization remained largely overlooked. One reason for this discrepancy may reside in the fact that considering head movement has been problematic for most approaches to sound localization. In studies relying on sounds delivered in a real environment, head movements have mostly been prevented, to ensure reproducibility of sound source position across trials and participants,

or remained uncontrolled. When virtual sounds were generated through HRTFs, implementation of head movement responses in real time is still a computationally challenging task.

The approach we proposed and tested in the present study (SPHERE, European patent n° EP 3,463,084 A1, (Salemme et al. 2021) is a valid tool to accurately sample spatial abilities in auditory perception all around the listener, with minimal constraints on the participant or experimenter. Most interestingly, SPHERE proved sensitive for detecting and quantifying the contribution of free head motion during sound emission, with improvements to sound localization accuracy and precision. The SPHERE approach has been used recently with adult and pediatric populations, on both normal-hearing participants and cochlear implant patients (Coudert et al. 2022; Valzolgher, et al. 2020a, b, 2020a; Valzolgher et al. 2022). It offers a highly versatile opportunity to assess normal and pathological sound localization performance in a more ecologically valid approach (for discussion see Russell 2022). Finally, our approach paves the way for future research, clinical and industrial applications that will leverage the full potential offered by having embedded a VR HMD in the SPHERE system.

Acknowledgements This work was supported by LabEx CORTEX (ANR-11-LABX-0042) of University of Lyon, the Medisite Foundation (France), ITMO New technologies for Neuroscience, IDEX Lyon 'Senses In Space Lab', ANR (16-CE17-0016-01, VIRTUALHEARING3D, France), CNRS-PICS (PHC Galilée), and from IHU CeSaMe ANR-10-IBHU-0003. We thank K.Reilly for her careful proofreading, and the administrative staff of the IMPACT team for their administrative and informatics support.

Author contributions VG, AF and FP conceived the study; RS, EK and CD developed the technical aspects of the system; VG, AC and FP analyzed the data and prepared a draft of the manuscript; VG, AC, ET, AF and FP revised and finalized the final version of the manuscript. Additionally, each of the authors confirms that this manuscript has not been previously published and is not currently under consideration by any other journal. All of the authors have approved the contents of this paper and have agreed to the Experimental Brain Research's submission policies. VG, RS, EK, AF and FP own a patent for the system reported in this study.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

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 licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence 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 licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Ahrens A, Lund KD, Marschall M, Dau T (2019) Sound source localization with varying amount of visual information in virtual reality. *PLoS ONE* 14(3):e0214603. <https://doi.org/10.1371/journal.pone.0214603>
- Alais D, Burr D (2004) The ventriloquist effect results from near-optimal bimodal integration. *Curr Biol* 14(3):257–262. <https://doi.org/10.1016/j.cub.2004.01.029>
- Andéol G, Simpson BD (2016) Editorial : how, and why, does spatial-hearing ability differ among listeners? What is the role of learning and multisensory interactions? *Front Neurosci*. <https://doi.org/10.3389/fnins.2016.00036>
- Bahill AT, McDonald JD (1983) Frequency limitations and optimal step size for the two-point central difference derivative algorithm with applications to human eye movement data. *IEEE Trans Biomed Eng, BME* 30(3):191–194
- Bahu H, Carpentier T, Noisternig M, Warusfel O (2016) Comparison of different egocentric pointing methods for 3D sound localization experiments. *Acta Acust Acust* 102(1):107–118. <https://doi.org/10.3813/AAA.918928>
- Brimijoin WO, Boyd AW, Akeroyd MA (2013) The contribution of head movement to the externalization and internalization of sounds. *PLoS ONE* 8(12):e83068. <https://doi.org/10.1371/journal.pone.0083068>
- Brungart DS, Durlach NI, Rabinowitz WM (1999) Auditory localization of nearby sources. II. Localization of a broadband source. *J Acoust Soc Am* 106(4):1956–1968. <https://doi.org/10.1121/1.427943>
- Bulkin DA, Groh JM (2006) Seeing sounds : visual and auditory interactions in the brain. *Curr Opin Neurobiol* 16(4):415–419. <https://doi.org/10.1016/j.conb.2006.06.008>
- Coudert A, Gaveau V, Gatel J, Verdelet G, Salemme R, Farne A, Pavani F, Truy E (2022) Spatial hearing difficulties in reaching space in bilateral cochlear implant children improve with head movements. *Ear Hear* 43(1):192–205. <https://doi.org/10.1097/AUD.0000000000001090>
- Dunai L, Peris-Fajarns G, Magal-Royo T, Defez B, Santiago V (2011) Virtual moving sound source localization through headphones. In: Strumillo P (ed) *Advances in sound localization*. InTech
- Freigang C, Richter N, Rübsem R, Ludwig AA (2015) Age-related changes in sound localisation ability. *Cell Tissue Res* 361(1):371–386. <https://doi.org/10.1007/s00441-015-2230-8>
- Gaveau V, Pélissier D, Blangero A, Urquizar C, Prablanc C, Vighetto A, Pisella L (2008) Saccade control and eye–hand coordination in optic ataxia. *Neuropsychologia* 46(2):475–486. <https://doi.org/10.1016/j.neuropsychologia.2007.08.028>
- Genovese A, Zalles G, Reardon G, Roginska A (2018, août) Acoustic perturbations in HRTFs measured on mixed reality headsets. In: Audio engineering society conference: 2018 AES international conference on audio for virtual and augmented reality. <http://www.aes.org/e-lib/browse.cfm?elib=19699>
- Genzel D, Firzlaff U, Wiegrebe L, MacNeilage PR (2016) Dependence of auditory spatial updating on vestibular, proprioceptive, and efference copy signals. *J Neurophysiol* 116(2):765–775. <https://doi.org/10.1152/jn.00052.2016>
- Genzel D, Schutte M, Brimijoin WO, MacNeilage PR, Wiegrebe L (2018) Psychophysical evidence for auditory motion parallax. *Proc Natl Acad Sci* 115(16):4264–4269. <https://doi.org/10.1073/pnas.1712058115>
- Goossens HHL, van Opstal AJ (1999) Influence of head position on the spatial representation of acoustic targets. *J Neurophysiol* 81(6):2720–2736. <https://doi.org/10.1152/jn.1999.81.6.2720>

- Groh JM, Sparks DL (1992) Two models for transforming auditory signals from head-centered to eye-centered coordinates. *Biol Cybern* 67(4):291–302. <https://doi.org/10.1007/BF02414885>
- Gupta R, Ranjan R, He J, Woon-Seng G (2018, août) Investigation of effect of VR/AR headgear on head related transfer functions for natural listening. In: Audio engineering society conference: 2018 AES international conference on audio for virtual and augmented reality. <http://www.aes.org/e-lib/browse.cfm?elib=19697>
- Haber L, Haber RN, Penningroth S, Novak K, Radowski H (1993) Comparison of nine methods of indicating the direction to objects : data from blind adults. *Perception* 22(1):35–47. <https://doi.org/10.1068/p220035>
- Honda A, Shibata H, Hidaka S, Gyoba J, Iwaya Y, Suzuki Y (2013) Effects of head movement and proprioceptive feedback in training of sound localization. *I-Perception* 4(4):253–264. <https://doi.org/10.1068/i0522>
- Kearney G, Gorzel M, Boland F, Rice H (2010) Depth perception in interactive virtual acoustic environments using higher order ambisonic soundfields. In: 2nd International Symposium on Ambisonics and Spherical Acoustics
- Kim J, Barnett-Cowan M, Macpherson EA (2013) Integration of auditory input with vestibular and neck proprioceptive information in the interpretation of dynamic sound localization cues. 050142-050142. <https://doi.org/10.1121/1.4799748>
- Kolarik AJ, Moore BCJ, Zahorik P, Cirstea S, Pardhan S (2016) Auditory distance perception in humans: a review of cues, development, neuronal bases, and effects of sensory loss. *Atten Percept Psychophys* 78(2):373–395. <https://doi.org/10.3758/s13414-015-1015-1>
- Kopčo N, Shinn-Cunningham BG (2011) Effect of stimulus spectrum on distance perception for nearby sources. *J Acoust Soc Am* 130(3):1530–1541. <https://doi.org/10.1121/1.3613705>
- Kuk F, Keenan DM, Lau C, Crose B, Schumacher J (2014) Evaluation of a localization training program for hearing impaired listeners. *Ear Hear* 35(6):652–666. <https://doi.org/10.1097/AUD.000000000000067>
- Kumpik DP, Campbell C, Schnupp JWH, King AJ (2019) Re-weighting of sound localization cues by audiovisual training. *Front Neurosci* 13:1164. <https://doi.org/10.3389/fnins.2019.01164>
- Lambert RM (1974) Dynamic theory of sound-source localization. *The Journal of the Acoustical Society of America* 56(1):165–171. <https://doi.org/10.1121/1.1903248>
- Lewald J, Ehrenstein WH (1996) The effect of eye position on auditory lateralization. *Exp Brain Res*. <https://doi.org/10.1007/BF00227270>
- Lewald J, Karnath H-O (2000) Vestibular influence on human auditory space perception. *J Neurophysiol* 84(2):1107–1111. <https://doi.org/10.1152/jn.2000.84.2.1107>
- Lewald J, Karnath H-O, Ehrenstein WH (1999) Neck-proprioceptive influence on auditory lateralization. *Exp Brain Res* 125(4):389–396. <https://doi.org/10.1007/s002210050695>
- Lida K (2019) Head-related transfer function and acoustic virtual reality. Springer Singapore. <https://doi.org/10.1007/978-981-13-9745-5>
- Litovsky RY, Parkinson A, Arcaroli J (2009) Spatial hearing and speech intelligibility in bilateral cochlear implant users. *Ear Hear* 30(4):419–431. <https://doi.org/10.1097/AUD.0b013e3181a165be>
- Maddox RK, Pospisil DA, Stecker GC, Lee AKC (2014) Directing eye gaze enhances auditory spatial cue discrimination. *Curr Biol* 24(7):748–752. <https://doi.org/10.1016/j.cub.2014.02.021>
- Majdak P, Goupell MJ, Laback B (2011) Two-dimensional localization of virtual sound sources in cochlear-implant listeners. *Ear Hear* 32(2):198–208. <https://doi.org/10.1097/AUD.0b013e3181f4dfe9>
- McAnally KI, Martin RL (2014) Sound localization with head movement: implications for 3-d audio displays. *Front Neurosci*. <https://doi.org/10.3389/fnins.2014.00210>
- Middlebrooks JC (2015) Sound localization. *Handbook of Clinical Neurology*, vol 129. Elsevier, pp 99–116
- Middlebrooks JC, Green DM (1991) Sound localization by human listeners. *Annu Rev Psychol* 42(1):135–159. <https://doi.org/10.1146/annurev.ps.42.020191.001031>
- Midway S, Robertson M, Flinn S, Kaller M (2020) Comparing multiple comparisons: practical guidance for choosing the best multiple comparisons test. *PeerJ* 8:e10387. <https://doi.org/10.7717/peerj.10387>
- Mueller MF, Meisenbacher K, Lai W-K, Dillier N (2014) Sound localization with bilateral cochlear implants in noise: how much do head movements contribute to localization? *Cochlear Implants Int* 15(1):36–42. <https://doi.org/10.1179/1754762813Y.0000000040>
- Nava E, Bottari D, Bonfilioli F, Beltrame MA, Pavani F (2009) Spatial hearing with a single cochlear implant in late-implanted adults. *Hear Res* 255(1–2):91–98. <https://doi.org/10.1016/j.heares.2009.06.007>
- Oldfield SR, Parker SPA (1984) Acuity of sound localisation: a topography of auditory space. I. Normal hearing conditions. *Perception* 13(5):581–600. <https://doi.org/10.1068/p130581>
- Parsehian G, Jouffrais C, Katz BFG (2014) Reaching nearby sources: comparison between real and virtual sound and visual targets. *Front Neurosci*. <https://doi.org/10.3389/fnins.2014.00269>
- Pastore MT, Natale SJ, Yost WA, Dorman MF (2018) Head movements allow listeners bilaterally implanted with cochlear implants to resolve front-back confusions. *Ear Hear* 39(6):1224–1231. <https://doi.org/10.1097/AUD.0000000000000581>
- Pavani F, Farnè A, Làdavas E (2003) Task-dependent visual coding of sound position in visuospatial neglect patients. *NeuroReport* 14(1):99–103. <https://doi.org/10.1097/00001756-200301200-00019>
- Pavani F, Husain M, Driver J (2008) Eye-movements intervening between two successive sounds disrupt comparisons of auditory location. *Exp Brain Res* 189(4):435–449. <https://doi.org/10.1007/s00221-008-1440-7>
- Pavani F, Meneghello F, Làdavas E (2001) Deficit of auditory space perception in patients with visuospatial neglect. *Neuropsychologia* 39(13):1401–1409. [https://doi.org/10.1016/S0028-3932\(01\)00060-4](https://doi.org/10.1016/S0028-3932(01)00060-4)
- Perrett S, Noble W (1997a) The contribution of head motion cues to localization of low-pass noise. *Percept Psychophys* 59(7):1018–1026. <https://doi.org/10.3758/BF03205517>
- Perrett S, Noble W (1997b) The effect of head rotations on vertical plane sound localization. *J Acoust Soc Am* 102(4):2325–2332. <https://doi.org/10.1121/1.419642>
- Pollack I, Rose M (1967) Effect of head movement on the localization of sounds in the equatorial plane. *Percept Psychophys* 2(12):591–596. <https://doi.org/10.3758/BF03210274>
- Populin LC (2006) Monkey sound localization : head-restrained versus head-unrestrained orienting. *J Neurosci* 26(38):9820–9832. <https://doi.org/10.1523/JNEUROSCI.3061-06.2006>
- Rabini G, Altobelli E, Pavani F (2019) Interactions between egocentric and allocentric spatial coding of sounds revealed by a multisensory learning paradigm. *Sci Rep* 9(1):7892. <https://doi.org/10.1038/s41598-019-44267-3>
- Rakerd B, Hartmann WM (1986) Localization of sound in rooms, III : onset and duration effects. *J Acoust Soc Am* 80(6):1695–1706. <https://doi.org/10.1121/1.394282>
- Russell MK (2022) Age and auditory spatial perception in humans : review of behavioral findings and suggestions for future research. *Front Psychol* 13:831670. <https://doi.org/10.3389/fpsyg.2022.831670>

- Salemme R, Gaveau V, Farnè A, Pavani F, Koun E, Belle A (2021) Method and apparatus for acquiring a spatial map of auditory perception of a subject (Patent N° 11176727)
- Seeber BU, Baumann U, Fastl H (2004) Localization ability with bimodal hearing aids and bilateral cochlear implants. *J Acoust Soc Am* 116(3):1698–1709. <https://doi.org/10.1121/1.1776192>
- Távora-Vieira D, De Ceulaer G, Govaerts PJ, Rajan GP (2015) Cochlear implantation improves localization ability in patients with unilateral deafness. *Ear Hear* 36(3):e93–e98. <https://doi.org/10.1097/AUD.000000000000130>
- Thurlow WR, Mergener JR (1970) Effect of stimulus duration on localization of direction of noise stimuli. *J Speech Hear Res* 13(4):826–838. <https://doi.org/10.1044/jshr.1304.826>
- Tollin DJ, Populin LC, Moore JM, Ruhland JL, Yin TCT (2005) Sound-localization performance in the cat : the effect of restraining the head. *J Neurophysiol* 93(3):1223–1234. <https://doi.org/10.1152/jn.00747.2004>
- Tyler RS, Witt SA, Dunn CC, Wang W (2010) Initial development of a spatially separated speech-in-noise and localization training program. *J Am Acad Audiol* 21(06):390–403. <https://doi.org/10.3766/jaaa.21.6.4>
- Valzolgher C, Alzhaler M, Gessa E, Todeschini M, Nieto P, Verdelet G, Salemme R, Gaveau V, Marx M, Truy E, Barone P, Farnè A, Pavani F (2020a) The impact of a visual spatial frame on real sound-source localization in virtual reality. *Curr Res Behav Sci* 1:100003. <https://doi.org/10.1016/j.crbeha.2020.100003>
- Valzolgher C, Todeschini M, Verdelet G, Gatel J, Salemme R, Gaveau V, Truy E, Farnè A, 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, Farnè A, Pavani F (2020b) Reaching to sounds in virtual reality : a multisensory-motor approach to promote adaptation to altered auditory cues. *Neuropsychologia* 149:107665. <https://doi.org/10.1016/j.neuropsychologia.2020.107665>
- Van Barneveld DCPBM, Binkhorst F, Van Opstal AJ (2011) Absence of compensation for vestibular-evoked passive head rotations in human sound localization : sound localization under passive vestibular stimulation. *Eur J Neurosci* 34(7):1149–1160. <https://doi.org/10.1111/j.1460-9568.2011.07844.x>
- Van Grootel TJ, Van Wanrooij MM, Van Opstal AJ (2011) Influence of static eye and head position on tone-evoked gaze shifts. *J Neurosci* 31(48):17496–17504. <https://doi.org/10.1523/JNEUROSCI.5030-10.2011>
- Vliegen J, Van Opstal AJ (2004) The influence of duration and level on human sound localization. *J Acoust Soc Am* 115(4):1705–1713. <https://doi.org/10.1121/1.1687423>
- Wallach H (1940) The role of head movements and vestibular and visual cues in sound localization. *J Exp Psychol* 27(4):339–368. <https://doi.org/10.1037/h0054629>
- Wightman FL, Kistler DJ (1992) The dominant role of low-frequency interaural time differences in sound localization. *J Acoust Soc Am* 91(3):1648–1661. <https://doi.org/10.1121/1.402445>
- Wightman FL, Kistler DJ (1999) Resolution of front–back ambiguity in spatial hearing by listener and source movement. *J Acoust Soc Am* 105(5):2841–2853. <https://doi.org/10.1121/1.426899>
- Yon D, de Lange FP, Press C (2019) The predictive brain as a stubborn scientist. *Trends Cogn Sci* 23(1):6–8. <https://doi.org/10.1016/j.tics.2018.10.003>
- Zahorik P (2002) Assessing auditory distance perception using virtual acoustics. *J Acoust Soc Am* 111(4):1832–1846. <https://doi.org/10.1121/1.1458027>
- Zahorik P, Wightman FL (2001) Loudness constancy with varying sound source distance. *Nat Neurosci* 4(1):78–83. <https://doi.org/10.1038/82931>
- Zahorik P, Brungart DS, Bronkhorst AW (2005) Auditory distance perception in humans : a summary of past and present research. *Acta Acust Acust* 91:409–420

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

Capturing Visual Attention With Perturbed Auditory Spatial Cues

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 Pavani^{1,2,7}

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



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

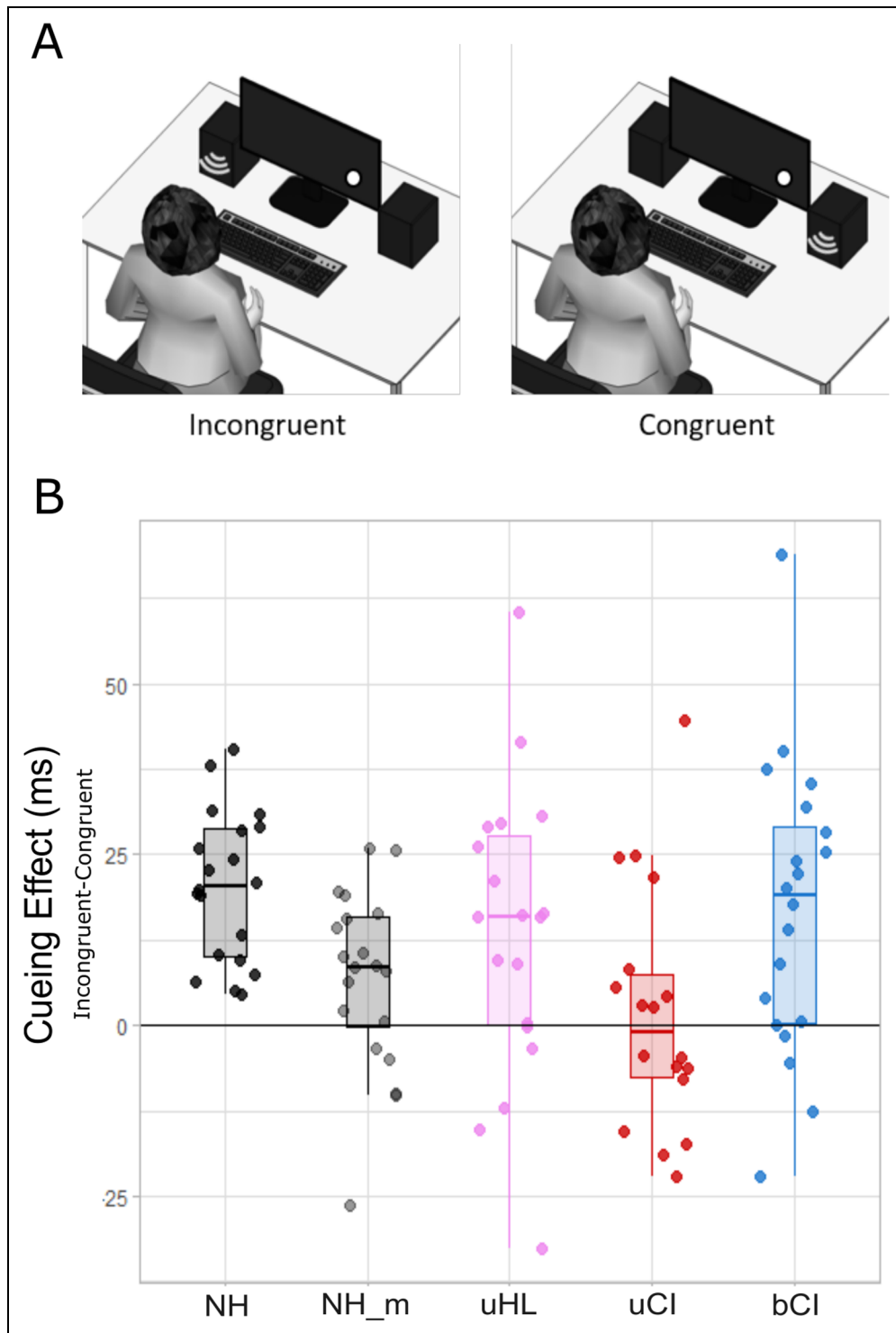


Figure 1. AV cueing. (A) Experimental setting: left: example of the incongruent audio-visual condition (the sound and the visual stimulus are presented on the opposite sides of the space); right: example of the congruent condition (the sound and the visual stimulus are presented on the same side of the 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

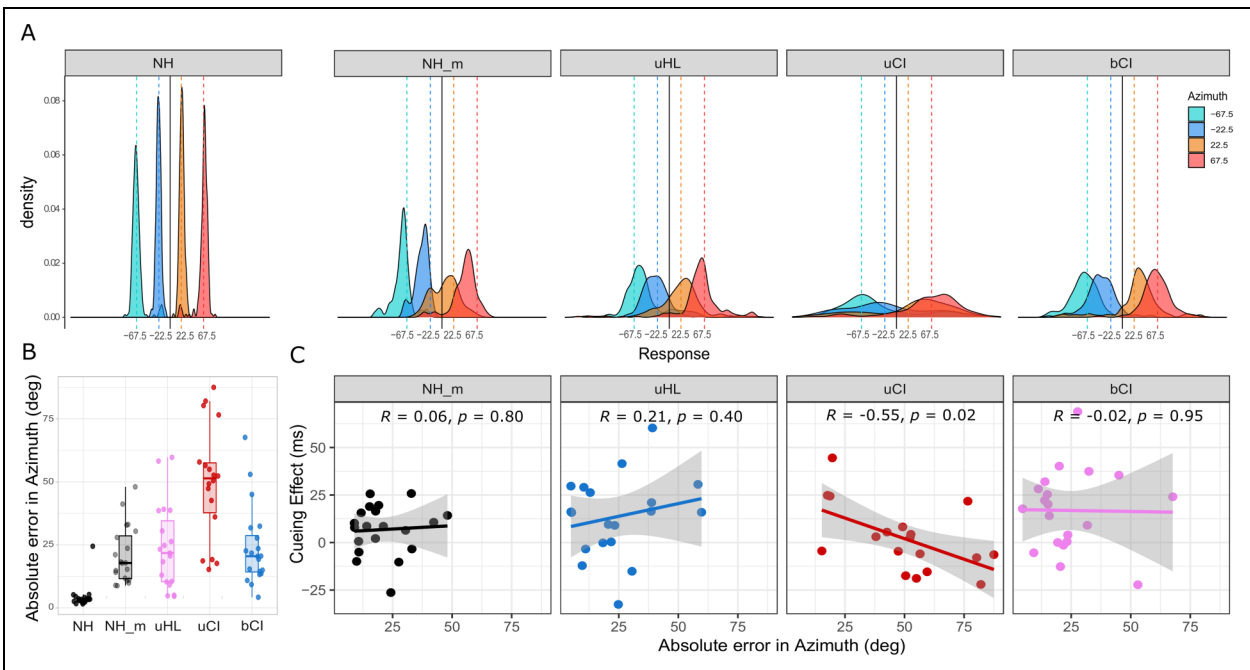


Figure 2. Sound-localization task. (A) All-trial-response density plotted using `geom_density` function in R-studio. This function draws a smoothed version of the histogram (kernel density estimate). Data are colored as a function of target azimuth positions (-67.5° , -22.5° , 22.5° , and 67.5°) and 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

1. 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>



A new tool for the assessment of speech understanding and spatial hearing difficulties in children: the Kid-SSQ questionnaire

Aurélie Coudert^{1,2,3} · Karen T. Reilly⁴ · Eric Truy^{1,2,3} · Annie Moulin⁵ · Valérie Gaveau¹

Received: 6 October 2022 / Accepted: 17 January 2023

© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2023

Abstract

Purpose To develop and validate a new questionnaire, the Kid-SSQ, for the rapid screening of hearing abilities in children with hearing impairment, aged 7–17 years.

Methods The questionnaire was constructed from two existing, validated versions of the ‘Speech, Spatial and Qualities of Hearing’ - (SSQ) questionnaire (pediatric form and adult short-form). The 12 selected items included auditory aspects from three subscales: speech perception, spatial hearing, and qualities of hearing. This new short form was then validated in 154 children with cochlear implants (100 bilaterally, and 54 unilaterally implanted children). Construct validity was assessed by testing relationships between Kid-SSQ scores and objective clinical parameters (e.g., age at test, pure-tone audiometry-PTA threshold, speech reception threshold-SRT, duration of binaural experience).

Results Completion time was acceptable for use with children (less than 10 min) and the non-response rate was less than 1%. Good internal consistency was obtained (Cronbach’s $\alpha = 0.78$), with a stable internal structure corresponding to the 3 intended subscales. External validity showed the specificity of each subscale: speech subscale scores were significantly predicted ($r = 0.32$, $p < 0.001$) by both 2 kHz PTA threshold ($\beta = 0.33$, $p < 0.001$) and SRT ($\beta = -0.23$, $p < 0.001$). Children with more binaural experience showed significantly higher scores on the spatial subscale than children with less binaural experience ($F(1,98) = 5.1$, $p < 0.03$) and the qualities of hearing subscale scores significantly depended on both age and SRT ($r = 0.32$, $p < 0.001$).

Conclusions The Kid-SSQ questionnaire is a robust and clinically useful questionnaire for self-assessment of difficulties in various auditory domains.

Keywords Cochlear implant · Quality of life · Speech perception · Spatial hearing · Health questionnaire · Hearing loss · Hearing aid

Abbreviations

BCI Bilateral cochlear implant
CI Cochlear implant
PTA Pure tone audiometry

Annie Moulin and Valérie Gaveau have contributed equally to this work.

✉ Annie Moulin
annie.moulin@cns.fr

¹ Université Claude Bernard Lyon 1, CNRS, INSERM, Centre de Recherche en Neurosciences de Lyon CRNL U1028 UMR5292, IMPACT (Integrative Multisensory Perception Action and Cognition Team, Bron, France

² Department of Pediatric Otolaryngology–Head and Neck Surgery, Femme Mere Enfant Hospital-Hospices Civils de Lyon, Lyon, France

³ Department of Otolaryngology–Head and Neck Surgery, Edouard Herriot Hospital-Hospices Civils de Lyon, Lyon, France

⁴ Université Claude Bernard Lyon 1, CNRS, INSERM, Centre de Recherche en Neurosciences de Lyon CRNL U1028 UMR5292, Trajectoires Team, Bron, France

⁵ Université Claude Bernard Lyon 1, CNRS, INSERM, Centre de Recherche en Neurosciences de Lyon CRNL U1028 UMR5292, Perception, Attention and Memory Team, 95 Bd Pinel, Bat 452, 69 675 Bron CEDEX, France

SRT Speech reception threshold
 SSQ Speech spatial and qualities of hearing scale
 UCI Unilateral cochlear implant

Introduction

The expansion of clinical indications for cochlear implantation in recent decades (e.g. [1]) has greatly improved treatment for severe to profound bilateral deafness in children. While this has notably improved these children's language development (for review, see [2]), many cochlear implant (CI) children still complain of difficulties in sound localization and speech understanding in noisy environments, even several years after implantation. These difficulties are largely undiagnosed and unreported during clinical follow-up, and can impair long-term social integration, making their detection and clinical care an important public health issue. Thanks to their ease of use and sensitivity, quality of life questionnaires are commonly used screening tools. However, most auditory quality of life questionnaires consider disability as a whole, and lack specificity for the diagnosis of auditory difficulties (see for example the KINDL Cochlear Implant Module questionnaire adapted for young CI users [3]). Designed for adult population with hearing-impairment, the Speech, Spatial, and Qualities of Hearing Scale (SSQ) [4], is aimed at encompassing the widest range of aspects of hearing in daily life of patients and, with its availability in multiple languages and its various forms, has been used in more than 180 international publications, showing its usefulness and relevance in the audiological field. This 49 item long scale consists in depicting a situation met in daily life (e.g., "You are in a group of about five people in a busy restaurant. You can see everyone else in the group. Can you follow the conversation?"), and inviting patients to answer on a 11 points visuo analog scale, graded from 0 ("Not at all") to 10 ("Perfectly"). Patients have the option to tick a "not applicable" (NA) box if they feel the situation depicted doesn't apply to them. The 49 items are organized in 3 subscales (i.e. speech perception, spatial hearing, and qualities of hearing), that have shown to provide, each, a specific index of hearing abilities [5, 6]. Unfortunately, the wealth of information brought by this more than 1000 word long scale comes at the price of a cognitive burden for the patients [7], that can compromise its acceptability, and its feasibility in clinical settings. To overcome this problem, several short forms have been built [8], the latest, in 15 items, (15iSSQ [9]) being created based on a data-driven approach, showing good psychometric properties and keeping the 3 original subscales.

Adapting adult scales for pediatric population involves adapting language, response types and length of scale [10–12]. Galvin and Noble [13] adapted the original adult

SSQ for the assessment of children, taking specific care in adapting the wording and the situations depicted to children's daily lives. While it explores a broad range of everyday auditory skills, through the same 3 subscales as the original SSQ, this pediatric adaptation, aimed at children above 10 years of age, consists of 33 items and takes approximately 30 min to complete. Responses are therefore subject to bias due to attentional difficulties, which are particularly prevalent in young children with hearing loss. Moreover, some items, in the long SSQ versions, have a high non-response rate due to the complexity of the everyday situations they describe [6, 14, 15]. Length of scale is part of the response burden of a scale [11], and should be optimized so that it doesn't compromise the acceptability, nor reliability of the scale, hence our choice of aiming for a 12 item long scale, allowing 4 items per subscale, with specific adaptation for children.

The primary objective of this study was thus to develop and validate a short version of the SSQ for children, based on Galvin and Noble's work [13], and using the same data-driven approach as that used to develop the 15iSSQ short form [9], that is in agreement with the international guidelines regarding translation and cultural adaptation of health questionnaires [16, 17], validation of health questionnaires [18] and creating short-forms [19, 20].

The second objective was to assess the construct validity of this new short-form, in particular how this scale relates to several clinical variables in children fitted with one or two CIs, to propose a new routine assessment tool for subjective hearing difficulties encountered in daily life.

Methods

Participants

One hundred bilateral cochlear implant children (BCI) aged 7.3–17.8 years (mean \pm standard deviation; 12.5 ± 2.9 years) and 54 unilateral cochlear implant children (UCI) aged 7.7–17.8 years (13.8 ± 2.2 years) were included in this study. Four children in the UCI group had bilateral implants before inclusion in the study but had opted to maintain only one CI, and 67% of UCI children had a contralateral hearing aid. For the BCI and UCI groups the three main hearing loss etiologies were "unknown" (42% and 39%), "genetic" (35% and 33%), and "acquired" (9% and 11%). In terms of schooling, 21% of the BCI children and 39% of the UCI children were enrolled in adapted schooling. The demographic and audiometric characteristics of each group are summarized in online resource, Appendix 1.

Questionnaire construction

We used two versions of the Speech Spatial and Qualities of hearing (SSQ) questionnaire as our reference: its short form for adult hearing assessment (the 15iSSQ [9]) and its long form for children (the SSQ33 [13]). Based on these two questionnaires, we developed a short pediatric form which we named the Kid-SSQ. A short questionnaire created from one or more forms of a validated questionnaire should [20]: (1) maintain the same internal structure as the reference forms, i.e. same number of subscales, (2) include between 3 and 5 items per subscale [21], and (3) exclude items with poor psychometric properties (e.g., high non response rates) and poor discriminability between pathological condition and controls. As the 15iSSQ was built using this data-driven approach, encompassing data and results from several international studies, we used the same criteria to create the Kid-SSQ. In addition, we also used data available from the long form for children SSQ [13] and our clinical experience of its use [14]. We included 4 items from each of the 3 subscales of the already-validated versions of the SSQ questionnaire, each representing a different hearing function (subscale A: speech perception, subscale B: spatial hearing, subscale C: qualities of hearing). Each item was assigned a letter (subscale) and a number (item) (e.g. for subscale A, each item was labelled from #A.1 to #A.4, for subscale B from #B.1 to #B.4, and for subscale C from #C.1 to #C.4). Items were selected based on their psychometric properties, on their performance in the SSQ33, and with the goals of covering a wide range of hearing situations [22] and being understood by children as young as 7 years old. The Kid-SSQ shares 9 of its 12 items with the 15iSSQ, and all 12 items are in the long pediatric version of the SSQ. An abbreviated text of

the communication situation depicted in each of the items is provided in Table 1 and details of the items included in the Kid-SSQ can be found in the Online Resource, Appendices 2 and 3.

We used the validated, French language version of the SSQ [7] to create a new, pediatric short-form. We retained the same rating scale as that used in the two reference forms: zero for "I can't do it at all" to ten for "I can do it perfectly". To assist younger children with the self-assessment task we added three smileys ("sad" for a rating of 0, "neutral" for a rating of 5, and "happy" for a rating of 10), as well as a color gradient (light for 0, to dark for 10). We also modified the shape of the scale to look like a funnel mimicking the increasing intensity of the response (see online resource, Appendix 4A). Despite these additions, pilot experiments revealed that a minority of children under 9 were still unable to use the numerical scale, so an additional scale was developed without numbers, based on the face scale used for pain assessment (Appendix 4B). The color scheme and the chosen smileys are widely used in French schools from nursery level onwards. Those modifications are among those typically recommended by international guidelines for pediatric health scales creation [10, 12]. The original questionnaire in French is available in online resource, Appendix 5.

The questionnaire was administered only once, using a paper and pen format during a face-to-face interview with each child. The items were read aloud by the experimenter (or by the child) and after reading the item, if necessary, the experimenter provided an example from everyday life so that the child could more easily imagine him/herself in the described situation. Children were asked to choose a whole number and were told they could choose

Table 1 List of the listening situations depicted in the 12 items of the Kid-SSQ (the text of each item has been abbreviated, adapted from Bahn et al. [23], and not the actual text presented to children)

Subscales	Item number Kid-SSQ	Kid-SSQ Listening situations (abbreviated text)	Item number SSQ33
Speech perception	A1	Talking with one person with TV on	1.1
	A2	Conversation with 5 people in noise with vision	1.4
	A3	Conversation with 5 people in noise without vision	1.6
	A4	Follow one conversation when many people talking	1.8
Spatial hearing	B1	Locate a door slam in a familiar house*	2.4
	B2	Locate dog barking	2.5
	B3	Locate vehicle from footpath	2.6
	B4	Identify approach or recede (voice or footsteps)	2.8
Qualities of hearing	C1	Identify different people by voice	3.3
	C2	Distinguish familiar music	3.4
	C3	Distinguish 2 different sounds (running water /fan)	3.5
	C4	Recognize different familiar sounds in the house	3.2

The correspondence of the items numbers with the item number of the original pediatric scale are mentioned in bold in the rightmost column (further correspondence between the different forms of the SSQ scale are in the online resource appendix 2)

the response "not applicable" (NA) if they were unable to answer the item.

Statistical analyses

The validation procedure involved the different steps recommended for statistical validation of health scales [18]. Internal consistency of the Kid-SSQ was assessed by combining data from all patients and calculating both item-to-total correlations and Cronbach's alpha for each factor determined by factor and cluster analyses. Exploratory factor analysis was performed using the same parameters as previously used for the SSQ [5, 7, 9], namely a factor extraction performed on a non-parametric correlation matrix with a maximum likelihood method and an oblique rotation method. The details of this analysis is in the online resource, Appendix 6.

The external validity of the Kid-SSQ was evaluated by Pearson correlation and multi-regression analysis, between the mean score per subscale and various clinical parameters (i.e. age at test, pure tone auditory threshold—PTA, speech reception threshold—SRT, duration of exposure with the first CI, duration of binaural experience, and inter-implant interval).

Results

Kid-SSQ responses

The number of non-responses or "not applicable" (NA) answers was very low (<1%). Most of the "NA" answers were distributed between two items: 5 for #C.2

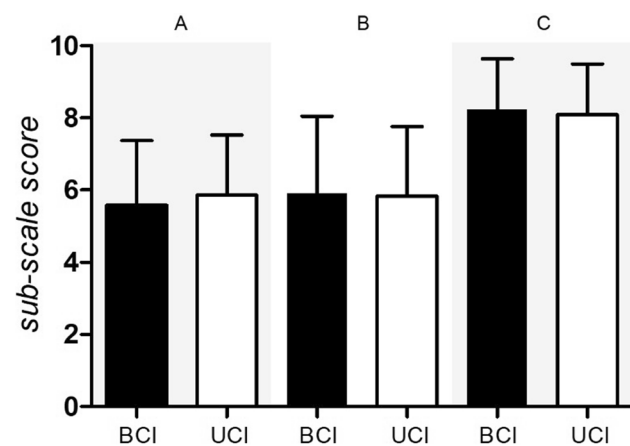


Fig. 1 Means and standard deviations on the three subscales for the bilateral cochlear implant (BCI, $n=100$) and unilateral cochlear implant (UCI, $n=54$) children (A: speech perception; B: spatial hearing; C: qualities of hearing)

(Identify different people by voice) and 4 for #B.4 (Identify approach or recede). Only 12 children had missing answers and 11 of these missed only one answer (one child missed 2). The median age of children with and without missing answers was similar (13 years). The face scale (Appendix 4B) was rarely necessary, only 5 BCI children under 9 years old used it.

The average total score on the Kid-SSQ for the whole sample was 6.6 (± 1.3) out of 10, and did not differ between the BCI and UCI groups (see ANOVA results below). Figure 1 and online resource Appendix 7 show the scores for each subscale and item separately for the BCI and UCI groups. A repeated measures ANOVA revealed that subscale scores differed between each other ($F(2, 304) = 131.86, p < 0.001, \eta_g^2 = 0.27$), but that scores on the three subscales were similar for both groups (no main effect of group nor a group by subscale interaction). Post-hoc tests revealed that subscale C (qualities of hearing) scores were higher than scores on the other two subscales ($p < 0.0001$, Bonferroni correction). Since the UCI group consisted of children with bimodal (i.e. fitted with a CI and a hearing aid) and monaural (i.e., with only a CI) configurations, scores for these two sub-groups are shown in Appendix 8. Monaural UCI children tended to have lower scores on all three subscales.

Factor analyses and internal consistency and reliability

The item to total correlations ranged from 0.22 (#C2) to 0.53 (average of 0.42), showing good internal consistency and Cronbach's alpha was 0.78 (0.76; 0.80) for the entire scale, and remained above 0.75 as each item was removed (Appendix 9).

The common factor analysis performed on 142 children explained 42% of the total variance with three clearly defined factors, and average communalities across items of 0.42, all communalities were above 0.21 (#C2) (factor loadings and communalities are detailed in Appendix 9). #A2 (Conversation in noise with vision) and #A3 (Conversation in noise without vision) showed a clear load on factor 1 (16% variance explained), with #A1 (Talking with one person with TV on) & #A4 (Follow one conversation when many people talking) cross-loading on factor 1 and 3 (#A4 loading more on factor 3). All B items loaded on factor 2 (13% of variance explained), with #B3 (Locate vehicle from footpath) and #B4 (Identify approach or recede) cross-loading on factors 2 and 3. #C2 (Identify familiar music), #C3 (Separation of two sounds) and #C4 (recognize different familiar sounds) clearly loaded on factor 3 (14% variance explained), with #C1 (Identify different people by voice) cross-loading on factors 2 and 3. This internal structure was confirmed by the cluster analysis (Fig. 2), which identified 3 higher-order

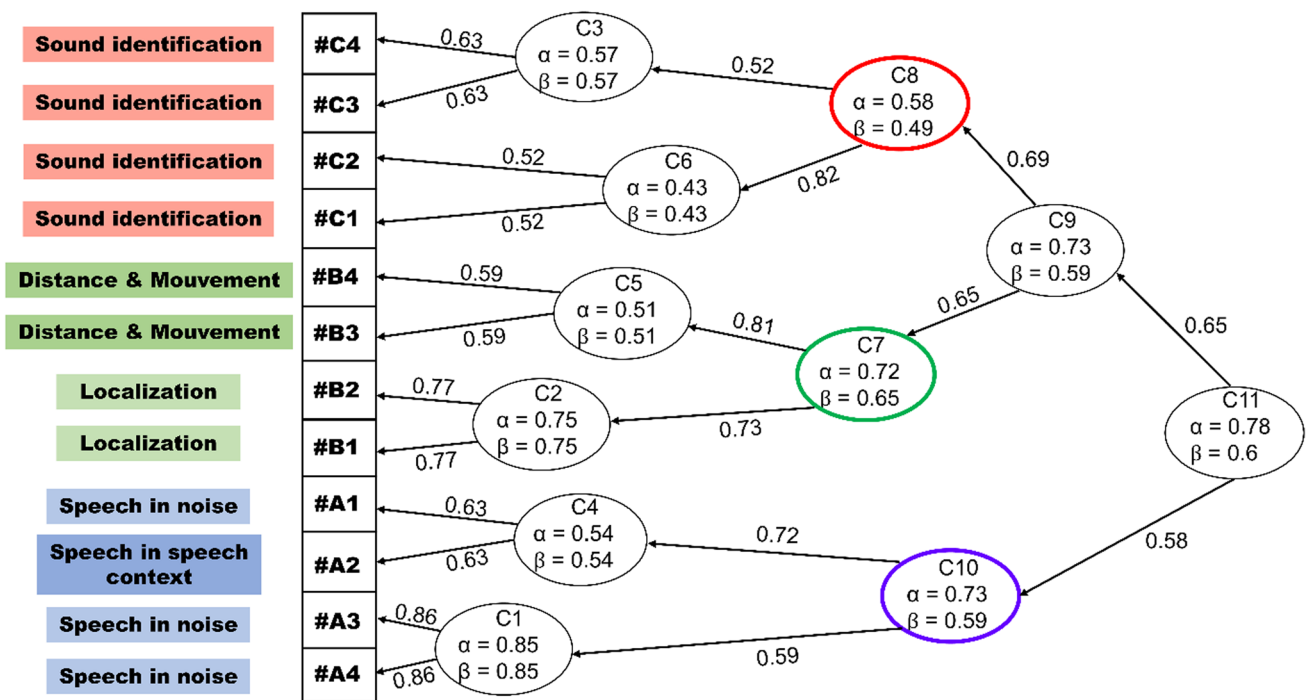


Fig. 2 Tree diagram of the cluster analysis. The 12 questionnaire items (4 from each subscale) are shown on the left in the colored rectangles and the most similar items are combined on the left of the figure.

The three main clusters are circled in red, green, and blue, and correspond to the C, B, and A subscales of the questionnaire. For each cluster, the alpha and beta coefficients are indicated

sub-clusters with good internal consistency and homogeneity, each comprising 4 items from the corresponding subscales of the questionnaire (cluster C10: #A1 to #A4; cluster C7: #B1 to #B4 and cluster C8: #C1 to #C4). Attempts to link these higher-order clusters together (i.e. C9, C10, and C11) resulted in a loss of homogeneity, evident in the larger differences between alpha and beta coefficients. This shows that our 3 higher-order clusters could not be combined into a single score, suggesting that the scale is not unidimensional, but instead explores 3 different aspects of hearing in our population. Correlations between the 3 different subscales ranged from $0.33, p < 5 \times 10^{-5}$ (A and C scales) to $0.41, p < 3 \times 10^{-7}$ (between B and C scales).

External validity—correlation with clinical parameters

For the BCI group, the Kid-SSQ total score did not correlate significantly with CI experience, PTA threshold, or SRT. A weak but significant increase in scores was observed with increasing age ($r=0.23, p=0.02$), mainly due to the significant correlation between age and the subscale C (qualities of hearing) score ($r=0.31, p < 0.002$), due to correlations between age and #C3 ($r=0.26, p=0.001$) and #C4 ($r=0.23, p=0.004$). Higher Kid-SSQ total scores were associated with higher binaural (averaged across right and left ears) PTA thresholds ($r=0.19, p=0.055$) and higher

Table 2 Pearson correlation coefficients between the different Kid-SSQ subscales and main auditory variables, in a population of 149 cochlear implant children (listwise deletion from the total population)

	Total Score	A speech score	B spatial score	C qualities score	SRT	PTA	2 kHz PTA
Speech score	0.75***						
Spatial score	0.82***	0.38***					
Qualities score	0.7***	0.33***	0.41***				
SRT	- 0.13	- 0.1	- 0.07	- 0.15			
PTA	0.04	0.15	- 0.03	- 0.04	0.48***		
2 kHz PTA	0.15	0.24*	0.05	0.04	0.38***	0.85***	
Age	0.22*	0.17	0.1	0.27**	0.04	0.1	0.12

Values in bold are statistically significant ($p < 0.05$), with * for $p < 0.01$, ** for $p < 0.001$ and *** for $p < 0.0001$

binaural 2 kHz PTA thresholds ($r=0.18$, $p=0.06$). Although those trends were not statistically significant, they can be attributed to the high correlation between subscale A scores and binaural PTA thresholds ($r=0.29$, $p<0.003$), and subscale A scores and binaural 2 kHz PTA threshold ($r=0.28$, $p<0.005$).

The different subscales showed different correlation patterns. First, subscale A (speech) scores correlated significantly with PTA threshold averaged over 0.5–4 kHz on both ears in the BCI group ($r=0.29$, $p=0.003$): scores improved as the average threshold increased. Similar results were obtained between speech scores and both left ear ($r=0.25$, $p=0.012$) and right ear hearing thresholds ($r=0.25$, $p=0.012$).

When considering the first implanted ear for the entire population (Table 2), the subscale A score tended to increase as the PTA increased ($r=0.15$, $p<0.07$). More detailed analyses revealed that this correlation was significantly stronger for the 2 kHz hearing threshold (HT) ($r=0.24$, $p<0.01$) than with for any lower frequency (e.g. $r=0.09$ for 1 kHz HT, $p=ns$, $z=2.25$, $p<0.025$) and that among the A items, A3 scores increased significantly with decreasing SRT thresholds ($r=-0.22$, $p=0.008$). The best explanatory model for A scores ($F(2, 148)=8.5$, $p<0.0004$, $r=0.32$) involved both SRT and 2 kHz HT as significant predictors, with 2 kHz HT being the strongest predictor ($\beta=0.33$, $t=3.9$, $p<0.0002$) versus SRT ($\beta=-0.23$, $t=-2.7$, $p=0.007$). When controlling for the 2 kHz HT, the A score decreased significantly (by 0.5 points) as SRT increased by 10 dB. Similar significant models were obtained for both A2 and A4, but they were not statistically significant for A1. For A3, SRT was the major predictor ($\beta=-0.32$) versus the 2 kHz HT ($\beta=0.27$), with a 0.94 points decrease in A3 score for each 10 dB increase in SRT.

Subscale B (spatial) scores tended to increase with longer experience with the second CI ($r=0.173$, $p=0.086$), and were not significantly correlated with age ($r=0.11$, $p=ns$) or experience with the first CI ($r=0.03$, $p=ns$), or ear asymmetry ($r=0.15$, $p=ns$).

When the group of BCI patients was median split based on amount of experience with bilateral CI, the Kid-SSQ total scores tended to be higher in the group with more bilateral CI experience ($F(1,98)=3.9$, $p=0.05$), and the subscale B (spatial) score was significantly higher in this group ($F(1,98)=5.1$, $p=0.026$). This was particularly true for B1 and B4.

Subscale C (qualities of hearing) scores increased significantly as age increased ($r=0.29$, $p<0.0004$), and with amount of experience with a first cochlear implant ($r=0.17$, $p<0.04$), but this latter effect was entirely mediated by age (direct effect: $z=0.5$, $p=ns$, indirect effect via age: $z=2.6$, $p<0.01$). The best predictive model for C scores ($F(2,148)=8.1$, $p<0.0005$, $r=0.32$) involved both age

($\beta=0.28$, $t=3.5$, $p<0.0007$) and, to a lesser degree, SRT ($\beta=-0.16$, $t=-2.1$, $p<0.04$) with a decrease in C score of 0.3 points as SRT increased by 10 dB, after having controlled for age. Detailed analysis of the subcluster formed by C1 (Identify different people by voice) and C2 (distinguish familiar music) showed a strong prediction only by SRT ($r=0.23$, $p=0.005$), with a decrease of 0.47 (se=0.16) by 10 dB increase in SRT. The subcluster formed by C3C4 showed a significant increase of 0.2 points (se=0.05) per year of age ($r=0.30$, $p<0.0002$).

Discussion

Assessment of quality of life in CI children is an essential part of their medical follow-up and clinical care. It is complementary to objective audiological testing and allows early detection of persistent hearing problems. In this study we developed a new short form of the pediatric SSQ, called the Kid-SSQ, and we validated it in a large cohort of 154 CI children. This new short form has 5 strengths: (1) it is specific to auditory disorders, (2) it is adapted for use with a large range of ages: from 7 to 17 years old, (3) completion is fast and simple, making it possible to perform during a medical visit, (4) it is sensitive enough to accurately quantify children's complaints, and (5) it has the same internal structure of 3 different subscales as the original SSQ, with each subscale being specific to a different aspect of hearing. In the literature, we did not find similar scales in other languages adapted in a short format for use in clinical practice.

Validation of a new short-form of the pediatric SSQ

The Kid-SSQ questionnaire consists of twelve items (3 subscales with 4 items each), selected from the original scale developed by Galvin and Noble (2013), using the same data-driven approach for short-form development as that used in developing the adult 15iSSQ short-form [9]. These items explore a wide variety of auditory functions: understanding speech in noise, spatial hearing of a stationary or moving source, identification of everyday sounds. Similar to both the original pediatric and adult scales, it has an internal structure with three well defined subscales (A—speech perception, B—spatial hearing, C—qualities of hearing), obtained by exploratory factor analysis and confirmed by cluster analysis. In our sample it achieved a Cronbach's α of 0.78, showing good reliability for a 12 item scale (as Cronbach's alpha decreases with the number of items [24]). The consistency indices revealed that items correlated well with each other, without multicollinearity, and thus showed a wide range of constructs.

The format and content of this new questionnaire were adapted to make it appropriate for self-assessment by

children less than 10 years old. Its completion time is less than 10 min and the numerical scale with smileys allowed children to easily rate their difficulties in a variety of daily-life situations (non-response rate was < 1%). In this study, we validated it using more than 150 children, the minimum requirement for stable exploratory factor analysis being 120 children (corresponding to ten children per item) [18, 25].

The multidimensional aspect of this short form is revealed by different response patterns on each subscale (i.e. notably a higher mean score on subscale C—qualities of hearing), by the cluster analysis, and by differences in correlation patterns between subscales and clinical parameters.

Subscale A (speech) scores were positively correlated with average PTA, and especially the 2 kHz hearing threshold (and to a lesser degree 4 kHz hearing threshold). While at first glance this result seems counterintuitive, since all items in subscale A depict situations of speech understanding in noisy environments, one possible explanation is that higher PTA at high frequencies leads to reduced perceived noise, an overall improvement in signal-to-noise ratio due to greater filtering of the higher frequencies, and thus better perception in noisy environments. Furthermore, after controlling for the 2 kHz hearing threshold, subscale A scores significantly increased (by 0.5 points) per 10 dB SRT decrease. This is particularly true for the #A3 item, pertaining to hearing in noise without visual clues: its score depends more on SRT than on 2 kHz hearing threshold, increasing by almost 1 point per 10 dB decrease of SRT: this shows the importance of speech perception when relying on purely auditory cues. For items #A2 (Conversation with 5 people noise with vision) and #A4 (Follow one conversation when many people talking), where visual clues are available and could help, the score is more closely linked to the 2 kHz hearing threshold, i.e., a threshold that allows a greater filtering-out of frequencies above 2 kHz, than to the SRT. This further emphasizes the importance of this type of questionnaire, i.e., evaluating situations encountered in daily life by children with cochlear implants, and the additional need to assess speech perception in environments closer to those frequently met every day (e.g. classroom conversations).

Subscale B (spatial hearing) scores increased with amount of binaural experience but did not correlate with inter-implant interval, contrary to the findings of a study using objective evaluation of spatial hearing abilities [14]. It is well known that inter-implant intervals greater than 18 months lead to asymmetric auditory function [26, 27] which could directly impair spatial hearing. However, it is possible that more than 24 months of auditory experience after the second CI could be sufficient to counteract auditory asymmetry and improve spatial hearing abilities [28]. Here, only #B1 (Locate a door slam at home) and #B4 (Identify approach or recede) showed significantly greater scores for children with more binaural experience: both these items

relate to localization in familiar and commonly-encountered situations (a door slamming at home, sounds made by moving sources such as people voices or steps), whereas #B2 (Locate dog barking) and #B3 (Locate vehicle from footpath) refer to localization in noisier and less familiar environments (street, outside). One possibility is that the situations depicted in items #B2 and #B3 are more difficult.

Subscale C (qualities of hearing) scores correlated positively with duration of exposure to the first CI and age at test. Mediation analysis showed that the entire effect of first cochlear implant experience was mediated by age. Cluster analysis identified 2 subclusters within scale C. The first one, C1C2 (Identify different people by voice and distinguish familiar music) significantly increased its scores (by almost 0.50 points) per 10 dB of SRT decrease, whereas the second one, C3C4 (distinguish and recognize different sounds), was heavily dependent on age. C1C2 deals with recognition of familiar sounds, and implies that this occurs in a quiet environment, which could explain the significant correlation with SRT, and not with PTA (by contrast with the A scores). Conversely, both #C3 and #C4 depended exclusively, and significantly, on age. These two items question the ability to separate and recognize two sounds close to one another, which is a particularly difficult skill. This suggests that this particular auditory function is driven by auditory maturation in childhood but that it also relies on early cochlear implantation, during the critical period. Indeed, many studies have shown that profound disruption of auditory function during the first year of life can jeopardize auditory development [29, 30] and lead to detrimental cortical reorganization. Furthermore, this evolution as a function of age reveals the importance of assessing Kid-SSQ scores across various ages in a normal-hearing population of children. Such data are necessary to ascertain if the strong correlation we observed here with age is specific to children with cochlear implants, or if it is linked to normal auditory neurodevelopment. Indeed, several studies have shown that children as old as 11 have significantly greater difficulty than young adults in perceiving speech in cocktail party noise [31]. Since only C3C4 items showed a strong correlation with age, it is unlikely that this correlation can be explained by difficulties younger children had in understanding the items. Indeed, the qualities SSQ items whose scores correlated with the number of years of education in an elderly population [6] (and therefore were more difficult to understand), have been eliminated from the 15iSSQ, and hence not included in the Kid-SSQ.

Comparison with the reference form

Galvin and Noble [13] were the first to develop a specific pediatric hearing questionnaire—the SSQ33. Initially conceived to evaluate hearing difficulties for all children older than 10 suffering from hearing loss, its use is largely limited

to clinical research, because it requires at least 30 min to complete. To our knowledge, no French deafness referral centre uses the SSQ33 (or a pediatric equivalent) as a clinical follow-up instrument for children with hearing aids or CI. This means that there is currently a lack of knowledge about the evolution of quality of life experienced by children themselves (and not their parents) over the years after fitting.

Rauch and colleagues [15] recently used the pediatric version of the SSQ33 to evaluate the long-term (3.5 years) outcomes of cochlear implantation in 11 children with profound hearing loss. Ten of the eleven children were between 5 and 9 years old, and only 5 children in the study successfully completed the questionnaire. Three years after cochlear implantation, the average scores of the SSQ33 were greater than what we observed for our 12-item version (subscale A: 8.6; subscale B: 7.3; subscale C: 8.2). This difference is attributable to the less severe hearing loss profile of the children in their study, who all had normal hearing in the ear contralateral to the CI, as well as to the fact that most of the children were less than 10 years old (younger than the lower limit fixed by Galvin and Noble [13]).

We recently evaluated spatial hearing performance in 18 BCI children between 8 and 17 years old using a 3D localization test and the pediatric SSQ33 [14]. No correlation emerged between objective spatial hearing abilities and children's self-assessments. Moreover, mean scores were greater than expected and similar across subscales (A: 7.2; B: 7.1; C: 7.6). It seems that the SSQ33 was not sensitive enough to highlight hearing difficulties in CI children because of its length and the inclusion of several items not relevant to children's understanding or lifestyle (e.g.: "speech in quiet", "sound naturalness"). Self-assessed hearing difficulties measured using the SSQ33 have probably been underestimated because of this scale's inclusion of insensitive items and items with high non-response rates.

Alternative forms to the long form of the pediatric SSQ and future research areas

Galvin and colleagues developed an alternative version of the SSQ for use by parents of children less than 6 years old [13, 32]. This questionnaire was adapted from the long form of the adult SSQ [4] and consists of 23 items. Its administration differs from the SSQ33 for children because the authors recommend that, to obtain accurate responses, parents should observe their child in daily life for 3 weeks (i.e. 1 week per subscale). In clinical practice, this type of monitoring is restrictive and runs a high risk of poor compliance.

The main use of this SSQ for parents has been in studies that aim to evaluate the benefits of bilateral cochlear implantation [33–36]. In these studies, parents completed the questionnaire before surgery (i.e., when children were in the UCI condition) and up to 2 years after (in the BCI condition). These studies

revealed that mean scores for the three subscales increased significantly after 2 years of binaural experience, with the greatest benefits observed on the spatial hearing items. However, in this young pediatric population, scores should be interpreted cautiously because some items can be difficult to score since they do not correspond to daily life of the youngest patients (e.g., "group conversations in a noisy environment" for children under 3 years old, see Lovett [34]). The test/retest reliability of the SSQ for parents and for children has recently been evaluated using a French language version of the pediatric SSQ scales [37], and has shown extremely good reproducibility for both children and parents ($r > 0.90$ for each subscale) with test/retest at a two weeks interval. Furthermore, scores from a group of normal hearing children have been found to be above 9 (out of 10) for the majority of items, with a standard deviation around 1.0, showing very good contrast with hearing impaired children. This is extremely encouraging as to the test/retest reproducibility of the Kid-SSQ and its discriminative power. As the majority of the Kid-SSQ items are shared with the 15iSSQ, and as the 15iSSQ has shown better sensitivity and specificity than the original SSQ long form as well as other short-forms [9], the Kid-SSQ should perform well in discriminating between pathological and control conditions. The results of the present study provide new, complementary information about self-perceived speech-in-noise abilities in children with cochlear implants. Speech scores improved by around 0.5 scale points (out of 10) per 10 dB of improvement in SRT. However, contrary to expectation, scores deteriorated with decreasing high frequency hearing threshold, probably because of a greater filtering-out of noise. An unusually bad score in the Speech scale associated with good SRT thresholds could trigger more detailed audiological assessment of speech in noise, to see if speech perception in noise could be improved by minor adjustments in high frequency pure-tone thresholds. Further validation studies are needed to investigate test-retest reliability, to establish sensitivity to changes in auditory skills across time, and to establish age-stratified norms in a normal-hearing population. It would also be interesting to expand this validation to children with a larger variety of auditory disorders, such as mild-to-moderate hearing loss and auditory processing disorders.

Conclusion

The 12-item Kid-SSQ questionnaire, presented and validated here in a large cohort of children with cochlear implants, showed good consistency, very good construct validity and feasibility in children. In addition, this questionnaire shares the same internal structure of the adult SSQ, with three subscales, each assessing a different aspect of hearing (speech in noise, spatial abilities and qualities of hearing). This will permit the assessment and follow-up of three different aspects of hearing. Furthermore, because the questionnaire reflects

difficulties encountered in daily life, it provides new and complementary insights into such difficulties to those provided by classic audiological tests. Not only can the Kid-SSQ be used in clinical practice to monitor the benefits of cochlear implantation and speech therapy rehabilitation over time, but it can also be used in children with a large variety of auditory disorders, such as mild-to-moderate hearing loss and auditory processing disorders. This would potentially enable all children with hearing difficulties to benefit from more effective clinical care and follow-up.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00405-023-07846-9>.

Acknowledgements A. Coudert was supported by a grant from Cochlear Corp. A. Moulin was supported by a research grant from the Auvergne Rhône Alpes Region (pack ambition recherche, project Effecbruit, 2018-2022) and by the CeLya labex (Lyon Acoustics Center, ANR-10-LABX-0060).

Data availability Data are available through the first and the corresponding authors upon reasonable request.

Declarations

Conflict of interest The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Compliance with ethical standard This clinical prospective study received ethical committee approval (Ile de France VIII, France, No. 19.09.02.63029 RIPH3) and was recorded in clinicaltrials.gov (NCT04106063).

References

- Simon F, Roman S, Truy E et al (2019) Guidelines (short version) of the French Society of Otorhinolaryngology (SFORL) on pediatric cochlear implant indications. *Eur Ann Otorhinolaryngol Head Neck Dis* 136:385–391. <https://doi.org/10.1016/j.anorl.2019.05.018>
- Gordon KA, Cushing SL, Easwar V et al (2017) Binaural integration: a challenge to overcome for children with hearing loss. *Curr Opin Otolaryngol Head Neck Surg* 25:514–519. <https://doi.org/10.1097/MOO.0000000000000413>
- Warner-Czyz AD, Loy B, Tobey EA et al (2011) Health-related quality of life in children and adolescents who use cochlear implants. *Int J Pediatr Otorhinolaryngol* 75:95–105. <https://doi.org/10.1016/j.ijporl.2010.10.018>
- Gatehouse S, Noble W (2004) The Speech, Spatial and Qualities of Hearing Scale (SSQ). *Int J Audiol* 43:85–99
- Akeroyd MA, Guy FH, Harrison DL, Suller SL (2014) A factor analysis of the SSQ (Speech, Spatial, and Qualities of Hearing Scale). *Int J Audiol* 53:101–114. <https://doi.org/10.3109/14992027.2013.824115>
- Moulin A, Richard C (2016) Sources of variability of Speech, Spatial, and Qualities of Hearing Scale (SSQ) scores in normal-hearing and hearing-impaired populations. *Int J Audiol* 55:101–109. <https://doi.org/10.3109/14992027.2015.1104734>
- Moulin A, Pautz A, Richard C (2015) Validation of a French translation of the Speech, Spatial, and Qualities of Hearing Scale (SSQ) and comparison with other language versions. *Int J Audiol* 54:889–898. <https://doi.org/10.3109/14992027.2015.1054040>
- Noble W, Jensen NS, Naylor G et al (2013) A short form of the Speech, Spatial and Qualities of Hearing Scale suitable for clinical use: the SSQ12. *Int J Audiol* 52:409–412. <https://doi.org/10.3109/14992027.2013.781278>
- Moulin A, Vergne J, Gallego S, Micheyl C (2019) A new Speech, Spatial, and Qualities of Hearing Scale short-form: factor, cluster, and comparative analyses. *Ear Hear* 40:938–950. <https://doi.org/10.1097/AUD.0000000000000675>
- Coombes L, Bristowe K, Ellis-Smith C et al (2021) Enhancing validity, reliability and participation in self-reported health outcome measurement for children and young people: a systematic review of recall period, response scale format, and administration modality. *Qual Life Res* 30:1803–1832. <https://doi.org/10.1007/s11136-021-02814-4>
- Eiser C, Morse R (2001) Quality-of-life measures in chronic diseases of childhood. *Health Technol Assess Winch Engl* 5:1–157. <https://doi.org/10.3310/hta5040>
- Matza LS, Patrick DL, Riley AW et al (2013) Pediatric patient-reported outcome instruments for research to support medical product labeling: report of the ISPOR PRO good research practices for the assessment of children and adolescents task force. *Value Health* 16:461–479. <https://doi.org/10.1016/j.jval.2013.04.004>
- Galvin KL, Noble W (2013) Adaptation of the speech, spatial, and qualities of hearing scale for use with children, parents, and teachers. *Cochlear Implants Int* 14:135–141. <https://doi.org/10.1179/1754762812Y.0000000014>
- Coudert A, Gaveau V, Gatel J et al (2021) Spatial hearing difficulties in reaching space in bilateral cochlear implant children improve with head movements. *Ear Hear* 43:192–205. <https://doi.org/10.1097/AUD.0000000000001090>
- Rauch A-K, Arndt S, Aschendorff A et al (2021) Long-term results of cochlear implantation in children with congenital single-sided deafness. *Eur Arch Oto-Rhino-Laryngol* 278:3245–3255. <https://doi.org/10.1007/s00405-020-06409-6>
- Acquadro C, Conway K, Hareendran A et al (2008) Literature review of methods to translate health-related quality of life questionnaires for use in multinational clinical trials. *Value Health* 11:509–521. <https://doi.org/10.1111/j.1524-4733.2007.00292.x>
- Epstein J, Santo RM, Guillemin F (2015) A review of guidelines for cross-cultural adaptation of questionnaires could not bring out a consensus. *J Clin Epidemiol* 68:435–441. <https://doi.org/10.1016/j.jclinepi.2014.11.021>
- Boateng GO, Neilands TB, Frongillo EA et al (2018) Best practices for developing and validating scales for health, social, and behavioral research: a primer. *Front Public Health* 6:149. <https://doi.org/10.3389/fpubh.2018.00149>
- Smith GT, McCarthy DM, Anderson KG (2000) On the sins of short-form development. *Psychol Assess* 12:102–111. <https://doi.org/10.1037/1040-3590.12.1.102>
- Widaman KF, Little TD, Preacher KJ, Sawalani GM (2011) On creating and using short forms of scales in secondary research. In: Trzesniewski KH, Donnellan MB, Lucas RE (eds) *Secondary data analysis: an introduction for psychologists*. American Psychological Association, Washington, pp 39–61
- Loewenthal K, Lewis CA (2015) *An introduction to psychological tests and scales*, 2nd edn. Psychology Press, London
- Gatehouse S, Akeroyd M (2006) Two-eared listening in dynamic situations: audición con dos oídos en situaciones dinámicas. *Int J Audiol* 45:120–124. <https://doi.org/10.1080/14992020600783103>
- Banh J, Singh G, Pichora-Fuller MK (2012) Age affects responses on the Speech, Spatial, and Qualities of Hearing Scale (SSQ) by

- adults with minimal audiometric loss. *J Am Acad Audiol* 23:81–91. <https://doi.org/10.3766/jaaa.23.2.2>. (quiz 139–140)
24. Sijtsma K (2009) On the use, the misuse, and the very limited usefulness of Cronbach's alpha. *Psychometrika* 74:107–120. <https://doi.org/10.1007/s11336-008-9101-0>
 25. Kass RA, Tinsley HEA (1979) Factor analysis. *J Leis Res* 11:120–138. <https://doi.org/10.1080/00222216.1979.11969385>
 26. Gordon KA, Wong DDE, Valero J et al (2011) Use it or lose it? Lessons learned from the developing brains of children who are deaf and use cochlear implants to hear. *Brain Topogr* 24:204–219. <https://doi.org/10.1007/s10548-011-0181-2>
 27. Gordon KA, Wong DDE, Papsin BC (2013) Bilateral input protects the cortex from unilaterally-driven reorganization in children who are deaf. *Brain J Neurol* 136:1609–1625. <https://doi.org/10.1093/brain/awt052>
 28. Easwar V, Sanfilippo J, Papsin B, Gordon K (2018) Impact of Consistency in daily device use on speech perception abilities in children with cochlear implants: datalogging evidence. *J Am Acad Audiol* 29:835–846. <https://doi.org/10.3766/jaaa.17051>
 29. Kral A, Tillein J, Heid S et al (2005) Postnatal cortical development in congenital auditory deprivation. *Cereb Cortex N Y N* 15:552–562. <https://doi.org/10.1093/cercor/bhh156>
 30. Trumpp NM, Kiefer M (2018) Functional reorganization of the conceptual brain system after deafness in early childhood. *PLoS ONE* 13:e0198894. <https://doi.org/10.1371/journal.pone.0198894>
 31. Ginzburg J, Moulin A, Fornoni L et al (2021) Development of auditory cognition in 5- to 10-year-old children: focus on musical and verbal short-term memory. *Dev Sci*. <https://doi.org/10.1111/desc.13188>
 32. Galvin KL, Mok M, Dowell RC (2007) Perceptual benefit and functional outcomes for children using sequential bilateral cochlear implants. *Ear Hear* 28:470–482. <https://doi.org/10.1097/AUD.0b013e31806dc194>
 33. Beijen J-W, Snik AFM, Mylanus EAM (2007) Sound localization ability of young children with bilateral cochlear implants. *Otol Neurotol* 28:479–485. <https://doi.org/10.1097/MAO.0b013e3180430179>
 34. Lovett RES, Kitterick PT, Hewitt CE, Summerfield AQ (2010) Bilateral or unilateral cochlear implantation for deaf children: an observational study. *Arch Dis Child* 95:107–112. <https://doi.org/10.1136/adc.2009.160325>
 35. Sparreboom M, Snik AFM, Mylanus EAM (2012) Sequential bilateral cochlear implantation in children: quality of life. *Arch Otolaryngol Head Neck Surg* 138:134–141. <https://doi.org/10.1001/archoto.2011.229>
 36. Galvin KL, Mok M (2016) Everyday listening performance of children before and after receiving a second cochlear implant: results using the parent version of the Speech, Spatial, and Qualities of Hearing Scale. *Ear Hear* 37:93–102. <https://doi.org/10.1097/AUD.0000000000000226>
 37. Gargula S, Simon F, C  lerier C et al (2022) French adaptation and validation of the Speech, Spatial and Qualities of Hearing Scale for parents (SSQ-P) and for children (SSQ-Ch). *Int J Audiol*. <https://doi.org/10.1080/14992027.2022.2084461>

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

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.



Training spatial hearing in unilateral cochlear implant users through reaching to sounds in virtual reality

Chiara Valzolgher^{1,2} · Sabrina Bouzaid² · Solene Grenouillet² · Julie Gatel³ · Laura Ratenet³ · Francesca Murenu² · Grégoire Verdelet^{2,4} · Romeo Salemmé^{2,4} · Valérie Gaveau² · Aurélie Coudert³ · Ruben Hermann³ · Eric Truy³ · Alessandro Farnè^{2,4} · Francesco Pavani^{1,2,5}

Received: 7 October 2022 / Accepted: 13 February 2023 / Published online: 11 March 2023
© The Author(s) 2023

Abstract

Background and Purpose Use of unilateral cochlear implant (UCI) is associated with limited spatial hearing skills. Evidence that training these abilities in UCI user is possible remains limited. In this study, we assessed whether a Spatial training based on hand-reaching to sounds performed in virtual reality improves spatial hearing abilities in UCI users

Methods Using a crossover randomized clinical trial, we compared the effects of a Spatial training protocol with those of a Non-Spatial control training. We tested 17 UCI users in a head-pointing to sound task and in an audio-visual attention orienting task, before and after each training. Study is recorded in clinicaltrials.gov (NCT04183348).

Results During the Spatial VR training, sound localization errors in azimuth decreased. Moreover, when comparing head-pointing to sounds before vs. after training, localization errors decreased after the Spatial more than the control training. No training effects emerged in the audio-visual attention orienting task.

Conclusions Our results showed that sound localization in UCI users improves during a Spatial training, with benefits that extend also to a non-trained sound localization task (generalization). These findings have potentials for novel rehabilitation procedures in clinical contexts.

Keywords Spatial hearing · Cochlear implant · VR training · Virtual reality · Reaching · Active listening · Head movements

Introduction

In case of neurosensory deafness, standard interventions often comprise the application of cochlear implants (CI). Although this surgery is indicated for people with bilateral hearing loss, many patients receive only one CI [1, 2]. Using only one CI and listening asymmetrically leads to difficulties

in sound localization. Sound localization is poorer with unilateral rather than bilateral CI [3]. Similarly, switching-off one implant in bilateral CI users (BCI) compromises sound localization in the horizontal dimension [4, 5]. Spatial hearing difficulties in unilateral CI users (UCI) have been attributed to the reduced availability of auditory cues. The CI alters auditory cues due to its restricted spectro-temporal processing [6]. In addition, it can modify sound intensity through automatic gain control (AGC) or alter auditory cues through noise reduction strategies or through filters emphasizing higher frequencies [10]. Wearing one CI also minimizes binaural inputs, which are crucial to localize sounds along the horizontal dimension [10–13]. Even when binaural hearing experience is only reduced and not absent, as in the case of bimodal stimulation (e.g., a cochlear implant in one ear and a contralateral hearing aid in the other), sound localization is perturbed [14]. Binaural cues are distorted and compromised by technology difference between device (e.g., the device delay mismatch [15]), while monaural spectral

✉ Chiara Valzolgher
chiara.valzolgher@unitn.it

¹ Center for Mind/Brain Sciences (CIMEC), University of Trento, Corso Bettini, 31 Rovereto, Trento, Italy

² Integrative, Multisensory, Perception, Action and Cognition Team (IMPACT), Lyon Neuroscience Research Center, University of Lyon 1, Lyon, France

³ Hospices Civils de Lyon, Lyon, France

⁴ Neuroimmersion, Lyon, France

⁵ Centro Interuniversitario di Ricerca “Cognizione, Linguaggio e Sordità” (CIRCLeS), Trento, Italy

pinna cues are poor or not preserved by hearing devices' features (i.e., microphone behind the ear) [16].

In this context of impoverished auditory cues, can CI users improve their sound localization skills? In people with normal hearing listening with one ear plugged, sound localization abilities can be trained [17–19]. Pioneering results suggest that this may also be possible in UCI [20] and BCI users [21]. Recently, we showed that a multisensory-motor training can rapidly change sound localization skills in BCI users [22, 23]. Taking advantage of virtual reality (VR) technologies, we promoted active interactions with sound sources through hand-reaching and head movements. We found that such an active exploration of the acoustic environment enhanced sound localization performance in normal hearing adults with one plugged ear [24] as well as BCI users [22]. These findings are in line with recent studies, which showed that CI users and people with hearing deficits can improve their sound localisation ability when head movements are allowed [25, 26]. Most importantly, we reported that training-related benefits can generalize, extending to a sound localization task in which both the response modality and stimulation positions were novel compared to the trained ones [22, 24].

In the present study, we leveraged such VR training protocol based on active interactions with the auditory scene. To test the efficacy of this training in 17 UCI users, we contrasted this Spatial training with a control procedure that did not entail processing of spatial features of the sound (i.e., the Non-Spatial training). Crucially, we compared these two VR trainings in a crossover experimental design, which allow us to test the effect of both training paradigms on each participant. Before and after each training paradigm, we tested participants in a head-pointing to sound localization task, which entails different sound positions and requires localizing sounds using a different effector (head instead of hand). In addition, to probe for training benefits when implicit sound localization is required, we tested participants in an audio-visual attention orienting task, in which they were asked to judge the elevation of a visual stimulus while listening a sudden sound.

Methods

Participants

Twenty UCI participants were recruited to participate in the study. Sample size was based on two previous experiments addressing a similar research question with an identical experimental design, but with different populations (normal hearing: [24]; bilateral CI users: [22]). Three participants were excluded from the analyses (one did not complete the second visit, one abandoned after wearing the Head Mounted Display (HMD), one did not fully match the inclusion criteria; mean age for the

included participants was 45.8 years, $SD=16.4$; 8 males, 13 right-handed). Three participants asked to interrupt the Spatial VR training because of fatigue (participants 5, 15 and 20 performed 104, 104 and 131 trials out of 156, respectively).

All participants were recruited at the ORL department of the civil hospital Edouard Herriot (HEH) in Lyon (France), and tested in a dedicated room inside the HEH premises. All had normal or corrected-to-normal vision and reported no movement or vestibular deficit, nor neurological or psychiatric history. Anamnestic and clinical data for individual UCI participants are provided in Table 1. During the experiment, participants used their daily sound processor settings (see Table 2 for details about CI model, processor strategy and microphone settings) and 10 of them wore hearing aid on the non-implanted ear. We let each participant perform the task with or without their hearing device in the ear contralateral to the CI, because we aimed to test their sound localization ability in the context of the acoustic stimulation they usually experience. Accordingly, in Table 2 we reported the pure tone average (PTA) threshold in the ear contralateral to the implant, as measured in the condition in which participants performed the experiment: i.e., with or without hearing aid. We calculated them by computing the average between the thresholds available in clinical record for each subject for 250, 500, 1000, 2000, 4000, 8000 Hz. The study was approved by a national Ethical Committee (Ile de France X, N° ID RCB 2019-A02293-54) and recorded in clinicaltrials.gov (NCT04183348). Before starting the experiment, each participant signed an informed consent.

Study design

The entire experiment was conducted inside VR environment. Participants wore a HMD (resolution: 1080×1200 px, Field Of View (FOV): 110° , Refresh rate: 90 Hz) that produced an immersive VR experience: participants always saw a reproduction of the room in which they were located. Importantly, the VR also allowed continuous tracking of their head posture and movements. All sounds were delivered from a real speaker, tracked in 3D space and moved by the experimenter's hand to pre-determined positions within the VR environment (identical to the methods adopted in [22–24] for use of this VR approach in CI users). Use of an actual sound source in the environment was motivated by the difficulty of creating replicable auditory virtual sounds for people using CIs or hearing aids (although extensive efforts have been made to render virtual acoustic scenes also for these hearing assisted populations [28–30]). The immersive VR gave full control over the multisensory cues of testing environment and sound source position, and permitted to provide audio-visual feedback in case of errors.

Participants performed each of the two VR training (Spatial and Non-Spatial) in a within-subject crossover design, in two separate experimental sessions (washout interval was at least 15 days, training order was counterbalanced across

Table 1 Anamnestic and clinical characteristics of CI participants

ID	Gender	Age	Etiology of deafness	Age at deaf diagnosis (y; m)	Age at first hearing aid (y; m)	Age at implantation (right ear) (y; m)	Age at implantation (left ear) (y; m)	Years with one CI
uCI 01	F	35	Unknown	7	9	33	–	2
uCI 02	M	63	Otosclerosis	24	24	60	–	3
uCI 03	F	61	Unknown	32	32	55	–	6
uCI 04	M	48	Radiotherapy/chemotherapy	37	37	–	44	4
uCI 05	M	60	Aminoglycoside treatment	54	55	56	–	4
uCI 06	F	46	Unknown	5	5	–	31	15
uCI 07	M	64	Otosclerosis	37	43	58	–	6
uCI 08	F	57	Genetic	6	17	50	–	7
uCI 09	F	24	Genetic	1	1	7	–	17
uCI 11	F	24	Unknown	1	1	4	–	20
uCI 12	F	21	Meningitis	0;3	1	–	12	9
uCI 14	F	32	Unknown	16	16	–	30	2
uCI 15	M	24	Unknown	0;4	1	2	–	22
uCI 16	F	60	Unknown	30	41	–	57	3
uCI 17	M	38	Genetic	1	2	33	–	5
uCI 18	M	56	Neonatal resuscitation	0	50	51	–	5
uCI 20	M	66	Meniere's disease	51	57	58	–	8

participants; see Fig. 1). Before and after each VR training session, participants completed testing phases that comprised two different auditory tasks: a head-pointing localization task and an audio-visual attention orienting task (the only task conducted outside VR).

Procedures

Testing phases

Head-pointing sound localization task In each trial, a single auditory target (3 s white noise burst) was presented from 8 possible pre-determined positions (5 repetitions each, resulting in 40 trials in each testing phase). The 8 positions were placed at 55 cm from the center of the subject's head and they varied along the azimuth dimension ($\pm 22.5^\circ$ and $\pm 67.5^\circ$ with respect to the midsagittal plane) and vertical dimension (5° and -15° with respect to the plane passing through the ears). The variation along the vertical dimension was introduced only to increase variability in the task, and we did not expect training-related changes in this dimension in which sound localization relies on monaural spectra cues. For this reason, we did not analyze performance along the vertical dimension or have hypothesis about participants errors along the vertical plane. While listening the sound, participants were not informed about the pre-determined target positions and were immersed in an empty virtual room (identical size to the real room, i.e., 3.6 m \times 3.9 m, height 2.7 m). Participants were instructed to

point with their head toward the perceived sound position, as soon as the sound finished, and validate their response using the VR controller they held in their right hand (Fig. 1B). In this specific task, the speaker was always invisible in the VR environment. Notably, while initial posture was identical for all participants and trials, head movements were unrestrained from target onset to the response. The task lasted approximately 15 min.

Audio-visual cueing task This task aimed to assess to what extent lateralized sound could capture the participant's audio-visual attention. The task was performed outside VR in the same room of the sound localization task, with participants sat at a desk in front of a computer monitor flanked by speakers. In each trial, a visual disk appeared above or below eye-level ($\pm 1.15^\circ$), on the left or right side (128 trials overall, equiprobable across the four possible positions). Participants were instructed to discriminate the vertical position of the disk as fast and accurately as possible, using up/down arrows keys on the keyboard (Fig. 1B). Each disk was paired with a task-irrelevant sound delivered from one of two loudspeakers flanking the screen. The sound was either delivered on the same side as the visual disk (spatially congruent trials) or from the opposite side (spatially incongruent trials), with equal probability. In normal hearing participants, this procedure results in automatic audio-visual orienting of selective attention, i.e., participants are faster and more accurate when responding to visual targets appearing on the same side as the preceding sound [31]. All

Table 2 Extended information about participants' cochlear implants

ID	Right				Left					
	Brand	Processor	Strategy	Microphone directionality	Number of active/total electrodes	Brand	Processor	Strategy	Microphone directionality	Number of active/total electrodes
uCI 01	Medel	Rondo	Fs4	NA	9/12				<i>No hearing aid [Hearing Threshold unaided 120 dB HL]</i>	
uCI 02	Oticon medical	Neuro 2	Crysalis	Omnidirectional	18/20				<i>Hearing aid [Hearing Threshold aided 46 dB HL]</i>	
uCI 03	Medel	Concerto	Fs4P	Omnidirectional	10/12				<i>Hearing aid [Hearing Threshold aided 39 dB HL]</i>	
uCI 04	<i>Hearing aid [Hearing Threshold aided 37 dB HL]</i>									
uCI 05	Cochlear	Cp910	Ace	Scan	22/22				Cochlear Cp910 Ace Scan	20/22
uCI 06	<i>Hearing aid [Hearing Threshold aided 48 dB HL]</i>								<i>No hearing aid [Hearing Threshold unaided 62 dB HL]</i>	
uCI 07	Cochlear	Cp1000	Speak	Omnidirectional	18/22				Cochlear Cp1000 Ace Omnidirectional	22/22
uCI 08	Medel	Rondo1	Fs4P	Omnidirectional	12/12				<i>Hearing aid [Hearing Threshold aided 45 dB HL]</i>	
uCI 09	Advanced bionics	Naida ci q70	Hires optima.s	Omnidirectional	16/16				<i>No hearing aid [Hearing Threshold unaided 71 dB HL]</i>	
uCI 11	Cochlear	Co910	Ace	Scan	20/22				<i>Hearing aid [Hearing Threshold aided 53 dB HL]</i>	
uCI 12	<i>Hearing aid [Hearing Threshold aided 35 dB HL]</i>								<i>No hearing aid [Hearing Threshold unaided 90 dB HL]</i>	
uCI 14	<i>Hearing aid [Hearing Threshold aided 40 dB HL]</i>								Cochlear Cp1000 Ace Omnidirectional	22/22
uCI 15	Cochlear	Cp1000	Speak	Omnidirectional	20/22				Medel Sonnet FS4 Natural	12/12
uCI 16	<i>Hearing aid [Hearing Threshold aided 69 dB HL]</i>								<i>No hearing aid [Hearing Threshold unaided 100 dB HL]</i>	
uCI 17	Advanced bionics	Naida ci q70	Hires optima.s	Omnidirectional	16/16				Medel Rondo 2 FS4 Omnidirectional	12/12
uCI 18	Cochlear	Cp1000	Ace	Scan	22/22				<i>Hearing aid [Hearing Threshold aided 68 dB HL]</i>	
uCI 20	Cochlear	Cp1000	Ace	Scan	22/22				<i>No hearing aid [Hearing Threshold unaided 100 dB HL]</i>	
									<i>No hearing aid [Hearing Threshold unaided 57 dB HL]</i>	

Note that the thresholds reported refer to PTA in the condition in which participant performed the experiment (with or without hearing aid)

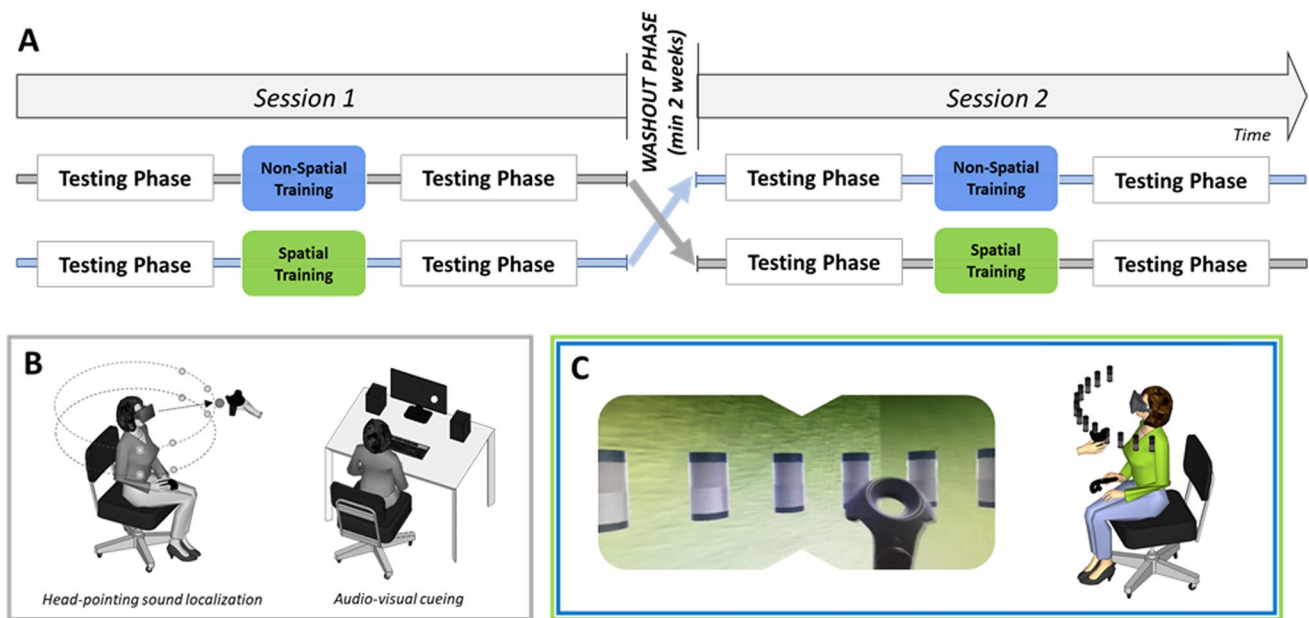


Fig. 1 Experimental procedure and setting. **A** Schematic description of the overall crossover design. Each session (Session 1 and Session 2) comprised two testing phases, separated by a training task: Non-Spatial VR in blue and Spatial VR in green. **B** Testing phases. Left: schematic representation of the participant wearing the HMD and holding the VR controller during the head-pointing sound localization task. The grey circles represent the 8 possible positions in which the real loudspeaker could be placed (shown here only for illustration purposes, as no visual cue to sound position was available in the VR environment). They were located 55 cm from the center of the subject's head, at different azimuth ($\pm 22.5^\circ$ and $\pm 67.5^\circ$ with

respect to the midsagittal plane) and vertical positions (5° and -15° with respect to the plane passing through the ears). Note that the real speaker was never visible in the VR environment. Right: schematic representation of the setting for the audio-visual cueing task (conducted entirely outside VR). **C** Training phase. Left: close-up of the scene as visible inside the HMD from participant's perspective. The virtual scenario comprised a room, 13 speakers and the VR controller held in participants' hands. Right: schematic representation of the participant wearing the HMD and holding the VR controller during the training tasks

head movements were restrained by a forehead and chin-rest. The task lasted approximately 10 min.

VR training tasks

Participants were immersed in the same virtual room as the head-pointing sound localization task, but saw 13 virtual loudspeakers spanning $\pm 72^\circ$ in front space (see Fig. 1C). In each trial, a sound was emitted by a real speaker moved by the experimenter, as in the head-pointing sound localization task (12 repetitions per loudspeakers, total 156 trials). Half of the sounds were amplitude modulated at 2 Hz, the remaining half at 3 Hz. Hence, irrespective of the VR training task (Spatial or Non-Spatial), the stimulation changed unpredictably in location and amplitude modulation on a trial-by-trial basis.

Spatial VR training Participants were instructed to reach the speaker emitting the sound using the VR controller they held in their right hand. The sound lasted until the participant reached and 'touched' the correct speaker. If they reached the wrong speaker, the correct loudspeaker started to display concentric red beams that expanded from the

correct position to reach, and the sound continued until the correct location was finally reached (a video that illustrates the training tasks is available in 22, <http://links.lww.com/EANDH/B44>).

Non-Spatial VR training Participants were instructed to identify the amplitude modulation in the target sound, and indicate their discrimination through a reaching movement using VR controller. For fast amplitude-modulated sounds, participants reached in front of them, aiming to touch the invisible virtual button placed above the central speaker. For slow amplitude-modulated sounds, participants reached instead the invisible virtual button placed below the same central speaker. As in the Spatial Training feedback procedure, the sound stopped only when a correct response was provided. If they reached the wrong button, a visual feedback was displayed and the sound continued until the correct button was finally touched. In both trainings, the feeling of touch was induced by making the controller vibrate as soon as it collided with objects (speakers or invisible buttons).

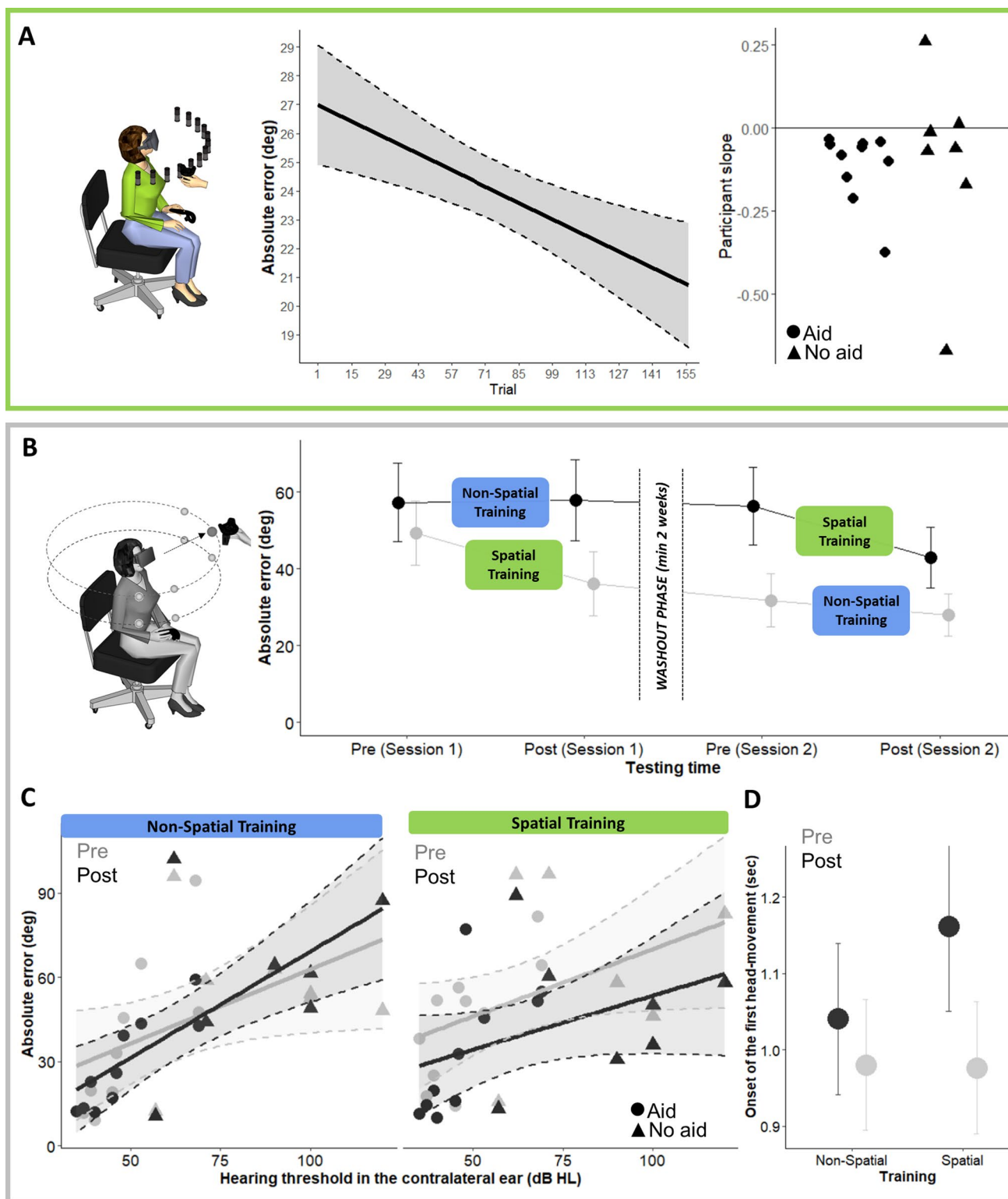


Fig. 2 Sound localization performance. **A** Absolute error along azimuth, as a function of trial in the Spatial training. Linear regression (solid line) with 95% confidence intervals (dashed lines). To the left, slope for each participant extracted from the LME model used in the analysis. **B** Absolute localization across the four testing sessions of the experimental design, separately for participants who completed the Spatial training on session 1 (grey) or session 2 (black line). Error bars represent standard errors. **C** Absolute localization error along azimuth

dimension as a function of training (Spatial: right and Non-Spatial: left), phase (Pre: grey and Post: black) and hearing threshold in the contralateral ear (x axis). **D** Onset of the first head-movements in seconds as a function of phase (pre-training in black and post-training in grey) and trainings. Error bars represent standard errors. In **A** and **C**, circles represent participants who wore hearing aid in the contralateral ear ($N=10$) and triangles who did not have hearing aid ($N=7$)

Statistical analysis

Linear mixed-effect modeling was used for all statistical analyses. Statistical analyses were run using R (version 1.0.143). For the linear mixed-effect (LME) model, we used the R-packages emmean, lme4, lmerTest in R Studio [32, 33]. The R-package car was used to obtain deviance tables from the LME models. For further details about “Materials and method”, see [22] and Supplementary Results for details on each of the analyses described below.

To study head movements, we extracted three dependent variables: number of head rotations, head-rotation extent and head-rotation bias [18]. All the variables concern head rotations around the vertical axis. To calculate the number of head movements, we counted all the detected peaks of velocity in the head trace expect for movements smaller than 2° , which were removed to exclude movements which are not indicators of spontaneous head intentional movements and not related to the task (i.e., micro-postural movements). To calculate head-rotation extent, we sum the absolute value of the rightward and leftward head rotation around the vertical axis extremity, while to calculated head-rotation bias we computed the signed sum of these two values.

Results

VR training

Performance

The spatial discrepancy between the stimulated and the reached speaker (i.e., absolute localization error in azimuth, calculated as difference in absolute value between speaker and response position in azimuth in degrees) was on average 24.0 degrees ($SD = 14.0$), with a numerical (but not-significant) bias toward the side contralateral to the CI (-2.6° , $SD = 7.6$; t -test against zero: $t(16) = -1.40$, $p = 0.19$). Importantly, absolute localization errors decreased across trials ($X^2(1) = 4.37$, $p = 0.04$), proving that participants improved localization abilities during the Spatial training (Table S1). Figure 2A shows changes in sound localization performance during the Spatial VR training across successive trials. The left panel depicts the overall change across participants; the right panel shows the value of the slope of the regression line for each individual participant, grouped as a function of whether they used a hearing aid in the ear contralateral to the CI or not. For all participants except two the slope was negative, i.e., their performance improved across trials.

We analyzed the influence of hearing asymmetries in hearing thresholds (PTA) between the implanted and

non-implanted ear on errors during the Spatial training. We found that the larger the asymmetry (i.e., the higher hearing thresholds in the contralateral ear), the larger the localization error ($X^2(1) = 75.55$, $p < 0.001$). Moreover, the bias toward the contralateral ear emerged as a function of hearing asymmetry ($X^2(1) = 4.26$, $p = 0.04$). Yet, individual asymmetries in hearing thresholds did not modulate the amount of improvement across time during the Spatial training (all $ps > 0.11$; see Table S2).

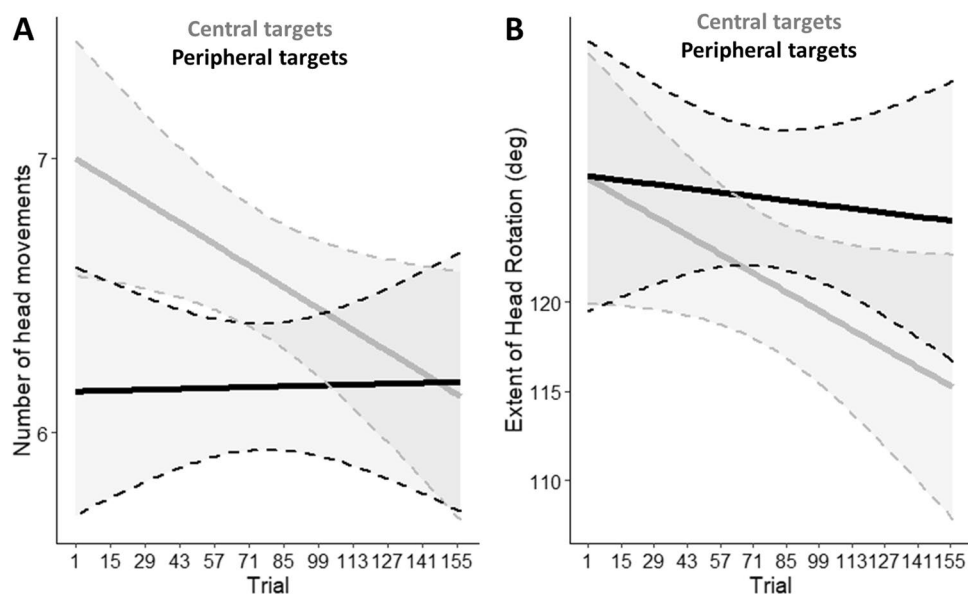
Performance in the Non-Spatial training was near ceiling for all participants (mean number of errors = 1.5%). During the Non-Spatial training, participants were also faster in completing the trial compared to the Spatial training (Non-Spatial: mean \pm SD = 2.5 ± 2.1 s; Spatial training: mean \pm SD = 16.1 ± 10.2 s; $t(16) = 11.41$, $p < 0.001$ on paired t -test).

Head movements

Head rotations were overall more frequent in the Spatial (6.51 ± 3.15) compared to the Non-Spatial VR training (1.5 ± 1.0 , $W = 152$, $p < 0.001$ on Wilcoxon signed rank test). Likewise, the extent of head rotation movements was larger during the Spatial ($123.9^\circ \pm 64.0^\circ$) compared to the Non-Spatial VR training ($15.2^\circ \pm 45.2^\circ$, $W = 153$, $p < 0.001$ on Wilcoxon signed rank test). Finally, head rotations were more biased toward the side contralateral to the CI during the Spatial ($-27.5^\circ \pm 35.8^\circ$) compared to the Non-Spatial training ($-0.8^\circ \pm 7.8^\circ$, $W = 17$, $p = 0.003$ on Wilcoxon signed rank test).

During the Spatial VR training, we observed that UCI users adapted their spontaneous head movements as a function of sound eccentricity as training trials progressed. Specifically, number of head rotations ($X^2(1) = 6.22$, $p = 0.01$) and head rotation extent ($X^2(1) = 10.06$, $p = 0.002$) diminished as a function of trial repetition, specifically when sounds were emitted by central sources (shown in grey in Fig. 3). This is compatible with participants requiring progressively fewer head movements and smaller extent of head rotation to identify central targets over time. We also observed that participants head-rotation bias changed as a function of sound side (ipsilateral vs. contralateral with respect to the CI, $X^2(2) = 1125.33$, $p < 0.001$). When the sound was delivered on the same side as the CI, head rotations were markedly biased toward the non-implanted side ($-46.7^\circ \pm 32.6^\circ$), as if participants aimed to exposed their CI to the sound energy. By contrast, this biased exploration was not evident for sounds delivered on the side opposite to the CI ($-4.2^\circ \pm 38.2^\circ$) (see Tables S3 and S4).

Fig. 3 Head rotation during the Spatial training. **A** Number of head movements across trial repetition as a function of sound eccentricity (central positions in grey and peripheral positions in black). **B** Extent of head rotation across trial repetition as a function of sound eccentricity (central positions in grey and peripheral positions in black)



Effects beyond the trained task

Head-pointing to sounds

Performance The Spatial VR training improved performance (i.e., decreased absolute error in azimuth) more than the Non-Spatial training, irrespective of the session in which it was completed (Spatial—pre: $52.6^\circ \pm 26.2^\circ$; post: $39.3^\circ \pm 23.5^\circ$; Non-Spatial training—pre: $43.8^\circ \pm 27.4^\circ$; post: $42.0^\circ \pm 27.6^\circ$, $X^2(1) = 23.36$, $p < 0.001$; see Fig. 2B). Interestingly, participants who performed the Spatial VR training in the first session ($N = 9$) maintained the improvement after the 2-week washout: errors in the pre-training phase were smaller on session 2 compared to session 1 ($X^2(1) = 94.86$, $p < 0.001$).

As documented above, individual asymmetries in hearing thresholds between the implanted and non-implanted ear influenced performance: the higher the threshold in the contralateral ear, the higher the error ($X^2(1) = 462.29$, $p < 0.001$). The effect of Spatial training was higher for participants who had higher threshold, as shown in Fig. 2C ($X^2(1) = 59.30$, $p < 0.001$) (see Table S5). Participants who wore a hearing aid at ear contralateral to the CI are indicated by circles in Fig. 2.

Trainings also influenced the response bias in sound localization (i.e., the signed error). In the pre-training session, no overall bias toward the side contralateral to the CI was measured (-1.1° , $SD = 39.2$; t -test against zero: $t(16) = 0.12$, $p = 0.91$). Yet, when the response bias was studied as a function of hearing asymmetry we found that, after each of the training, the participants' responses changed. The Non-Spatial training increased the bias toward the side of the CI, especially for participants with higher hearing

asymmetry ($X^2(1) = 4.18$, $p = 0.04$). Conversely, the Spatial training decreased the bias for all participants, especially for participants with higher level of hearing asymmetry ($X^2(1) = 24.97$, $p < 0.001$, see Table S6 for further details about the analysis).

To further examine the effect of training on head-pointing to sounds, we analyzed the direction of the first head-rotation in each trial. This measure captures the immediate orienting response toward the sound. We found that participants discriminated the side of sounds source (ipsilateral: 24.09 ± 39.42 ; contralateral: -17.18 ± 47.19 , $X^2(1) = 58.84$, $p < 0.001$). Their accuracy in discriminating sources' side increased selectively after the Spatial training as compared to the Non-Spatial one ($X^2(1) = 3.09$, $p = 0.05$). Another head-movement variable worth considering is the onset of the first head movement of correct responses, which has been used as indicator of the ability to discriminate the side of the sound direction [22, 34]. Irrespective of VR training type, onset of the first head movement decreased after training ($X^2(1) = 7.00$, $p = 0.008$). Yet, this reduction was more pronounced after the Spatial (before: 1.16 ± 0.46 ; after: 0.98 ± 0.35 ; $t = 7.36$, $p < 0.001$), as compared to the Non-Spatial training (before: 1.04 ± 0.41 ; after: 0.98 ± 0.35 ; $t = 2.65$, $p = 0.008$; $X^2(1) = 11.92$, $p < 0.001$). This reduction was more pronounced for participants who have higher asymmetry ($X^2(1) = 22.08$, $p < 0.001$) (Fig. 2D, see Table S7).

Head movements In order to describe changes in head movement after training, we measured number of head rotations, head-rotation extent, head-rotation bias and direction of the first head movements during the sound (see Analysis for a description of these variables). We report here the

main findings, but see Supplementary Materials for further details (Table S8).

Participants changed their head-related behavior after each of the training. In the post-training session, they increased the number of movements (Pre: 1.90 ± 0.65 ; Post: 2.06 ± 0.52 , $X^2(1) = 5.16$, $p = 0.02$) and increased head-rotation extent (Pre: 99.21 ± 48.77 ; Post: 101.35 ± 48.39 , $X^2(1) = 3.78$, $p = 0.05$). Importantly, only after the VR Spatial training they turned their heads toward the contralateral side of the implant to bring the implanted ear toward the sounds: head-bias toward the contralateral space increased more after the Spatial (Pre: -8.84 ± 50.54 ; Post: -30.11 ± 51.05 , $t = 5.42$, $p < 0.001$) compared to the Non-Spatial training (Pre: -15.07 ± 51.05 ; Post: -18.35 ± 54.20 , $t = 0.84$, $p = 0.40$, $X^2(1) = 10.48$, $p = 0.001$) (Fig. 4).

Head movements also changed as a function of hearing asymmetry. Participants with greater hearing asymmetry increased their head-rotation extent after the Spatial compared to Non-Spatial VR training more than participants with less hearing asymmetry ($X^2(1) = 18.03$, $p < 0.001$). Likewise, they directed their first head movement toward the contralateral space after the Spatial, but not after the Non-Spatial VR training ($X^2(1) = 15.35$, $p < 0.001$).

Audio-visual attention orienting task When people with normal hearing are asked to make a visual discrimination (here on the elevation of the visual target, up vs. down), they are faster when the visual target is preceded by a sound located in the same (congruent) vs. opposite side of the space (incongruent) [35, 36]. This well-documented behavioral effect reveals to what extent a sound can capture auditory and visual attention to its location in space, resulting in perceptual processing benefits across sensory modalities.

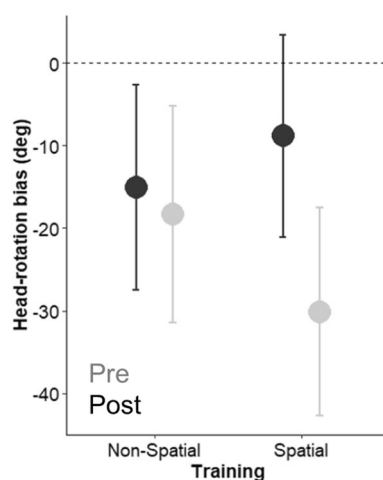


Fig. 4 Head-rotation bias during the Head-pointing to sounds task, as a function of training (Spatial Training and Non-Spatial training) and Phase (Before training in black and post-training in grey). Error bars represent standard errors

Notably, UCI participants tested in the present study did not benefit of the congruency between the sound and the visual target, as revealed by the absence of a congruency effect ($X^2(1) = 1.03$, $p = 0.31$), nor this effect emerged after training ($X^2(1) = 0.41$, $p = 0.52$). The only change between pre- and post-training session was in the overall speed of the response, irrespective of the congruency between the sound and the visual target ($X^2(1) = 9.73$, $p = 0.002$), which is compatible with a practice effect. Intriguingly, the difference in reaction time between congruent and incongruent conditions (often termed audio-visual cueing, [35, 36]) changed as a function of hearing asymmetry: the lower the asymmetry, the higher is the cueing effect. After both training paradigms, audio-visual cueing increased specifically for people who had higher asymmetry ($X^2(1) = 4.49$, $p = 0.03$, see Table S9).

Discussion

We observed that UCI users can improve their sound localization abilities, despite the substantial impoverishment of the available auditory cues. Thus, acoustic space perception improvement is possible also for people using a single CI, at least in the experimental context we have examined. Specifically, we showed that sound localization of UCI users can improve across trials while engaged in a Spatial VR training and that error reduction extended beyond the trained task. Localization errors decreased after training, as compared to before training, and this decrement (about 13°) was greater after the Spatial compared to the Non-Spatial VR training. Further analyses revealed that hearing asymmetry (as described by PTA at the non-implanted ear) modulated training benefit. Generalization effect of Spatial training was more pronounced for participants with higher hearing asymmetry (i.e., higher hearing threshold at the contralateral ear). Finally, the Spatial VR training had no impact on the audio-visual orienting task, which involves the ability to localize sounds sources indirectly. A possible explanation for the lack of this effect is that active listening was prevented during this task as, for experimental reasons, it was performed using a chin-rest. Although this test represents a firm attempt to test the generalization effect, further studies are needed to investigate the transferability of training effects to tasks in which the ability to localize sounds is implicitly involved.

A previous study by Luntz and colleagues already suggested that training UCI spatial hearing skill is possible, but this early report suffered from methodological limitations. They tested only few participants ($N = 9$) and the spatial positions of targets did not vary between the test and training situations [20]. The present study represents a step forward compared to previous literature because a larger number of UCI users were involved and, most importantly, we tested

the effects of training also in a different sound localization task to assess generalization. Furthermore, in the present study we adopted a within-subject experimental design, which gave us the opportunity to directly compare the effects of the experimental training with that of the control training in the exact same participants. We also demonstrated the possibility of improvement using a training of only 156 trials, which is considerably shorter compared to previous studies (four to eight training sessions over a period of 6.5 weeks in [20]; or 8 sessions spanning over 4 weeks as in [23]).

Interestingly, our short training produced effects also in a sound localization task entailing different sound positions, a different response method (i.e., use of the head as pointer instead of reaching sound sources using the hand), and less visual cues available (i.e., potential sound sources invisible during the test). The difference between the trained task and the test task (i.e., head-pointing to sound) is clearly evident also in the different performance achieved by the participants in the two sound localization procedures. While in the trained task, the absolute error for participants in the Spatial training group was 24° on average, in the test task, they started from an average absolute error of 52.6° to achieve a performance of 39.3° at the end of training. This difference is likely the consequence of the different priors about sound position in the two procedures: during training, all possible sound positions were visually identified, whereas during test, no visual cue helps participants to locate the sound sources.

This result highlights the importance of assessing generalization effect when testing the efficacy of training protocols (see also [37]). In addition, it encourages to pursue testing objectives which deviate even more from the training task, in order to fully examine the potentials of the generalization processes (e.g., using words or syllables as stimuli, see [23]). Notably, our crossover experimental design provides initial evidence of relatively long-lasting effects. Participants who performed the Spatial VR training in the first session maintained the improvement after the 2-week washout. This finding corroborates recent evidence in BCI users that also documented a persistent effect after 2 weeks [22], and prompts the implementation of longitudinal experimental designs to test this aspect thoroughly in future studies. Despite these encouraging results, it is important to note that the uCI users we tested still showed a limited sound localization ability (their errors are still around 40° after the training). Hence, it should be considered as a further opportunity for improvement, but not an alternative to other solutions that could improve spatial hearing (e.g., bilateral implantation).

Given the large interindividual variability in terms of hearing experience, we investigated if hearing asymmetry influenced sound localization performance and training effects. We observed an increase in the effectiveness of the Spatial training for UCI users with higher levels of hearing

asymmetry. This finding supports the idea that it is possible to improve localization of sounds even when auditory cues available are primarily monaural intensity cues, and opens the possibility to offer a similar training to people with unilateral hearing loss. Since the larger training effectiveness was documented in individuals who primarily listen monaurally using their CI, rather than individuals with bimodal experience providing binaural cues, this finding suggests that our training primarily changed the way in which participants exploited the available intensity monaural cues (but see [22] for evidence showing the efficacy of the Spatial training BCI users). In the present study, however, the analyses that examined the role of hearing asymmetry were exploratory. The recruitment of participants was not conducted taking this aspect into account, as it was not included in our original research questions, hence we could not fully control this dependent variable (i.e., hearing asymmetry level). For this reason, we investigated asymmetry along a continuum, considering for each participant the hearing threshold at the ear contralateral to the CI. Future studies should address this aspect more systematically, for instance by manipulating the degree of asymmetry of the participants' hearing or by considering separately patients with effective bimodal experience (i.e., CI plus an effective hearing aid) and patients with a clear unilateral hearing loss using a CI. Furthermore, even if it was beyond the scope of the present work, a further element which may play a role when training acoustic space perception is deafness onset. Investigating this aspect remains a key question for future research.

A further contribution of the present work concerns the study of head movements' behavior. First, during training, participants required progressively fewer head movements to perform the task and reduced the extent of their head rotation when responding to central targets. This corroborates the observation of a trial-by-trial improvement, that we described above in terms of progressive reduction of performance errors. Second, after training, head movements changed between the pre- and post-training measurements. When we focused on the first head-movement onset during the head-pointing to sounds test, we observed that the correct direction of the sound was identified faster after the Spatial compared to the Non-Spatial VR training. Third, participants also started to spontaneously implement novel head-movement behaviors after the training. Specifically, they increased the number of movements and explored a larger portion of space with the head. This was particularly evident after the Spatial VR training, hinting at the possibility that they moved the head strategically to bring their CI toward the sounds. This strategy might have favored the extraction and use of monaural intensity variation at the CI—pointing again to an advantage of the Spatial training mostly related to the use of monaural cues available at the unilateral CI. This strategy has been already documented in

previous studies testing people with normal hearing in monaural listening conditions [18, 24]. These findings highlight the importance of moving the head and engaging in active and exploratory listening behavior when aiming to improve sound localization abilities [25, 26, 38, 39] and to foster relearning processes [24]. They also point to the importance of measuring head movements when assessing sound localization skills, and the notion that promoting head-orienting strategies may play a key role in protocols aimed at training sound localization.

Conclusion

Using a novel VR training based on reaching to sounds, audio-visual feedback and free head movements during listening, we documented that training sound localization ability in UCI users is possible. While these observations emerged in laboratory setting, they have direct translational implications for the clinical context because the observed improvements did not result from changes in hearing settings and hearing thresholds of the participants. Instead, they were likely the result of recalibration processes and self-regulatory behavior, triggered by a combination of multisensory feedback and actions directed to sounds (with the hand and the head). In turn, these allowed participants to better exploit the residual auditory cues when processing auditory space.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00405-023-07886-1>.

Acknowledgements We are grateful to all the participants who accepted to participate in this study. C. V. was supported by a grant of the Università Italo-Francese/Université Franco-Italienne, the Zegna Founder's Scholarship and Associazione Amici di Claudio Demattè. F. P., A.F. 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 the IHU CaSaMe ANR-10-UBHU-0003 and ANR 2019CE37 Blind Touch. We thank J.L. Borach and S. Alouche for administrative support, E. Koun for informatics support and Giordana Torresani for graphical support (Figs. 1 and 2).

Author contributions CV, FP, and AF conceptualized and designed the experiment and managed funding acquisition and project administration. CV, SB, SG, and FM performed the experiment. CV and FP analyzed data and wrote the paper. GV and RS worked on software. FP, AF and ET supervised the project and managed funding. LR, JG, AC, and VG helped in recruiting the patients.

Funding Open access funding provided by Università degli Studi di Trento within the CRUI-CARE Agreement.

Data availability Data can be retrieved osf.io/mbshq.

Declarations

Conflict of interest The authors have no conflicts of interest to declare.

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 licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence 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 licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Moore DR, Shannon RV (2009) Beyond cochlear implants: awakening the deafened brain. *Nat Neurosci* 12(6):686–691. <https://doi.org/10.1038/nn.2326>
2. Wilson BS (2019) The remarkable cochlear implant and possibilities for the next large step forward. *Acoustics Today* 15(1):53–61
3. Smulders YE, van Zon A, Stegeman I, Rinia AB, Van Zanten GA, Stokroos RJ et al (2016) Comparison of bilateral and unilateral cochlear implantation in adults: a randomized clinical trial. *JAMA Otolaryngol* 142(3):249–256. <https://doi.org/10.1001/jamaoto.2015.3305>
4. Asp F, Mäki-Torkko E, Karltorp E, Harder H, Hergils L, Eskilsson G, Stenfelt S (2012) Bilateral versus unilateral cochlear implants in children: Speech recognition, sound localization, and parental reports. *Int J Audiol* 51(11):817–832. <https://doi.org/10.3109/14992027.2012.705898>
5. Litovsky RY, Parkinson A, Arcaroli J (2009) Spatial hearing and speech intelligibility in bilateral cochlear implant users. *Ear Hear* 30(4):419–431. <https://doi.org/10.1097/AUD.0b013e3181a165be>
6. Majdak P, Goupell MJ, Laback B (2011) Two-dimensional sound localization in Cochlear Implantees. *Ear Hear* 32(2):198. <https://doi.org/10.1097/AUD.0b013e3181f4dfe9>
7. Dorman MF, Loiselle L, Stohl J, Yost WA, Spahr A, Brown C, Cook S (2014) Interaural level differences and sound source localization for bilateral cochlear implant patients. *Ear Hear* 35(6):633–640. <https://doi.org/10.1097/AUD.000000000000057>
8. Archer-Boyd AW, Carlyon RP (2019) Simulations of the effect of unlinked cochlear-implant automatic gain control and head movement on interaural level differences. *J Acoust Soc Am* 145(3):1389–1400. <https://doi.org/10.1121/1.5093623>
9. Archer-Boyd AW, Carlyon RP (2021) Further simulations of the effect of cochlear-implant pre-processing and head movement on interaural level differences. *J Acoust Soc Am* 150(1):506–525. <https://doi.org/10.1121/10.0005647>
10. Seeber BU, Fastl H (2008) Localization cues with bilateral cochlear implants. *J Acoust Soc Am* 123(2):1030–1042. <https://doi.org/10.1121/1.2821965>
11. Grantham DW, Ashmead DH, Ricketts TA, Labadie RF, Haynes DS (2007) Horizontal-plane localization of noise and speech signals by postlingually deafened adults fitted with bilateral cochlear implants. *Ear Hear* 28(4):524–541. <https://doi.org/10.1097/AUD.0b013e31806dc21a>

12. Grieco-Calub TM, Litovsky RY (2010) Sound localization skills in children who use bilateral cochlear implants and in children with normal acoustic hearing. *Ear Hear* 31(5):645. <https://doi.org/10.1097/AUD.0b013e3181e50a1d>
13. van Hoesel RJ (2004) Exploring the benefits of bilateral cochlear implants. *Audiol Neurotol* 9(4):234–246. <https://doi.org/10.1159/000078393>
14. Seeber B, Baumann U, Fastl H (2004) Localization ability with bimodal hearing aids and bilateral cochlear implants. *J Acoust Soc Am* 116(3):1698–1709. <https://doi.org/10.1121/1.1776192>
15. Zirn S, Angermeier J, Arndt S, Aschendorff A, Wesarg T (2019) Reducing the device delay mismatch can improve sound localization in bimodal cochlear implant/hearing-aid users. *Trends Hearing* 23:2331216519843876. <https://doi.org/10.1177/2331216519843876>
16. Veugen LC, Hendrikse MM, van Wanrooij MM, Agterberg MJ, Chalupper J, Mens LH et al (2016) Horizontal sound localization in cochlear implant users with a contralateral hearing aid. *Hearing Res* 336:72–82. <https://doi.org/10.1016/j.heares.2016.04.008>
17. Valzolgher C, Campus C, Rabini G, Gori M, Pavani F (2020) Updating spatial hearing abilities through multisensory and motor cues. *Cognition* 204:104409. <https://doi.org/10.1016/j.cognition.2020.104409>
18. 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>
19. Strelnikov K, Rosito M, Barone P (2011) Effect of audiovisual training on monaural spatial hearing in horizontal plane. *PLoS ONE* 6(3):e18344. <https://doi.org/10.1371/journal.pone.0018344>
20. Luntz M, Brodsky A, Watad W, Weiss H, Tamir A, Pratt H (2005) Sound localization in patients with unilateral cochlear implants. *Cochlear Implants Int* 6(1):1–9. <https://doi.org/10.1002/cii.15>
21. Tyler RS, Witt SA, Dunn CC, Wang W (2010) Initial development of a spatially separated speech-in-noise and localization training program. *J Am Acad Audiol* 21(06):390–403. <https://doi.org/10.3766/jaaa.21.6.4>
22. Valzolgher C, Gatel J, Bouzaid S, Grenouillet S, Todeschini M, Verdelet G et al (2022) Reaching to sounds improves spatial hearing in bilateral cochlear implant users. *Ear Hear*. <https://doi.org/10.1097/AUD.0000000000001267>
23. Coudert A, Verdelet G, Reilly KT, Truy E, Gaveau V (2022) Intensive training of spatial hearing promotes auditory abilities of bilateral cochlear implant adults: a pilot study. *Ear Hear*. <https://doi.org/10.1097/AUD.0000000000001256>
24. Valzolgher C, Todeschini M, Verdelet G, Gatel J, Salemme R, Gaveau V et al (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>
25. Pastore MT, Natale S, Yost W, Dorman MF (2018) Head movements allow listeners bilaterally implanted with cochlear implants to resolve front-back confusions. *Ear Hear* 39(6):1224. <https://doi.org/10.1097/AUD.0000000000000581>
26. Coudert A, Gaveau V, Gatel J, Verdelet G, Salemme R, Farné A et al (2022) Spatial hearing difficulties in reaching space in bilateral cochlear implant children improve with head movements. *Ear Hear* 43(1):192. <https://doi.org/10.1097/AUD.0000000000001090>
27. Gaveau V, Coudert A, Salemme R, Koun E, Desoche C, Truy E et al (2022) Benefits of active listening during 3D sound localization. *Exp Brain Res*. <https://doi.org/10.1007/s00221-022-06456-x>
28. Fichna S, Biberger T, Seeber BU and Ewert SD (2021) Effect of acoustic scene complexity and visual scene representation on auditory perception in virtual audio-visual environments. In: 2021 Immersive and 3D Audio: from Architecture to Automotive (I3DA) (pp 1–9). IEEE. <https://doi.org/10.1109/I3DA48870.2021.9610916>
29. Sechler SD, Valdes AL, Waechter SM, Simoes-Franklin C, Viani L and Reilly RB (2017) Virtual reality sound localization testing in cochlear implant users. In: 2017 8th International IEEE/EMBS Conference on Neural Engineering (NER) (pp 379–382). IEEE. <https://doi.org/10.1109/NER.2017.8008369>
30. Vickers D, Salorio-Corbetto M, Driver S, Rocca C, Levto Y, Sum K, Parmar B, Dritsakakis G, Dlores JA, Jiang D, Mahon M, Early F, Van Zalk N, Picinali L (2021) Involving children and teenagers with bilateral cochlear implants in the design of the BEARS (both EARS) virtual reality training suite improves personalization. *Front Digital Health*. <https://doi.org/10.3389/FGDGH.2021.759723>
31. Koelewijn T, Bronkhorst A, Theeuwes J (2009) Competition between auditory and visual spatial cues during visual task performance. *Exp Brain Res* 195(4):593–602. <https://doi.org/10.1007/s00221-009-1829-y>
32. Bates D, Maechler M, Bolker BM, Walker S (2015) Fitting linear mixed-effects models using lme4. *J Statist Softw* 67:1–48
33. Fox J and Weisberg S (2020) Using car and effects functions in other functions 3, pp 1–5. <https://mran.revolutionanalytics.com/snap-shot/2020-05-18/web/packages/car/vignettes/embedding.pdf>
34. Isaiyah A, Hartley DE (2015) Can training extend current guidelines for cochlear implant candidacy? *Neural Regen Res* 10(5):718–720. <https://doi.org/10.4103/1673-5374.156964>
35. Pavani F, Venturini M, Baruffaldi F, Artesini L, Bonfioli F, Frau GN, van Zoest W (2017) Spatial and non-spatial multisensory cueing in unilateral cochlear implant users. *Hear Res* 344:24–37. <https://doi.org/10.1016/j.heares.2016.10.025>
36. Spence C, Driver J (1997) Audiovisual links in exogenous covert spatial orienting. *Percept Psychophys* 59(1):1–22. <https://doi.org/10.3758/BF03206843>
37. Mendonça C (2014) A review on auditory space adaptations to altered head-related cues. *Front Neurosci* 8:219. <https://doi.org/10.3389/fnins.2014.00219>
38. Wallach H (1940) The role of head movements and vestibular and visual cues in sound localization. *J Exp Psychol* 27(4):339. <https://doi.org/10.1037/h0054629>
39. Mueller MF, Meisenbacher K, Lai WK, Dillier N (2014) Sound localization with bilateral cochlear implants in noise: How much do head movements contribute to localization? *Cochlear Implants Int* 15(1):36–42. <https://doi.org/10.1179/1754762813Y.0000000040>

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

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

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

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

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

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

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

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

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

(Ear & Hearing 2022;XX:00–00)

INTRODUCTION

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

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

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

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

Copyright © 2022 The Authors. Ear & Hearing is published on behalf of the American Auditory Society, by Wolters Kluwer Health, Inc. This is an open-access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

Bolognini et al. 2005), and many studies have highlighted its benefits in terms of sound localization (Strelnikov et al. 2011) and speech discrimination (e.g., Middelweerd & Plomp 1987; MacLeod & Summerfield 1987; Schwartz et al. 2004).

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

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

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

During the early stages after surgery, auditory rehabilitation focuses on sound detection, with this focus shifting to speech discrimination in quiet and noisy situations several months after surgery. Sound localization, however, is rarely included in routine rehabilitation protocols. Experimental spatial hearing rehabilitation protocols could provide a good starting point

for developing training protocols adapted to clinical practice. Indeed, experimental approaches that exploit multisensory interactions to promote learning suggest that rehabilitation of spatial hearing and spatial attention in CI patients is feasible (Tyler et al. 2010; Nawaz et al. 2014). There are, however, several methodological challenges that must be overcome in order to sufficiently control the proposed training sessions. First, since spatial hearing is 3D, it is essential to work in the whole space surrounding the patient, and not only in front space with azimuth (as is commonly the case, e.g., Tyler et al. 2010; Firszt et al. 2015). To do this, using a traditional setup would require many loudspeakers in front and back space at different elevations, which would make the system very bulky and difficult to use in everyday clinical practice. Furthermore, since spatial hearing is a multisensory process, strongly based on visual information (particularly in hearing impaired patients), control of the visual environment is critical. In most training protocols, participants have visual information about the actual positions of loudspeakers around them, which could bias their sound localization responses (Tyler et al. 2010; Kuk et al. 2014; Nawaz et al. 2014; Firszt et al. 2015).

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

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

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

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

than unisensory (Strelnikov et al. 2011), we hypothesized that training-related improvement would be greater in the group that received audiovisual feedback.

MATERIALS AND METHODS

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

Participants

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

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

Experimental Protocol

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

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

training protocol regardless of their level of performance at inclusion.

Material for Delivery of Auditory and Visual Stimuli

The idea behind the training sessions was to promote spatial hearing rehabilitation by providing participants with trial-by-trial feedback in the form of real sounds and visual information congruent with sound sources, or with visual information only. Visual information was delivered using a virtual reality (VR) apparatus and sounds were delivered using a single loudspeaker (mini speaker model JBL GO Portable from HARMAN International Industries, Northridge, California USA; $68.3 \times 82.7 \times 30.8 \text{ mm}$, Output Power 3.0 W; frequency response: 180 Hz–20 kHz) that was moved in space by the experimenter. For each sound localization trial, the experimenter stood next to the participant (who was sitting on a rotating chair) and silently moved the loudspeaker by hand to the desired sound position guided by a weak echo radar signal for elevation placement and visual feedback on a computer screen for direction placement. Further details of the experimental setup are available in Coudert et al. (2021).

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

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

Training Sessions

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

TABLE 1. Demographics and device information for the 12 bilateral cochlear implant users included in the training protocol.

Group	Id	Sex	Testing Age (y)	Diagnosis	Age at	Age at	First cochlear implant						Second cochlear implant									
							Etiology	Hearing aids (y)	Age at	Internal Part,	Strategy	Micro	Threshold (dB HL)	Word Recognition (%)	CI2 (y)	Processor, Ear Side	Strategy	Micro	Threshold (dB)	Word Recognition (%)	Interim-plant	Duration of
A	A01	M	18	Genetic	6	6	12	Nucleus CI422, CP1000, R	Processor, Ear Side	ACE	Omni	25	96	17	Nucleus CI522, CP1000, L	ACE	Omni	25	98	5	1	
	A02	F	46	Genetic	3	3	42	Nucleus CI422, CP910, L	Processor, Ear Side	ACE	Omni	25	86	45	Nucleus CI522, CP1000, R	ACE	Omni	30	90	3	1	
	A03	F	45	Meningitis	2	5	39	Concerto, RONDO2, R	Processor, Ear Side	Fs4p	Direc	35	88	44	Synchrony, RONDO2, L	Fs4p	Direc	35	96	5	1	
	A04	F	23	Ototoxicity	9	9	15	Concerto, SONNET, L	Processor, Ear Side	Fs4p	Omni	25	96	17	Concerto, OPUS2, R	Fs4p	Omni	25	97	2	6	
	A05	F	49	Unknown	6	6	42	CONCERTO, SONNET, R	Processor, Ear Side	Fs4p	Omni	25	70	46	Synchrony, SONNET, L	Fs4p	Omni	25	70	4	3	
	A06	M	31	Meningitis	2	2	2	Nucleus, CP910, R	Processor, Ear Side	SPEAK	Omni	25	72	28	Nucleus CI522, CP910, L	ACE	Omni	30	100	26	3	
			35.3±		4.7±	5.1±	25.3±					26.7±	84.7±	32.8±			28.3±	91.8±	7.5±	2.5±		
			13.2		2.8	2.5	17.7					4.1	11.4	13.9			4.1	11.2	9.1	2.0		
B	B01	F	69	Unknown	55	58	64	Digitonic SP EVO, SAPHYR	Processor, Ear Side	Crystalline XDP	Omni	35	92	67	Digitonic Ziti EVO, NEURO2, L	Crystalline CAP	Omni	35	84	3	2	
	B02	M	40	Ototoxicity	2	6	34	Digitonic SP, R	Processor, Ear Side	Crystalline XDP	Omni	30	82	35	Digitonic SP EVO, SAPHYR	Crystalline SPD	Omni	35	96	1	5	
	B03	F	46	Genetic	10	12	42	Concerto, SONNET, R	Processor, Ear Side	Fs4	Omni	25	100	42	Synchrony, SONNET, L	Fs4	Omni	25	97	0.9	4	
	B04	F	48	Unknown	13	30	38	Sonata, RONDO, R	Processor, Ear Side	Fs4	Direc	30	87	39	Sonata, RONDO2, L	Fs4	Direc	30	90	1	9	
	B05	M	55	Genetic	1	1	50	HiRes90K, NAIDA, R	Processor, Ear Side	HiRes Optima-S	Omni	20	94	52	HiRes90K, NAIDA, L	HiRes Optima-S	Omni	20	84	2	3	
	B06	M	24	Unknown	3	3	17	Sonata, SONNET, R	Processor, Ear Side	Fs4	Omni	30	87	20	Synchrony, SONNET, L	Fs4	Omni	30	99	3	4	
			47.0±		14.0±	18.3±	40.8±					28.3±	90.3±	42.5±			29.2±	91.7±	1.8±	4.5±		
			15.0		20.7	22.1	15.8					5.2	6.3	15.9			5.8	6.7	1.0	2.4		

Values in bold represent group averages with standard deviations. CI, cochlear implant; Direc, directional; Group A, audiovisual feedback; Group B, visual feedback; Id, patient identification; L, left; Omni, omnidirectional; R, right.

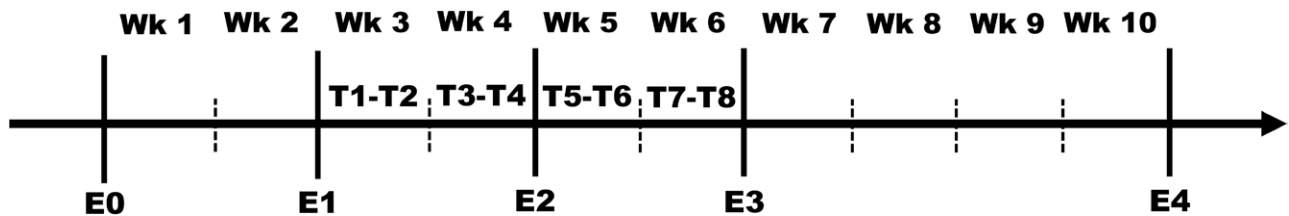


Fig. 1. Flowchart of the rehabilitation training protocol conducted over 10 weeks (Wk 1 to Wk 10), with five evaluation sessions (E0 to E4) distributed across the 10 weeks and 8 training sessions (T1 to T8) distributed across 4 weeks (weeks 3 to 6).

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

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

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

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

If the participant did not respond during the 15-second trial duration feedback appeared automatically at the end of the trial. Half of the participants (group A) received both visual and

auditory feedback, in the form of an image of the loudspeaker in virtual reality in a position spatially congruent with its actual position in space, while the sound stimulus continued playing from the loudspeaker's unchanged position. The other half (group B) received exactly the same visual feedback as group A (indicating the spatial position of the sound source) but without any sound (i.e., no spatial auditory information). The added value of providing multisensory information was investigated by comparing audiovisual feedback with visual-only feedback because the visual system dominates over the auditory system in perception of the environment (Witten & Knudsen 2005).

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

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

Evaluation Sessions

To control for the possibility that the evaluation session itself led to performance improvements, two evaluation sessions (E0 and E1) were performed before the start of training. The two evaluations were separated by 2 weeks and performance in the two sessions was compared (see Results). To evaluate the immediate benefits of the training sessions we performed an evaluation session in the middle of the training sessions (E2) and after the last training session (E3). A final evaluation (E4) was performed 1 month after the end of the last training session to evaluate whether any observed effects of training persisted over time. Each evaluation session consisted of a quality of life questionnaire and two objective tests: a 3D sound localization test and a speech comprehension test in noise.

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

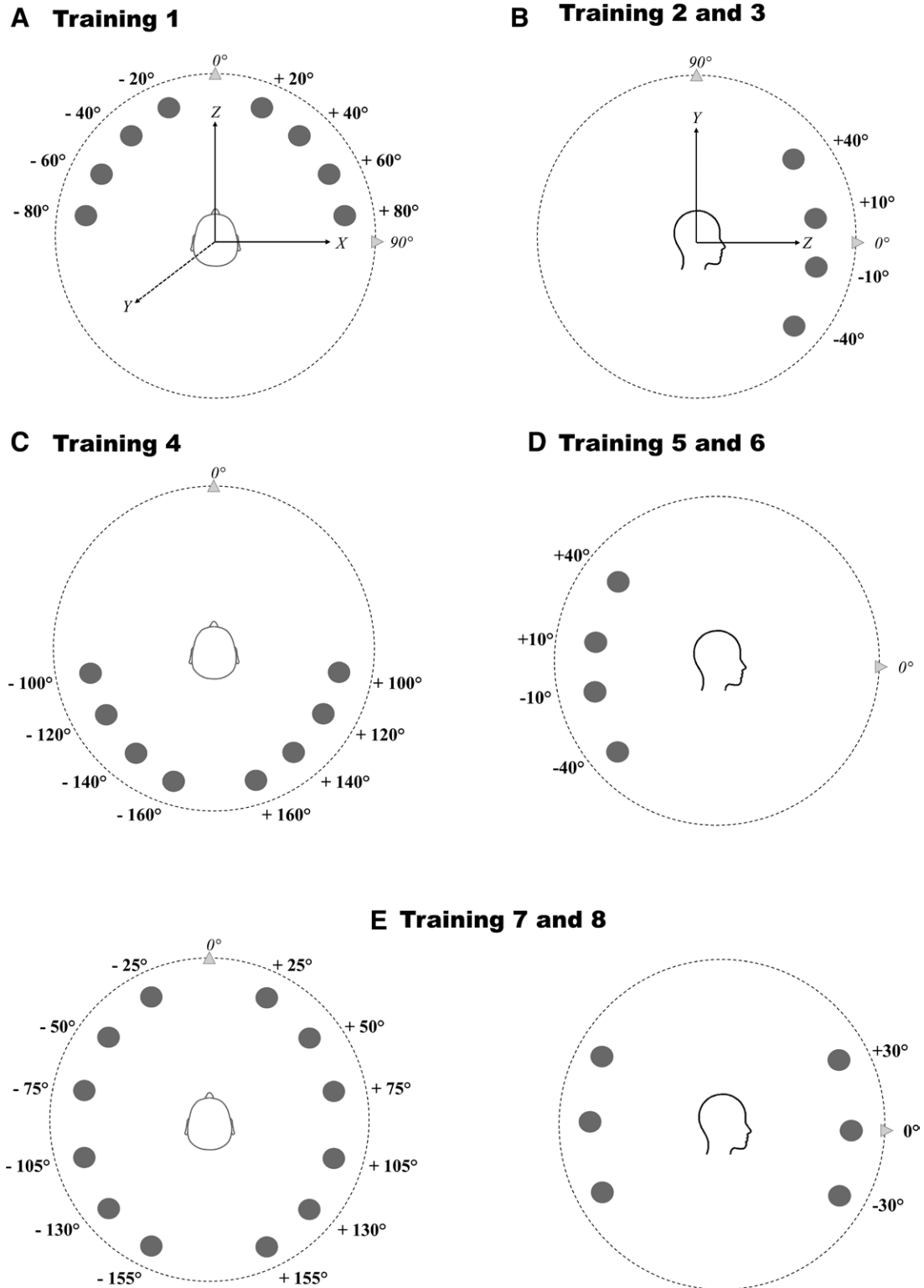


Fig. 2. Configuration of loudspeaker positions during the 8 training sessions (T1 to T8). A and B, Training sessions with sound sources placed in front space; (C) and (D) in back space; (E) in both front and back space with different azimuthal and elevation sound sources. Three axes were defined according to the reference-frame (i.e. participant head-centered): X, azimuth; Y, elevation; and Z, distance. Note that for trainings T1 and T4 elevation remained fixed at ear level; for trainings T2 and T3 azimuthal stimulations were fixed at -45° and $+45^\circ$; and for trainings T5 and T6 azimuthal stimulations were fixed at -135° and $+135^\circ$.

contains specific questions on sound localization and speech comprehension in noise. The short form of the SSQ has 15 items, divided into three subscales: speech perception; spatial hearing; and other qualities of hearing. Each question is scored

from 0 (not at all) to 10 (perfectly) and higher scores represent greater perceived ability in everyday situations. The questionnaire was administered during a face-to-face interview with an ENT doctor.

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

Speech Comprehension in Noise Task • Speech comprehension in noise was evaluated using the French Matrix Test (Jansen et al. 2012). This test was performed in the sound field in a calibrated room ($4.1 \text{ m} \times 2.8 \text{ m} \times 2.7 \text{ m}$, reverberation time RT60: 0.39 seconds). Two loudspeakers were placed at ear level, at -45° and $+45^\circ$ azimuth angle, and at 1 m from the participant's head. One loudspeaker delivered speech sentences (all based on the same syntactic structure: name-verb-number-object-color), while the other delivered noise (i.e., a stationary long-term average speech spectrum made by Oldenburg Measurement Application, www.hoertech.de). Sentences were always presented on the side of the better ear. The test consisted of three blocks: a training block of 10 sentences (results were

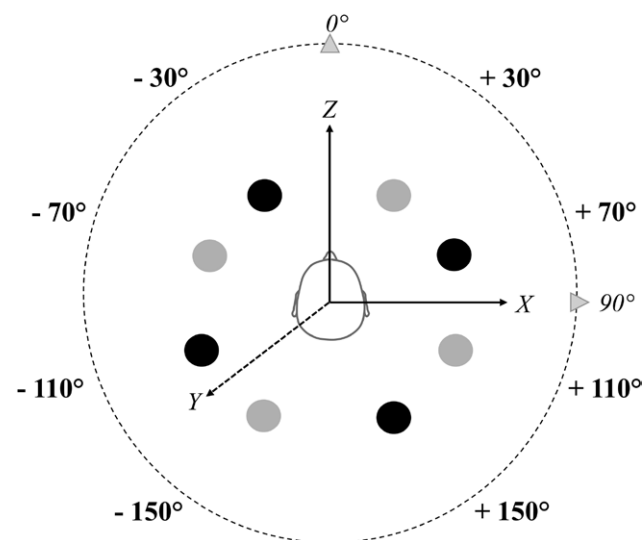


Fig. 3. Setup during sound localization evaluation sessions. A total of eight positions were tested. Black and gray circles indicate two target elevations in the near-field, at $+25^\circ$ and -25° , respectively. Distance was fixed at 55 cm. Three axes were defined according to the reference-frame (i.e.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

not taken into account in the statistical analyses), followed by two 20-sentence blocks. For each block, we used an adaptive procedure (see Jansen et al. 2012 for details) where noise was fixed at 60 dB SPL and speech level varied to obtain a signal-to-noise ratio (SNR) with 50% correct word recognition: used to determine the speech recognition threshold (SRT). The lower the SRT, the better the ability to discriminate speech in the presence of high background noise.

Data Analysis

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

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

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

RESULTS

Training Feasibility

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

Auditory Performance Before Training

To investigate the possibility that participation in the evaluation sessions had a beneficial effect on auditory performance, we compared performance at E0 and E1. The delay between these two sessions was 2 weeks and both were conducted before training began (see Fig. 1). Three separate Wilcoxon signed-rank tests indicated no significant performance differences

between E0 and E1 on any of the tests: sound localization test ($V = 29.5$, $p = 0.121$), matrix test ($V = 56$, $p = 0.051$), and SSQ questionnaire ($V = 40$, $p = 0.563$). These results suggest that the content of the evaluation sessions did not induce any learning. For all subsequent analyses, baseline performance was defined as performance at E1, as this session was the closest in time to the training sessions.

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

Auditory Performance Mid-training and Immediately After Training

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

The four middle panels of Figure 4 show pointing responses half-way through (E2) and immediately after the end of the training (E3), while the bottom panels show responses 4 weeks after the end of training (E4). The left-hand panels show that, compared with E1, participants' performance for back sound sources progressively improved, with the median percentage of front-back confusions decreasing from 25.8% during E2 to 14.8% during E3 and stabilizing at 14.1% during E4 (Fig. 5A). The group-level improvement at E3 was driven by eight of 12 bilateral CI users, all of whom had fewer than 25% of confusions, although it is important to note that at E4 nine of 12 participants also had fewer than 25% of confusions. A Friedman test including data from all four sessions found a significant decrease in front-back confusions between sessions ($\chi^2(3) = 10.9$, $p = 0.012$, effect size $W = 0.31$). Pairwise comparisons (Wilcoxon test with Bonferroni adjustment) showed that this improvement was significant between E1 and E3 ($p = 0.017$) and between E1 and E4 ($p = 0.023$) but not between E1 and E2 ($p = 0.169$).

The right-hand panels in the middle of Figure 4 show that even after 8 weeks of training up-down discrimination remained difficult, and the bottom right-hand panel shows that this difficulty was still present 4 weeks after the end of the protocol. This can also be seen in Figure 5B, which shows only a slight decrease in up-down confusions across the four evaluation sessions (from 46.1% during E2 to 42.2% during E4). Interestingly, even if the median number of up-down confusions at E4 was

similar to that at E3, several participants continued to improve between E3 and E4, and the improvements in two participants (A01 and B03) put them below the chance level (see Table 2). A Friedman test revealed no significant differences between the four evaluation sessions ($\chi^2(3) = 6.54$, $p = 0.088$).

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

One participant (A05) was never able to repeat the sentences in the Matrix test even after 8 weeks of training. Thus, Figure 6 shows SRT scores from all four evaluation sessions for 11 of 12 participants. The median SRT score decreased from 1.5 dB at E1 to 0.2 dB SNR at E2 and -0.5 dB at E3 then stabilized at -0.7 dB SNR at E4. In addition to clearly showing the gradual decrease in SRT scores across evaluation sessions, Figure 6 also shows a large decrease in between-participant variability across sessions. SRT scores ranged from 6.8 to -4.7 dB SNR at E1 and from 1.3 to -4.8 dB SNR at E4. SRT scores improved between E1 and E2 for eight of 11 participants and between E2 and E3 for 10 of 11 participants. Some participants continued to improve between E3 and E4, although these improvements were generally small and were observed in only five of 11 participants (see Table 3 for individual data). A Friedman test revealed a significant decrease in SRT scores across all 4 evaluation sessions ($\chi^2(3) = 12.6$, $p = 0.006$, effect size $W = 0.39$). Pairwise comparisons (Wilcoxon test with Bonferroni adjustment) showed that this improvement was significant between E1 and E3 ($p = 0.029$), and between E1 and E4 ($p = 0.018$) but not between E1 and E2 ($p = 0.194$). It is important to note that at the end of training, eight bilateral CI users were able to discriminate sentences with a *negative* SNR, meaning that they were able to correctly repeat sentences with a higher noise than the speech level. In summary, the data from the Matrix test show that SRT improved moderately with training and that this improvement was still present 4 weeks after the end of training.

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

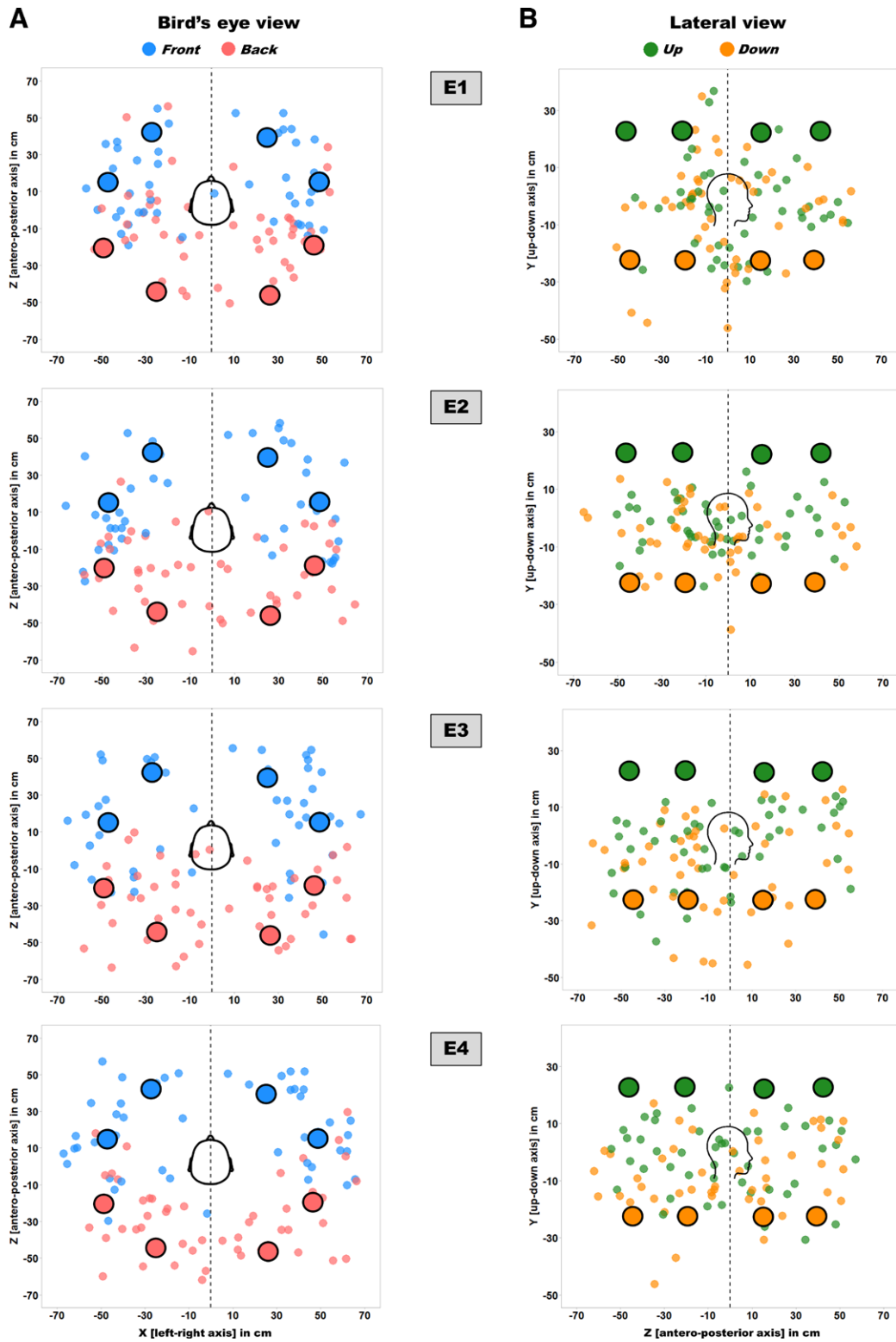


Fig. 4. Sound localization performance of all 12 bilateral cochlear implant users during each evaluation session: before training (E1), after 4 (E2), and 8 (E3) training sessions, and 1 month after the end of the training sessions (E4). Black outlined symbols represent the sound sources and colored symbols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

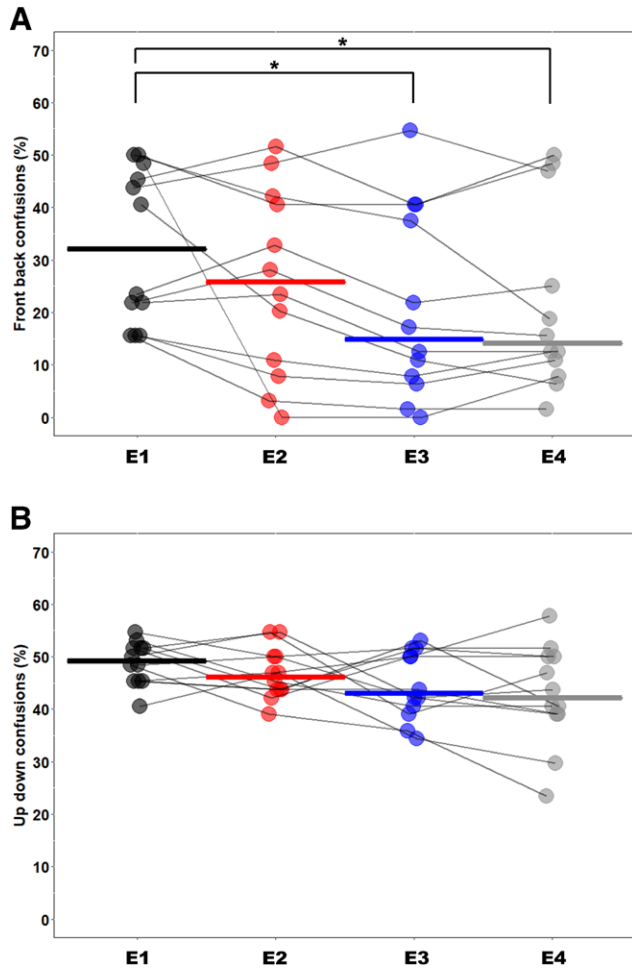


Fig. 5. Sound localization performance during evaluation sessions E1 to E4. Thick lines represent the median percentage of confusions for all 12 participants: (A) front-back confusions in azimuth and (B) up-down confusions in elevation. Colored dots correspond to the percentage of confusions for each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

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

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

Auditory Performance with Audiovisual or Visual-only Feedback During Training

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

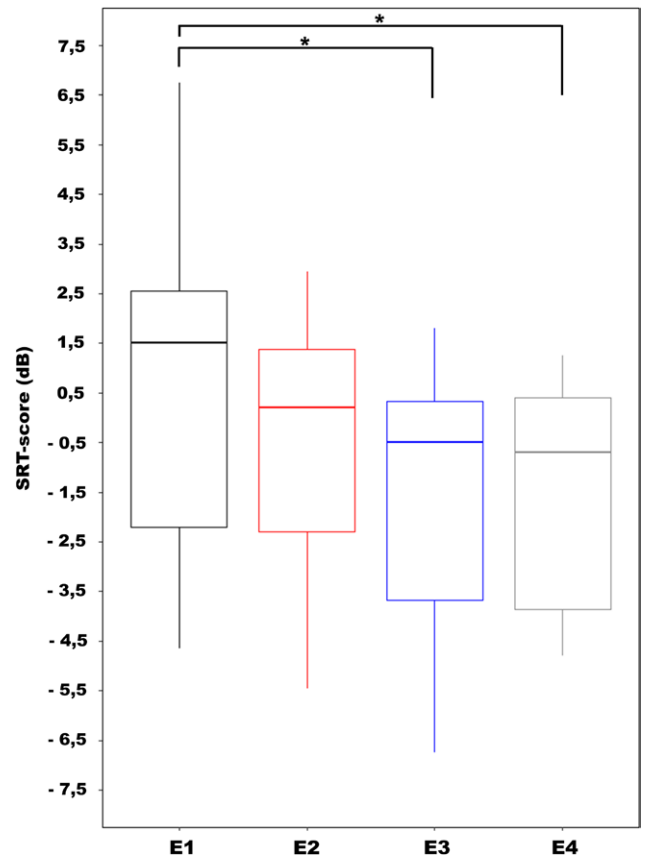


Fig. 6. SRT scores during evaluation sessions E1 to E4. Lines are medians, box limits 25th–75th percentiles and error bars 95% confidence limits ($n = 11$ for each session). Asterisks indicate significant differences (Wilcoxon test, $*p < 0.05$). SRT, speech recognition threshold.

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

DISCUSSION

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

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

Group	Id	Front-back Confusions (%)				Up-down Confusions (%)			
		E1	E2	E3	E4	E1	E2	E3	E4
A	A01	15.6	3.1	1.6	1.6	48.4	39.1	35.9	23.4
	A02	21.9	28.1	17.2	15.6	50.0	54.7	39.1	46.9
	A03	43.8	48.4	54.7	46.9	51.6	42.2	50.0	50.0
	A04	50.0	42.2	37.5	18.8	45.3	46.9	50.0	57.8
	A05	15.6	10.9	7.8	12.5	54.7	50.0	51.6	51.6
	A06	23.4	32.8	21.9	25.0	53.1	45.3	40.6	40.6
	Mean all	28.4	27.6	23.5	20.1	50.5	46.4	44.5	45.1
B	B01	45.3	51.6	40.5	48.4	48.4	50.0	42.2	43.8
	B02	50.0	40.6	40.6	50.0	45.3	43.8	51.6	50.0
	B03	15.6	7.8	6.3	10.9	40.6	46.9	34.4	29.7
	B04	40.6	20.3	10.9	6.3	51.6	54.7	42.2	39.1
	B05	21.9	23.4	12.5	12.5	45.3	43.8	43.8	39.1
	B06	48.4	0.0	0.0	7.8	51.6	43.8	53.1	40.6
	Mean all	37.0	24.0	18.5	22.7	47.1	47.2	44.6	40.4

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

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

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

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

et al. 2011), others virtual sounds (Mendonça et al. 2013, Steadman et al. 2019), some performed the training in a laboratory setting while others implemented at-home training (Tyler et al. 2010). Strelnikov et al. (2011) were one of the first groups to propose training sessions spread over several days, and this was done with NH adults wearing a monaural plus. Studies conducted with hearing impaired people are rare, with only three published studies of note. These included either patients with unilateral hearing loss (Firszt et al. 2015), bilateral moderate deafness (Kuk et al. 2014), or a single case study of a bilateral CI user (Tyler et al. 2010). The present study is thus the first to evaluate the feasibility and efficacy of a standardized rehabilitation protocol spread over 4 weeks for a group of bilateral CI users.

Spatial Hearing Improvement

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

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

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

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

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

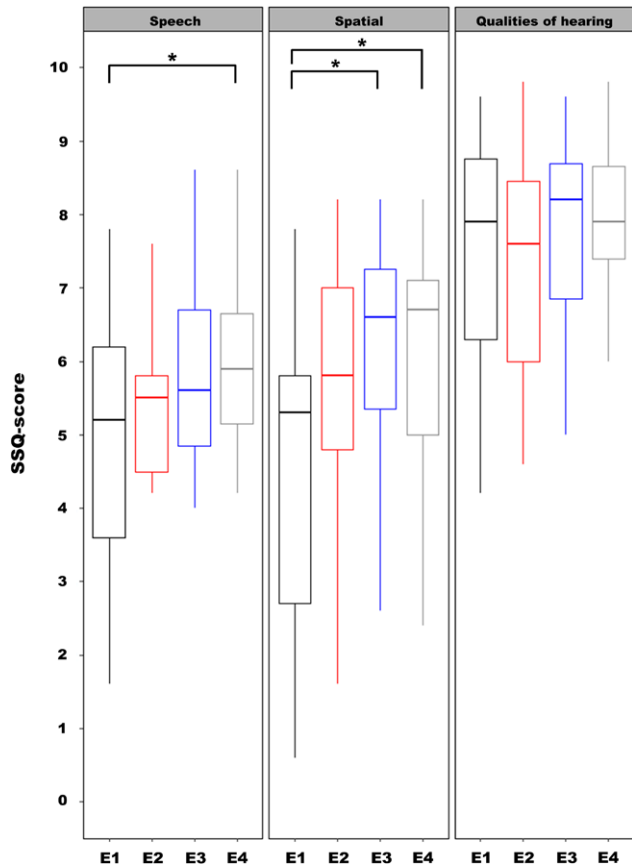


Fig. 7. SSQ-score during evaluation sessions E1 to E4 shown separately for each subscale (SSQ). Lines are medians, box limits 25th–75th percentiles and error bars 95% confidence limits ($n = 12$ for each session). Asterisks indicate significant differences (Wilcoxon test, $*p < 0.05$). SSQ, speech, spatial,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

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

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

Group	Id	Subscale	E1	E2	E3	E4	
A	A01	A	4.2	4.2	4	4.2	
		B	5.8	5	7.2	5.4	
		C	8.8	7	9	9.2	
	A02	A	2.4	4.6	5.2	5	
		B	0.6	1.4	2.2	2.4	
		C	8.8	8	8.6	8	
	A03	A	6.2	5.8	6.4	6.8	
		B	5.6	6.8	6.4	6.6	
		C	7.4	7.4	7.8	7.8	
	A04	A	3.6	5.6	5.6	5.2	
		B	7	8.2	7.4	7	
		C	8.4	8.4	8	7.8	
	A05	A	6.2	5.8	6.6	6.2	
		B	3	5.8	6.8	6.8	
		C	4.2	6	6.4	6.6	
	A06	A	5.6	4.8	5.6	6.6	
		B	1.8	1.6	2.6	3.8	
		C	9.6	9.8	9.6	9.8	
	B	B01	A	1.6	1.8	0.8	1.4
			B	0.6	4.2	4	3.8
			C	5.4	6	5	6
		B02	A	4.8	5.4	5	5.4
			B	5.4	5.8	5.8	6.2
			C	6.6	6	7	7.6
B03		A	7.8	8.6	8.6	8.6	
		B	7.8	8	8.2	7.4	
		C	8.4	8.6	9	8.6	
B04		A	3.6	4.2	4.4	5.6	
		B	4.8	5	6.2	7.6	
		C	4.8	4.6	6.2	6.8	
B05		A	7.8	7.6	8.2	8.2	
		B	5.8	7.6	8.2	8.2	
		C	8.8	8.6	8.6	8.8	
B06		A	6	5.8	7	6.4	
		B	5.2	6.2	6.8	6.8	
		C	7.4	7.8	8.4	8	

SSQ, speech, spatial, and qualities of hearing.

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

There is little data available on sound localization in elevation in CI users, but one study by Majdak et al. (2011) reported results similar to our pretraining results. That is, near-chance sound localization accuracy in elevation, large inter-individual variability, and localization clustered around the level of the CI. While sound localization in azimuth relies on binaural cues, detecting a sound in elevation is based on monaural spectral indices from upper body filtration (i.e., the pinna and to a lesser extent the head and shoulders) above 3 kHz (e.g., Musicant & Butler 1984; Perrett & Noble 1997). CI processors do not correctly provide these spectral cues, since all incident sound waves are directly caught by the microphone and the range of upper frequencies encoded by the processor is limited. It is therefore not surprising that unlike our results for front-back performance, only two participants improved their up-down discrimination after training, while the other 10 remained at the chance level. The improvement shown by these two participants raises the question of whether the rehabilitation protocol might

have trained them to identify and extract auditory information that enabled them to learn new monaural coordinate cues. This would be consistent with the suggestion by Algazi et al. (2001) that sound localization in elevation is possible from low spectral cues when sound sources are positioned laterally. It is also possible that the position of the microphone (i.e. behind or off the ears) could affect up-down discrimination by modifying monaural cues. However, the two best performers had their processors behind the ears, which is the most detrimental situation compared to a microphone placed inside the external auditory canal. Further investigations are clearly needed to better understand the effect of microphone position on spatial hearing performance.

Rehabilitation Benefits Beyond Spatial Hearing

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

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

While the 3D sound localization test and the matrix test were important for objectively assessing performance, they were performed under artificial experimental conditions. For this reason, we decided to evaluate the impact of the training sessions on participants' daily life using a validated quality of life questionnaire; the short form of the SSQ (Moulin et al. 2019). In line with the content of the training protocol, bilateral CI users' scores on questions about sound localization improved (from 5.3 to 6.6). In line with their improvement on the Matrix test, their scores for speech comprehension in noise also improved (from 5.2 to 5.6), but less than for sound localization. It is difficult to assess whether these changes are clinically/behaviorally relevant, as there are currently no norms for CI users on

the SSQ questionnaire. This questionnaire has most often been used to evaluate the benefits of cochlear implantation in the first 2 years after surgery (Hassepass et al. 2013; Zhang et al. 2015), and no data exist at longer delays, when performance and device settings are stable. Based on the categorization system developed by Noble and Gatehouse (2006), we found that one third of participants significantly improved their "speech perception score" and two thirds their "spatial hearing score." Moreover, when compared to the results of a large cohort of patients suffering from moderate hearing loss, that is, a maximum loss of 55 dB HL (Moulin et al. 2019), our mean scores per subscale one month after the end of training were close to theirs (speech perception: 5.9 and 6.7, spatial hearing: 6.7 and 6.5, other qualities of hearing: 7.9 and 8.2, respectively, in our study and theirs). This suggests that our rehabilitation protocol allowed bilateral CI users to reach a similar hearing-related quality of life to that reported by patients with a less disabling hearing deficit.

Persistent Benefits

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

Multisensory Stimulation and Feedback

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

Recent studies have shown that multisensory training can promote subsequent unisensory learning (Shams et al. 2011; Isaiah & Hartley 2015), and that adding redundant information from other intact sensory modalities (e.g., vision) does not make the task too easy but instead reduces the effort involved and promotes better learning (Strelnikov et al. 2011; Isaiah & Hartley 2015). In everyday situations, patients with hearing deficits rely heavily upon the visual system to compensate for the lack of information from the auditory system (Rouger et

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

Clinical Implications

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

next step in developing our training protocol could be the addition of more complex stimuli that simulate real-life situations (e.g., adding background noise and varying its sound level). This type of rehabilitation protocol could also be proposed to patients wearing hearing aids, who also experience spatial hearing difficulties.

CONCLUSION

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

ACKNOWLEDGMENTS

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

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

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

The authors have no conflicts of interest to disclose.

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

Received August 10, 2021; accepted May 13, 2022

REFERENCES

- Algazi, V. R., Avendano, C., Duda, R. O. (2001). Elevation localization and head-related transfer function analysis at low frequencies. *J Acoust Soc Am*, *109*, 1110–1122.
- Alhanbali, S., Dawes, P., Lloyd, S., Munro, K. J. (2017). Self-reported listening-related effort and fatigue in hearing-impaired adults. *Ear Hear*, *38*, e39–e48.
- Anderson, S., & Kraus, N. (2010). Sensory-cognitive interaction in the neural encoding of speech in noise: A review. *J Am Acad Audiol*, *21*, 575–585.
- Angell JR, Fite W. (1901a). The monaural localization of sound. *Psychol Rev*, *8*, 225–246.
- Angell JR, Fite W. (1901b). Further observations on the monaural localization of sound. *Psychol Rev*, *8*, 449–458.
- Archer-Boyd, A. W., & Carlyon, R. P. (2019). Simulations of the effect of unlinked cochlear-implant automatic gain control and head movement on interaural level differences. *J Acoust Soc Am*, *145*, 1389.
- Aronoff, J. M., Yoon, Y. S., Freed, D. J., Vermiglio, A. J., Pal, I., Soli, S. D. (2010). The use of interaural time and level difference cues by bilateral cochlear implant users. *J Acoust Soc Am*, *127*, EL87–EL92.

- Ausili, S. A., Agterberg, M. J. H., Engel, A., Voelter, C., Thomas, J. P., Brill, S., Snik, A. F. M., Dazert, S., Van Opstal, A. J., Mylanus, E. A. M. (2020). Spatial hearing by bilateral cochlear implant users with temporal fine-structure processing. *Front Neurol*, *11*, 915.
- Bolognini, N., Frassinetti, F., Serino, A., Lådavas, E. (2005). "Acoustical vision" of below threshold stimuli: Interaction among spatially converging audiovisual inputs. *Exp Brain Res*, *160*, 273–282.
- Brimjoin, W. O., McShefferty, D., Akeroyd, M. A. (2010). Auditory and visual orienting responses in listeners with and without hearing-impairment. *J Acoust Soc Am*, *127*, 3678–3688.
- Brimjoin, W. O., McShefferty, D., Akeroyd, M. A. (2012). Undirected head movements of listeners with asymmetrical hearing impairment during a speech-in-noise task. *Hear Res*, *283*, 162–168.
- Brungart, D. S. (1999). Auditory localization of nearby sources. III. Stimulus effects. *J Acoust Soc Am*, *106*, 3589–3602.
- Brungart, D. S., & Rabinowitz, W. M. (1999). Auditory localization of nearby sources. Head-related transfer functions. *J Acoust Soc Am*, *106*(3Pt 1), 1465–1479.
- Bulkin, D. A., & Groh, J. M. (2006). Seeing sounds: Visual and auditory interactions in the brain. *Curr Opin Neurobiol*, *16*, 415–419.
- Coudert, A., Gaveau, V., Gatel, J., Verdet, G., Salemme, R., Farne, A., Pavani, F., Truy, E. (2022). Spatial hearing difficulties in reaching space in bilateral cochlear implant children improve with head movements. *Ear Hear*, *43*, 192–205.
- Da Silva, J. A. (1985). Scales for perceived egocentric distance in a large open field: Comparison of three psychophysical methods. *Am J Psychol*, *98*, 119–144.
- Dincer D'Alessandro, H., Ballantyne, D., Boyle, P. J., De Seta, E., DeVincentiis, M., Mancini, P. (2018). Temporal fine structure processing, pitch, and speech perception in adult cochlear implant recipients. *Ear Hear*, *39*, 679–686.
- Dorman, M. F., Loiselle, L., Stohl, J., Yost, W. A., Spahr, A., Brown, C., Cook, S. (2014). Interaural level differences and sound source localization for bilateral cochlear implant patients. *Ear Hear*, *35*, 633–640.
- Firszt, J. B., Reeder, R. M., Dwyer, N. Y., Burton, H., Holden, L. K. (2015). Localization training results in individuals with unilateral severe to profound hearing loss. *Hear Res*, *319*, 48–55.
- Fischer, T., Schmid, C., Kompis, M., Mantokoudis, G., Caversaccio, M., & Wimmer, W. (2020). Pinna-imitating microphone directionality improves sound localization and discrimination in bilateral cochlear implant users. *Ear and Hearing*, *42*, 214–222.
- Gatehouse, S., & Noble, W. (2004). The speech, spatial and qualities of hearing scale (SSQ). *Int J Audiol*, *43*, 85–99.
- Giraud, A. L., Price, C. J., Graham, J. M., Truy, E., Frackowiak, R. S. (2001). Cross-modal plasticity underpins language recovery after cochlear implantation. *Neuron*, *30*, 657–663.
- Grantham, D. W., Ashmead, D. H., Ricketts, T. A., Labadie, R. F., Haynes, D. S. (2007). Horizontal-plane localization of noise and speech signals by postlingually deafened adults fitted with bilateral cochlear implants. *Ear Hear*, *28*, 524–541.
- Hassepass, F., Schild, C., Aschendorff, A., Laszig, R., Maier, W., Beck, R., Wesarg, T., Arndt, S. (2013). Clinical outcome after cochlear implantation in patients with unilateral hearing loss due to labyrinthitis ossificans. *Otol Neurotol*, *1278*–1283.
- Hughes, S. E., Hutchings, H. A., Rapport, F. L., McMahon, C. M., Boisvert, I. (2018). Social connectedness and perceived listening effort in adult cochlear implant users: A grounded theory to establish content validity for a new patient-reported outcome measure. *Ear Hear*, *39*, 922–934.
- Isaiah, A., & Hartley, D. E. (2015). Can training extend current guidelines for cochlear implant candidacy? *Neural Regen Res*, *10*, 718–720.
- Jansen, S., Luts, H., Wagener, K. C., Kollmeier, B., Del Rio, M., Dauman, R., James, C., Fraysse, B., Vormès, E., Frachet, B., Wouters, J., van Wieringen, A. (2012). Comparison of three types of French speech-in-noise tests: A multi-center study. *Int J Audiol*, *51*, 164–173.
- Kerber, S., & Seeber, B. U. (2012). Sound localization in noise by normal-hearing listeners and cochlear implant users. *Ear Hear*, *33*, 445–457.
- Kolarik, A. J., Moore, B. C., Zahorik, P., Cirstea, S., Pardhan, S. (2016). Auditory distance perception in humans: A review of cues, development, neuronal bases, and effects of sensory loss. *Atten Percept Psychophys*, *78*, 373–395.
- Kuk, F., Keenan, D. M., Lau, C., Crose, B., Schumacher, J. (2014). Evaluation of a localization training program for hearing impaired listeners. *Ear Hear*, *35*, 652–666.
- Laback, B., Egger, K., Majdak, P. (2015). Perception and coding of interaural time differences with bilateral cochlear implants. *Hear Res*, *322*, 138–150.
- Lafon, J.C. (1964). Le test phonétique et la mesure de l'audition, éditions centrex – eindhoven, 144–146.
- Loomis, J. M., Klatzky, R. L., Philbeck, J. W., Golledge, R. G. (1998). Assessing auditory distance perception using perceptually directed action. *Percept Psychophys*, *60*, 966–980.
- MacLeod, A., & Summerfield, Q. (1987). Quantifying the contribution of vision to speech perception in noise. *Br J Audiol*, *21*, 131–141.
- Majdak, P., Goupell, M. J., Laback, B. (2011). Two-dimensional localization of virtual sound sources in cochlear-implant listeners. *Ear Hear*, *32*, 198–208.
- Mendonça, C., Campos, G., Dias, P., & Santos, J. A. (2013). Learning auditory space: Generalization and long-term effects. *PLOS ONE*, *8*, e77900.
- Middelweerd, M. J., & Plomp, R. (1987). The effect of speechreading on the speech-reception threshold of sentences in noise. *J Acoust Soc Am*, *82*, 2145–2147.
- Middlebrooks, J. C. (2015). Sound localization. *Handb Clin Neurol*, *129*, 99–116.
- Mo, B., Lindbaek, M., Harris, S. (2005). Cochlear implants and quality of life: A prospective study. *Ear Hear*, *26*, 186–194.
- Moore, B. C. J. (2008). The role of temporal fine structure processing in pitch perception, masking, and speech perception for normal-hearing and hearing-impaired people. *J Assoc Res Otolaryngol: JARO*, *9*, 399–406.
- Moulin, A., Vergne, J., Gallego, S., & Michey, C. (2019). A new speech, spatial, and qualities of hearing scale short-form: Factor, cluster, and comparative analyses. *Ear and Hearing*, *40*, 938–950.
- Mueller, M. F., Meisenbacher, K., Lai, W. K., Dillier, N. (2014). Sound localization with bilateral cochlear implants in noise: How much do head movements contribute to localization? *Cochlear Implants Int*, *15*, 36–42.
- Musicant, A. D., & Butler, R. A. (1984). The influence of pinnae-based spectral cues on sound localization. *J Acoust Soc Am*, *75*, 1195–1200.
- Nawaz, S., McNeill, C., Greenberg, S. L. (2014). Improving sound localization after cochlear implantation and auditory training for the management of single-sided deafness. *Otol Neurotol*, *35*, 271–276.
- Noble, W., & Gatehouse, S. (2006). Effects of bilateral versus unilateral hearing aid fitting on abilities measured by the speech, spatial, and qualities of hearing scale (SSQ). *Int J Audiol*, *45*, 172–181.
- O'Neill, E. R., Kreft, H. A., Oxenham, A. J. (2019). Cognitive factors contribute to speech perception in cochlear-implant users and age-matched normal-hearing listeners under vocoded conditions. *J Acoust Soc Am*, *146*, 195.
- Pastore, M. T., Natale, S. J., Yost, W. A., Dorman, M. F. (2018). Head movements allow listeners bilaterally implanted with cochlear implants to resolve front-back confusions. *Ear Hear*, *39*, 1224–1231.
- Pavani, F., Husain, M., Driver, J. (2008). Eye-movements intervening between two successive sounds disrupt comparisons of auditory location. *Exp Brain Res*, *189*, 435–449.
- Perreau, A. E., Ou, H., Tyler, R., Dunn, C. (2014). Self-reported spatial hearing abilities across different cochlear implant profiles. *Am J Audiol*, *23*, 374–384.
- Perrett, S., & Noble, W. (1997). The effect of head rotations on vertical plane sound localization. *J Acoust Soc Am*, *102*, 2325–2332.
- Rouger, J., Lagleyre, S., Fraysse, B., Deneve, S., Deguine, O., Barone, P. (2007). Evidence that cochlear-implanted deaf patients are better multi-sensory integrators. *Proc Natl Acad Sci U S A*, *104*, 7295–7300.
- Roy, A. T., Carver, C., Jiradejvong, P., Limb, C. J. (2015). Musical sound quality in cochlear implant users: A comparison in bass frequency perception between fine structure processing and high-definition continuous interleaved sampling strategies. *Ear Hear*, *36*, 582–590.
- Schwartz, J. L., Berthommier, F., Savariaux, C. (2004). Seeing to hear better: Evidence for early audio-visual interactions in speech identification. *Cognition*, *93*, B69–B78.
- Seeber, B. U., Baumann, U., Fastl, H. (2004). Localization ability with bimodal hearing aids and bilateral cochlear implants. *J Acoust Soc Am*, *116*, 1698–1709.
- Shams, L., Wozny, D. R., Kim, R., Seitz, A. (2011). Influences of multisensory experience on subsequent unisensory processing. *Front Psychol*, *2*, 264.
- Shinn-Cunningham, B., Best, V., Lee, A. K. (2017). Auditory object formation and selection. In *The Auditory System at the Cocktail Party*. Springer, J.C. Middlebrooks, J.Z. Simon, A.N. Popper, R.R. Fay. 7–40.

- Smulders, Y. E., van Zon, A., Stegeman, I., Rinia, A. B., Van Zanten, G. A., Stokroos, R. J., Hendrice, N., Free, R. H., Maat, B., Frijns, J. H., Briaire, J. J., Mylanus, E. A., Huinck, W. J., Smit, A. L., Topsakal, V., Tange, R. A., Grolman, W. (2016). Comparison of bilateral and unilateral cochlear implantation in adults: A randomized clinical trial. *JAMA Otolaryngol Head Neck Surg*, *142*, 249–256.
- Steadman, M. A., Kim, C., Lestang, J. H., Goodman, D. F. M., Picinali, L. (2019). Short-term effects of sound localization training in virtual reality. *Sci Rep*, *9*, 18284.
- Strelnikov, K., Rosito, M., Barone, P. (2011). Effect of audiovisual training on monaural spatial hearing in horizontal plane. *PLoS One*, *6*, e18344.
- Tyler, R. S., Perreau, A. E., Ji, H. (2009). Validation of the spatial hearing questionnaire. *Ear Hear*, *30*, 466–474.
- Tyler, R. S., Witt, S. A., Dunn, C. C., Wang, W. (2010). Initial development of a spatially separated speech-in-noise and localization training program. *J Am Acad Audiol*, *21*, 390–403.
- Valzolgher, C., Alzhaler, M., Gessa, E., Todeschini, M., Nieto, P., Verdelet, G., Salemmé, R., Gaveau, V., Marx, M., Truy, E., Barone, P., Farnè, A., Pavani, F. (2020a). The impact of a visual spatial frame on real sound-source localization in virtual reality. *Current Research in Behavioral Sciences*, *1*, 100003.
- Valzolgher, C., Verdelet, G., Salemmé, R., Lombardi, L., Gaveau, V., Farnè, A., Pavani, F. (2020b). Reaching to sounds in virtual reality: A multi-sensory-motor approach to promote adaptation to altered auditory cues. *Neuropsychologia*, *149*, 107665.
- van Hoesel, R. J. M. (2004). Exploring the benefits of bilateral cochlear implants. *Audiol. Neurootol*, *9*, 234–246.
- van Hoesel, R. J., & Tyler, R. S. (2003). Speech perception, localization, and lateralization with bilateral cochlear implants. *J Acoust Soc Am*, *113*, 1617–1630.
- van Zon, A., Smulders, Y. E., Stegeman, I., Ramakers, G. G., Kraaijenga, V. J., Koenraads, S. P., Zanten, G. A., Rinia, A. B., Stokroos, R. J., Free, R. H., Frijns, J. H., Huinck, W. J., Mylanus, E. A., Tange, R. A., Smit, A. L., Thomeer, H. G., Topsakal, V., Grolman, W. (2017). Stable benefits of bilateral over unilateral cochlear implantation after two years: A randomized controlled trial. *Laryngoscope*, *127*, 1161–1168.
- Verdelet, G., Desoche, C., Volland, F., Farnè, A., Coudert, A., Hermann, R., Truy, E., Gaveau, V., Pavani, F., Salemmé, R. (2019). *Assessing Spatial and Temporal Reliability of the Vive System as a Tool for Naturalistic Behavioural Research. 2019 International Conference on 3D Immersion (IC3D)*, 1–8. <https://doi.org/10.1109/IC3D48390.2019.8975994>
- Wallach, H. (1940). The role of head movements and vestibular and visual cues in sound localization. *J. Exp. Psychol*, *27*, 339–368.
- Wightman, F. L., & Kistler, D. J. (1999). Resolution of front-back ambiguity in spatial hearing by listener and source movement. *J Acoust Soc Am*, *105*, 2841–2853.
- Witten, I. B., & Knudsen, E. I. (2005). Why seeing is believing: Merging auditory and visual worlds. *Neuron*, *48*, 489–496.
- Won, J. H., Lorenzi, C., Nie, K., Li, X., Jameyson, E. M., Drennan, W. R., & Rubinstein, J. T. (2012). The ability of cochlear implant users to use temporal envelope cues recovered from speech frequency modulation. *The Journal of the Acoustical Society of America*, *132*, 1113–1119.
- Zaltz, Y., Bugannim, Y., Zechoval, D., Kishon-Rabin, L., Perez, R. (2020). Listening in noise remains a significant challenge for cochlear implant users: Evidence from early deafened and those with progressive hearing loss compared to peers with normal hearing. *J Clin Med*, *9*, E1381.
- Zhang, J., Tyler, R., Ji, H., Dunn, C., Wang, N., Hansen, M., Gantz, B. (2015). Speech, spatial and qualities of hearing scale (ssq) and spatial hearing questionnaire (shq) changes over time in adults with simultaneous cochlear implants. *Am. J. Audiol*, *24*, 384–397.

Reaching to Sounds Improves Spatial Hearing in Bilateral Cochlear Implant Users

Chiara Valzolgher,^{1,2} Julie Gatel,³ Sabrina Bouzaid,¹ Solene Grenouillet,¹
Michela Todeschini,⁴ Gregoire Verdelet,^{1,5} Romeo Salemme,^{1,5} Valerie Gaveau,^{1,5}
Eric Truy,^{1,3} Alessandro Farnè,^{1,2,5} and Francesco Pavani^{1,2,4}

Objectives: We assessed if spatial hearing training improves sound localization in bilateral cochlear implant (BCI) users and whether its benefits can generalize to untrained sound localization tasks.

Design: In 20 BCI users, we assessed the effects of two training procedures (spatial versus nonspatial control training) on two different tasks performed before and after training (head-pointing to sound and audiovisual attention orienting). In the spatial training, participants identified sound position by reaching toward the sound sources with their hand. In the nonspatial training, comparable reaching movements served to identify sound amplitude modulations. A crossover randomized design allowed comparison of training procedures within the same participants. Spontaneous head movements while listening to the sounds were allowed and tracked to correlate them with localization performance.

Results: During spatial training, BCI users reduced their sound localization errors in azimuth and adapted their spontaneous head movements as a function of sound eccentricity. These effects generalized to the head-pointing sound localization task, as revealed by greater reduction of sound localization error in azimuth and more accurate first head-orienting response, as compared to the control nonspatial training. BCI users benefited from auditory spatial cues for orienting visual attention, but the spatial training did not enhance this multisensory attention ability.

Conclusions: Sound localization in BCI users improves with spatial reaching-to-sound training, with benefits to a nontrained sound localization task. These findings pave the way to novel rehabilitation procedures in clinical contexts.

Key words: Active listening, Cochlear implant, Head movements, Reaching, Spatial hearing, Training, Virtual reality.

(Ear & Hearing 2022;XX;00–00)

INTRODUCTION

Cochlear implants (CIs) have become standard intervention for neurosensory deafness (Moore & Shannon 2009; Wilson 2019). Bilateral CIs (BCIs) are increasingly common, thus allowing partial access to interaural auditory cues (Aronoff et al. 2010; Gifford & Stecker 2020). Benefits of CIs for speech understanding are now well-established in children (Allen et al. 1998; Calmels et al. 2004), adults (Wilson et al. 1991; Stickney et al. 2004; Wilson 2017), and aging individuals (Forli et al. 2019). In addition, binaural hearing improves sound localization in the

horizontal dimension for BCI users, compared to unilateral users (between group comparison) (Smulders et al. 2016), or to conditions in which one of the two CIs is switched off (within-group comparison) (Litovsky et al. 2009; Asp et al. 2012).

Yet, spatial hearing skills of BCI users are substantially poorer compared to listeners with normal hearing, when locating stationary (Kerber & Seeber 2012; Jones et al. 2014) as well as moving sound sources (Moua et al. 2019). A key factor causing such poor performance is the reduced availability of auditory cues. CIs limit input resolution in the temporal and frequency domains (Moore & Shannon 2009), alter sound levels through automatic gain control (AGC) (see, for instance, Dorman et al. 2014; Archer-Boyd & Carlyon 2019, 2021), and change the auditory cues by the effect of microphone filtering strategies (both related to noise reduction and to emphasize higher frequencies) (Seeber & Fastl 2008). In addition, the two CI processors work in isolation from one another, leading to independent acoustic streams to the brain (for discussion, see Verhaert et al. 2012; Roskoth-Kuhl et al. 2021). Hence, BCI users can only access a degraded set of auditory cues to localize sound, among which the most salient ones are the interaural-level differences (ILDs) (Van Hoesel 2004; Grantham et al. 2007; Seeber & Fastl 2008; Grieco-Calub & Litovsky 2010).

All these limitations pertain to peripheral auditory signals. Can we rely on more central auditory learning processes to better use the degraded signal? Relearning spatial hearing skills is possible in deafened ferrets with BCI after multisensory training with interleaved auditory and visual stimuli (Isaiah et al. 2014). Training-induced improvements were observed in the responsiveness of auditory cortical neurons and in their sensitivity to ILDs. Improvement emerged also in behavior: both in the sound localization responses and in sound-evoked head-orienting movements made by early-deafened ferrets with BCI (Isaiah et al. 2014). Studies with normal-hearing participants exposed to altered auditory cues also suggest that relearning spatial hearing is possible. Training protocols leveraging multisensory perception and active interactions with sounds proved effective in promoting adaptation to altered auditory cues and improved spatial hearing skills (Strelnikov et al. 2011; Carlile 2014; Mendonça 2014; Keating et al. 2015; Rabini et al. 2019; Valzolgher et al. 2020c). Finally, preliminary results in unilateral cochlear implant users (Luntz et al. 2005) and BCI users (Tyler et al. 2010) suggest that feedback-based training could also improve sound localization in this population. Yet, these pioneering works involved a limited number of participants and—critically—they did not explore if training effects could generalize beyond the trained task, which constitutes a hallmark of learning processes.

¹Integrative, Multisensory, Perception, Action and Cognition Team, Lyon Neuroscience Research Center, University Lyon 1, France; ²Center for Mind/Brain Sciences, University of Trento, Italy; ³Hospices Civils de Lyon, Lyon, France; ⁴Department of Psychology and Cognitive Sciences (DiPSCo), University of Trento, Italy; ⁵Neuroimmersion, Lyon, France

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

The aim of the present study was to assess whether spatial hearing can be trained in BCI users and whether training effects can generalize to untrained spatial auditory tasks. Building upon our previous work in normal-hearing adults listening with one ear plugged (Valzolgher et al. 2020c, 2022), we trained BCI users with a multisensory-motor task that combined reaching to near-by sounds, audiovisual feedback about sound position, and free-head movements. Reaching to sound could enhance spatial coding of sound position by incorporating the coordination of different effectors (eyes, head, hand) into a common reference frame (Cohen & Andersen 2002; Boyer et al. 2013). In addition, it could help directing attention toward the position occupied by the sound source and make the task more engaging, which is a fundamental feature in learning (Dehaene 2020). Finally, in monaurally plugged normal-hearing individuals, reaching to sounds encourage additional head movements in response to sounds, which appear to serve the purpose to explore the auditory scene more widely and to expose the unplugged ear to the sound sources (Valzolgher et al. 2020c).

To assess training specificity, we used a crossover experimental design (Fig. 1A; see Video File in Supplemental Digital Content 1, <http://links.lww.com/EANDH/B44>) in which we compared the experimental spatial training with a control nonspatial training that used the same stimuli and multisensory feedback, but different task demands (i.e., discrimination of amplitude modulations in the sound). Thus, all participants performed both the experimental training and the control one. Crossover paradigms are particularly suited for clinical applications (Senn 2002), because they permit the involvement of the totality of participants in the experimental training without the need of testing a control group (which could not benefit from training). Furthermore, taking advantage of this design, we minimized any potential intergroup differences. Specifically, we assessed if the improvement induced by the experimental training (spatial training) extends (i.e., generalizes) to a head-pointing sound localization task, which entails different sound positions and requires localizing sounds using a different effector (the head instead of the hand). In addition, to probe generalization effects also when implicit sound localization is required, we tested participants in an audiovisual attention-orienting task, in which they were asked to judge the elevation of a visual stimuli while listening for a sudden sound.

Based on work by Valzolgher et al. (2020c, 2022), we predicted observing an improvement in sound localization across trials during the spatial training. In addition, we expected greater generalization effects after the spatial compared to the nonspatial training (as recently documented by Valzolgher et al. 2022, using a similar approach in normal-hearing adults listening with one ear plugged). Finally, because our apparatus allowed for unrestrained but measurable head movements, in line with Isaiah et al. (2014) we also studied changes in sound-related head-orienting movements. We predicted finding changes in sound-related head-orienting movements during (Valzolgher et al. 2022) and after training (Isaiah et al. 2014), consistent with the expected error reductions.

MATERIALS AND METHODS

Study Design

All participants (BCI users) took part in two experimental sessions, with at least 15 days of washout interval. Each session

comprised two testing phases and one training (Fig. 1A); they were identical in all respects except for the training task. Participants performed both training types, which were counterbalanced across participants in a within-subject crossover design. In one session, they engaged in the experimental spatial training task; in the other, they performed the control nonspatial training task. Testing phases included two auditory tasks: a head-pointing localization task (conducted in virtual reality (VR)) and an audiovisual attention-orienting task (conducted outside VR, see below). The head-pointing localization task differed with respect to training both in terms of targets spatial position and response effector (the head instead of the hand). The rationale was to minimize any potential effect of sensorimotor adaptation. The audiovisual attention-orienting task was instead an implicit sound localization task, in which sounds served as lateralized attention-orienting cues for the discrimination of visual targets (i.e., an audiovisual analog to the classic attention cueing paradigm) (see below for detailed descriptions of each tasks).

Participants

Twenty BCI participants (age: $M = 45.6$, $SD = 13.1$, 7 males, 15 right-handed) took part in the study (see the Materials for power calculation in Supplemental Digital Content 2, <http://links.lww.com/EANDH/B43>).

They were recruited at the otorynolaryngology department of the civil Hospital Edouard Herriot in Lyon (France), and tested in a dedicated room inside the Hospital Edouard Herriot premises. Participants signed an informed consent before starting the experiment, which was conducted according to the criteria of the Declaration of Helsinki (1964, amended in 2013) and it was approved by the Ethical Committee (Ile de France X, N° ID RCB 2019-A02293-54), and recorded in clinicaltrials.gov (NCT04183348). All had normal or corrected-to-normal vision and reported no movement or vestibular deficit, nor neurological or psychiatric history. Age, sex and clinical data for individual BCI participants are provided in Table 1. During the experiment, participants used their daily sound processor settings (see Table 1 in Supplemental Digital Content 2, <http://links.lww.com/EANDH/B43> for details about CI model, processor strategy and microphone settings), to preserve the acoustic experience that each participant developed with the CIs.

Procedures

Testing Phases

Head-Pointing Sound Localization Task

This task was carried out entirely in VR, always using *real sounds* delivered in free field from predetermined positions computed on each trial based on the initial head position (Coudert et al. 2022; Valzolgher et al. 2020a,c). Our VR approach brings real sounds into visual VR and coordinates everything using Unity (see the Materials in Supplemental Digital Content 2, <http://links.lww.com/EANDH/B43> for details about the apparatus and also Gaveau et al. 2020). Participants were immersed in a virtual room that matched the size of the real one but was devoid of any objects. Participants were instructed to localize sound sources and informed that sound could be presented anywhere in the surrounding space. At the beginning of each trial, participants were asked to direct their gaze in front of them, by aligning their head with a central fixation cross. When the

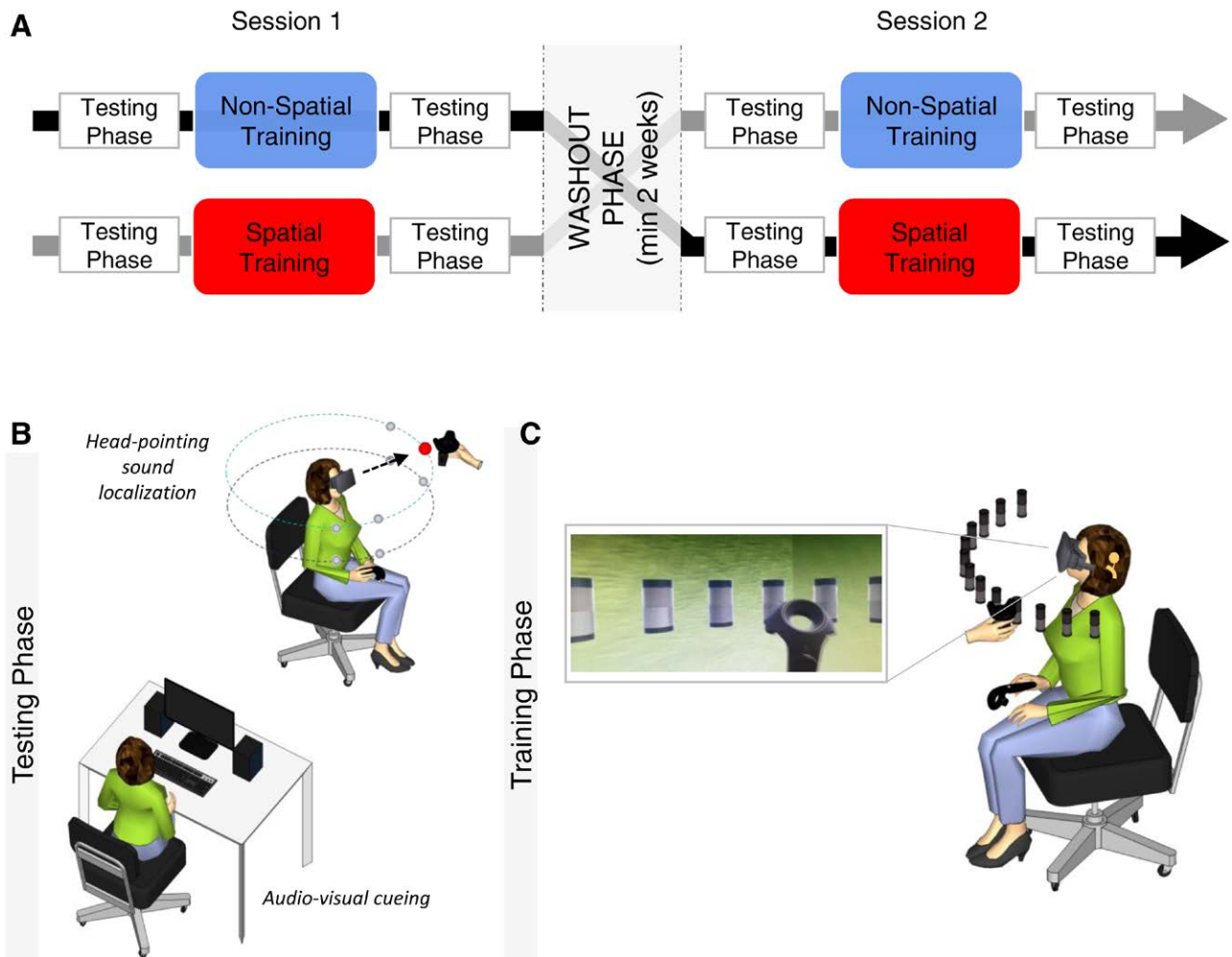


Fig. 1. Experimental procedure and setting. A, Schematic description of the overall crossover design. Each session (session 1 and session 2) comprised by two testing phases, with a training task in between. B, Testing phases. Top: schematic representation of the participant wearing the Head-mounted display (HMD) and holding the virtual reality (VR) controller during the head-pointing sound localization task. The eight spheres in front of the participant indicate predetermined speaker positions (not visible in the HMD). The hand holding the speaker represented near the red point was the experimenter's hand who move the speaker and bring it to the target position following instruction provided on a monitor. Bottom: schematic representation of the setting for the audiovisual attention-orienting task (conducted entirely outside VR). C, Training phase. Schematic representation of the participant wearing the HMD and holding the VR controller. The cylindrical speakers were visible only in VR (see inset) and indicated the possible sound positions during both training types. Inset: close-up of the scene as visible inside the HMD from participant's perspective. During the task participant saw a virtual scenario comprising a room, the speakers and the VR controller they were holding in their hand. See Video File in Supplemental Digital Content 1, <http://links.lww.com/EANDH/B44> for animated description of the two training types.

correct posture was reached, the fixation cross turned from white to blue. In the meanwhile, the experimenter placed the speaker in one of the possible eight predetermined positions, using visual indications provided on a dedicated monitor (see Verdelet et al. 2019 and Gaveau et al. 2020 for a validation of this method and Valzolgher et al. 2020a,c for examples). The predetermined positions (shown in the top part of Fig. 1B) result from the combination of 4 different azimuths in front space (-67.5° , -22.5° , 22.5° , or 67.5° ; with respect to the participant's body midline), a single distance (55 cm from the center of the head) and 2 different elevations (5° above and 15° below ear level). Target elevation was changed only with the purpose of introducing variability in the auditory stimulation, and the asymmetrical arrangement elevation positions resulted from the fact that the software for placing the speakers at predetermined positions was based on the position of the VIVE tracker located above the speaker, rather than the position of the

actual speaker. Crucially, the target sound was delivered only when the participant was in the correct start posture, and when the loudspeaker was positioned within a sphere with diameter of 5 cm centered on the predetermined location. This ensured that each stimulation position was fully replicable at the beginning of each trial. The target sound consisted of 3 sec of white noise, amplitude-modulated at 2.5 Hz, and delivered at about 65 dB SPL as measured from the participant's head. Participants were instructed to indicate sound position using their heads as a pointer. Note that head and trunk movements remained unconstrained during and after sound emission, allowing spontaneous active listening behavior. Note that participants were informed about their head projection due to a marker (white cross) and that they were only allowed to validate their response after the sound ended, by pressing a button on their handheld controller. To emphasize this instruction, the head projection marker (the cross) was white throughout the stimulation phase and

TABLE 1. Biographic and anamnestic and clinical characteristics of CI participants

ID	Sex	Age, y	Etiology of deafness	Age at deaf diagnosis (y;m)	Age at first hearing aid (y;m)	Age at implantation (right ear)	Age at implantation (left ear)	Years with one CI	Years with two CIs
1	F	22	Unknown	0;5	0;7	2	15	13	7
2	F	65	Unknown	45	49	54	53	1	11
3	F	60	Genetic	10	10	40	51	11	9
4	F	46	Meningitis	1	5	39	44	5	2
5	F	47	Genetic	10	12	42	43	1	4
6	M	55	Genetic/sudden deafness	1	44	50	52	2	3
7	M	41	Ototoxicity	2	6	34	35	1	6
8	F	47	Genetic	16	21	42	39	3	5
9	M	32	Meningitis	2	2	3	28	25	4
10	M	25	Unknown	3	3	18	21	3	4
11	F	49	Unknown	13	30	39	39	0	10
12	M	59	Head trauma	52	No	53	53	0	6
13	F	47	Genetic	3	3	45	42	3	2
14	F	54	Genetic	42	42	50	51	1	3
15	M	24	Genetic	2;6	3	18	12	6	6
16	M	65	Otosclerosis	35	55	59	60	1	5
17	F	32	Otosclerosis of the cochlea	10	12	23	25	2	7
18	F	38	Genetic	13	18	26	30	4	8
19	F	51	Unknown	6	6	43	46	3	5
20	F	52	Genetic	6	23	43	51	8	1

CI indicates cochlear implant; F, female; M, male.

turned red only when response validation was allowed. Since the speaker was invisible in VR, participants did not have visual cues about its position. Each head-pointing sound localization task comprised 40 trials (five repetitions, for each of the eight sound positions) delivered in random order. Five practice trials were also introduced but discarded from the analyses. The task lasted approximately 15 min. See the Materials for details about the apparatus and general setting in Supplemental Digital Content 2, <http://links.lww.com/EANDH/B43>.

Audiovisual Attention-Orienting Task

This task was carried out entirely outside VR. Participants sat at a table placed inside the same room in which the VR experiments are carried out. Unlike the VR tasks, in this audiovisual attention-orienting task participants rested their heads on a chinrest facing the screen where visual stimuli were displayed; hence, head movements were not allowed. Participants were instructed to indicate the elevation of a visual disk (i.e., appearing on the screen either above or below their eye level) as quickly and accurately as possible. Crucially, in each trial a task-irrelevant sound was presented simultaneously from one of the loudspeakers flanking the screen, either on the same side as the visual disk (spatially congruent trials), or from the opposite side (spatially incongruent trials). Each trial started with a white fixation cross appearing in the center of the screen and remaining visible until response. After a random delay (450–600 ms), an auditory stimulus (white noise, duration 100 ms) was emitted from one of the two loudspeakers, at approximately 60 dB SPL as measured from head position. After 100 ms from sound delivery, the visual target was presented. This consisted of a white disk (20 pixels radius, 0.5° of visual angle), appearing on a black background for 140 ms 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). In half of the trials, the visual target and the sound appeared on the same hemispace (congruent trials), whereas on the remaining half, the visual target and the sound appeared on opposite hemispaces (incongruent trials). Participants were asked to keep their gaze towards central fixation throughout the task and to indicate as quickly and accurately as possible the elevation of the visual target. Up/down responses were given using the up/down arrows keys on an external keyboard, using the index/middle fingers of the right hand (Fig. 1B). Participants had a timeout of 2000 ms to give their answer and received feedback on accuracy (percentage of correct responses) and mean response time (in ms) only at the end of each block. Importantly, they were also explicitly informed that sounds were always entirely task-irrelevant. The audiovisual attention-orienting task comprised 128 trials (16 repetitions, for each of the four disk positions and each sound side, divided in two blocks of 64 trials each) delivered in random order. Eight practice trials were also introduced but discarded from the analyses. The task lasted approximately 10 min. See the Materials for details about the apparatus and general setting in Supplemental Digital Content 2, <http://links.lww.com/EANDH/B43>.

Training Tasks

The training phases were carried out entirely in VR and took place in the same virtual room used for the other tasks detailed above. During both training tasks, the participant saw 13 virtual loudspeakers inside the Head-mounted display. The loudspeakers were arranged in a semicircle centered on the head (radius 50 cm), at –5° below ear level (due to the uncorrected offset between the tracker and the center of the speaker), and spanned between ±72° (with 12° of separation between each of them; see Fig. 1C and Video File in Supplemental Digital Content 1, <http://links.lww.com/EANDH/B44>). The target stimulus consisted of a white noise, with two possible amplitude modulations (2 or

3 Hz). Targets sounds were delivered from each of the virtual speaker in random order (12 repetition for each of the 13 speakers, resulting in 156 trials overall, divided into three blocks of 52 trials each). In each trial, a single sound was presented from one of the speakers. Sounds were delivered when the experimenter brought the real speaker into correspondence with one of the virtual ones, following a computer-controlled procedure based on continuous position tracking. Crucially, the same stimuli were delivered in the two training tasks, thus making the auditory component of the trainings identical in this respect (i.e., sounds with two amplitude modulations were delivered from positions that changed on a trial-by-trial basis). In addition, both training tasks involved a similar motor response: a reaching movement. However, in the spatial training the reaching movement served to indicate the perceived sound position, whereas in the nonspatial training served to indicate the perceived amplitude modulation in the target sound. Participants always held the controller with their right hand. At the beginning of each trial, the controller was positioned on their sternum. See the Materials in Supplemental Digital Content 2, <http://links.lww.com/EANDH/B43> for details about the apparatus and general setting.

Spatial Training

Participants were instructed to reach the speaker emitting the sound using the handheld VR controller. The sound lasted until the participant reached and “touched” the correct speaker. If they reached the wrong speaker, the correct speaker started flashing while the sound continued to play. Specifically, a series of red concentric circles (1024×1024 px, 2 circles per second) expanded intermittently, irradiating from the correct sound source (see Video File in Supplemental Digital Content 1, <http://links.lww.com/EANDH/B44>). The rationale for using wide concentric circles was to capture the participants’ attention even in case they were looking away from the correct sound source (e.g., toward the opposite hemisphere). The visual feedback and the target sound stopped only when the subject reached the correct speaker position with the controller. This has two implications: first, a sense of agency was associated with the correct response (Valzolgher et al. 2020a,b,c); second, whenever the wrong speaker was originally selected, a combination of visual and auditory signals guided the participant to the correct sound source. To enhance the feeling that the virtual speaker was touched, the controller always produced a brief vibration when coming into “contact” with one of the virtual speakers. This did not serve as feedback for performance, because the vibration occurred irrespective of whether the response was correct or incorrect. To prevent participants from colliding with the real speaker (held by the experimenter), the virtual speakers were always presented with a 5-cm offset in depth. The task lasted approximately 45 min (considering the breaks).

Nonspatial Training

Participants were instructed to identify whether the target sound contained a fast (3 Hz) or slow (2 Hz) amplitude modulation. For fast amplitude-modulated sounds, participants directed the controller in front of them, aiming to touch an invisible virtual button above the central speaker. For slow amplitude-modulated sounds, participants reached toward an invisible virtual button below the same central speaker instead. The position and effects of these virtual buttons was demonstrated to participants during the practice trials. A vibration from the controller always indicated that one of two buttons was correctly reached. As in the Spatial Training feedback procedure, a visual feedback was delivered in the case of wrong responses. This was a

series of red concentric semicircles that expanded intermittently from above or below the central speaker to indicate the correct response. Recall that target sounds were presented from different spatial positions also during the nonspatial training, but this spatial information was totally task-irrelevant (see Video File in Supplemental Digital Content 1, <http://links.lww.com/EANDH/B44>). The task lasted approximately 20 min (including breaks).

Statistical Analysis

Linear mixed-effect (LME) modeling was used for all statistical analyses. Statistical analyses were run using R (version 1.0.143). For the LME model, we used the R-packages *emmean*, *lme4*, *lmerTest* in R Studio (Bates et al. 2015; Fox & Weisberg 2020). The R-package *car* was used to obtain deviance tables from the LME models. When appropriate, we calculated Cohen’s d_{av} as index of effect size (Lakens 2013). Data and R codes can be retrieved from osf.io/cqtmtd.

The dependent variables considered in our analyses were sound localization azimuth error for the head-pointing sound localization task and for spatial training, and reaction time for the audiovisual attention-orienting task. Reaction time were instead analyzed for the audiovisual attention-orienting task. To study head movements, we extracted two dependent variables: number of head rotations and head-rotation extent (Valzolgher et al. 2020c).

RESULTS

Changes in Sound Localization and Head Movements During Spatial Training

Sound localization absolute error in azimuth at the beginning of the study was $24.1^\circ \pm 15.6^\circ$ (head-pointing sound localization before training in the first session), matching previous studies in documenting a low auditory spatial resolution, for example, when compared to normal-hearing performance (Kerber & Seeber 2012; Jones et al. 2014).

To examine if sound localization improved during the spatial training, we studied changes in absolute error in the horizontal dimension across successive trials using an LME model with trial number as fixed effect (Valzolgher et al. 2020c). To account for the variability related to individual participants and session in which the Spatial training was completed (first or second), we also included the participant (intercept and slope) and session (intercept) as random effects in the model. As shown in Figure 2A, BCI users reduced their absolute error across trials (main effect of trial number, $X^2(1) = 15.84$, $p < 0.001$). Performance in the nonspatial training was near ceiling (mean number of errors = 1.4%; see cumulative discrimination error across participants in Fig. 2B). During the nonspatial training, participants were also faster in completing the trial compared to the spatial training (nonspatial: mean \pm SD = 1.68 ± 0.25 sec; spatial training: mean \pm SD = 9.62 ± 6.99 ; $t(19) = 5.15$, $p < 0.001$ on paired t-test; Cohen’s $d_{av} = 2.19$).

During training the head was unrestrained and the target lasted until correct response, allowing active listening behavior (i.e., spontaneous head movements). To study head movements, we extracted two dependent variables: the number of head rotations and head-rotation extent (Valzolgher et al. 2020c) (see Materials for details in Supplemental Digital Content 2, <http://links.lww.com/EANDH/B43>). As expected by design, head-rotation patterns differed in the two training protocols: more head rotations occurred in the spatial (5.72 ± 3.24)

compared to the nonspatial training (1.15 ± 0.41 , $t(19) = 5.72$, $p < 0.001$ on paired t-test); furthermore, the extent of head rotations was larger during the spatial ($89.06^\circ \pm 41.86^\circ$) compared to the nonspatial training ($4.00^\circ \pm 1.91^\circ$, $t(19) = 9.20$, $p < 0.001$ on paired t-test). Interestingly, head-behavior evolved during the course of the spatial training tasks. BCI users progressively adapted their spontaneous head movements as a function of sound eccentricity: the number of head rotations diminished for central sounds (two-way interaction, $X^2(1) = 5.35$, $p = 0.02$), while head-rotation extent increased for more peripheral sounds ($X^2(1) = 4.78$, $p = 0.03$; see Materials in Supplemental Digital Content 2, <http://links.lww.com/EANDH/B43> and Fig. 3).

Effects of Spatial Training on Head-Pointing Sound Localization

Figure 3A shows absolute localization errors in head-pointing sound localization across the two sessions of the experiment. The black line and circles indicate BCI users who performed the spatial training before the nonspatial training; the gray line and

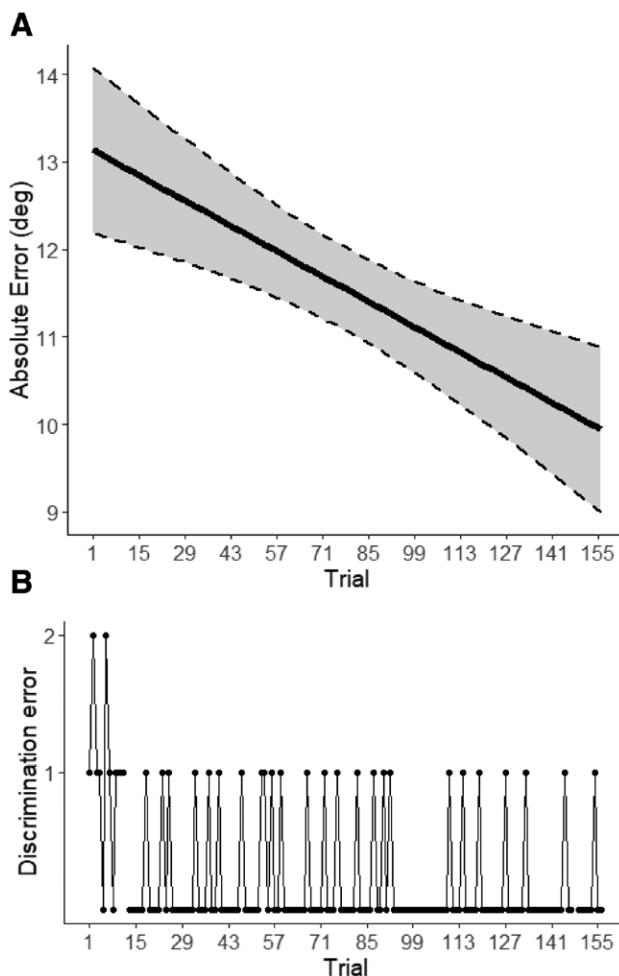


Fig. 2. Participants' performance during training. A, Reduction of absolute error in the horizontal dimension, as a function of trial in the spatial training. Linear regression (solid line), with 95% confidence intervals (dashed lines). B, Cumulative discrimination error across participants as a function of trials in the nonspatial training (i.e., number of participants who made a mistake in the trial).

circles indicate BCI users who completed the training in the reversed order. The spatial training improved performance to a greater extent compared to the nonspatial training, irrespective of the session in which it was completed.

To study the effect of spatial training on head-pointing sound localization, we entered absolute error in azimuth into an LME analysis with phase (pre or post training), training (spatial or nonspatial) and azimuth as fixed effects. We also included participant (intercept and slope) and testing session (intercept) as random effects, to account for the variability across participants and the order of training type. We found a main effect of training ($X^2(1) = 4.89$, $p = 0.03$). Crucially, the 2-way interaction between phase and training reached significance ($X^2(1) = 24.14$, $p < 0.001$). The absolute error in azimuth decreased after the spatial (Pre: 25.2 ± 16.8 ; Post: 19.3 ± 11.1 ; $t = 5.47$, $p < 0.001$, Cohen's $d_{av} = 0.40$, Fig. 3B, C), but not after the nonspatial training (Pre: 22.8 ± 14.7 ; Post: 24.5 ± 17.2 , $t = 1.48$, $p = 0.14$, Cohen's $d_{av} = 0.11$, Fig. 3D, E). Density plots of raw data (responses) in azimuth as a function of target position and training are shown in Figure S1 in Supplemental Digital Content 2, <http://links.lww.com/EANDH/B43>.^{*} For completeness, we also report the analyses on absolute error in elevation (Table 1 in Supplemental Digital Content 2, <http://links.lww.com/EANDH/B43>).

To explore if participants who performed the spatial training in the first session ($N = 10$) maintained the improvement after the 2-weeks washout, we entered on their absolute errors in azimuth during the pretraining phase into an LME analysis with time (session 1 versus session 2) as fixed effects. Errors in the pre training phase were smaller on session 2 compared to session 1 ($X^2(1) = 4.64$, $p = 0.03$). Albeit this effect was small (3°), it may provide initial evidence that the error reduction induced by the Spatial training may persist after the 2-week washout.

Following Isaiah et al. (2014) we also studied immediate head-orienting movements (i.e., movements occurring around 0.90 sec after the beginning of the sound), which reflect the first spontaneous orienting response toward the sound. Saccade-like head-orienting response were more accurate after the spatial training, as revealed by the significant interaction between training and phase on response density at target location ($F(1,19) = 6.52$, $p = 0.02$, $\eta^2 = 0.11$; see Materials and Figure S1 in Supplemental Digital Content 2, <http://links.lww.com/EANDH/B43>).

Finally, we examined if training benefits were modulated by clinical variables. We correlate the training effect (difference in absolute error before and after the spatial training) with the participant's age at the moment of deafness diagnosis, years spent with one CI only and years spent with two CIs. No significant correlation emerged (all $ps > 0.35$).

^{*}To examine to what extent changes in the amount of front-back confusions during the head-pointing task could impact on the observed training effects, we entered front-back confusions as variable in an LME with training and phase as fixed effects. We found a training by phase interaction ($X^2(1) = 8.61$, $p = 0.003$). Before training, front-back confusions did not differ between spatial (12.2%) and nonspatial conditions (11.8%). After training, there were less front-back confusions in the spatial (4.6%) compared to the nonspatial condition (9.3%). Yet, this front-back confusion reduction did not fully explain our training related improvement. When we repeated the analysis of absolute azimuth error after removing all trials in which a front-back confusion was present, a training \times phase interaction was still present ($X^2(1) = 14.64$, $p < 0.001$).

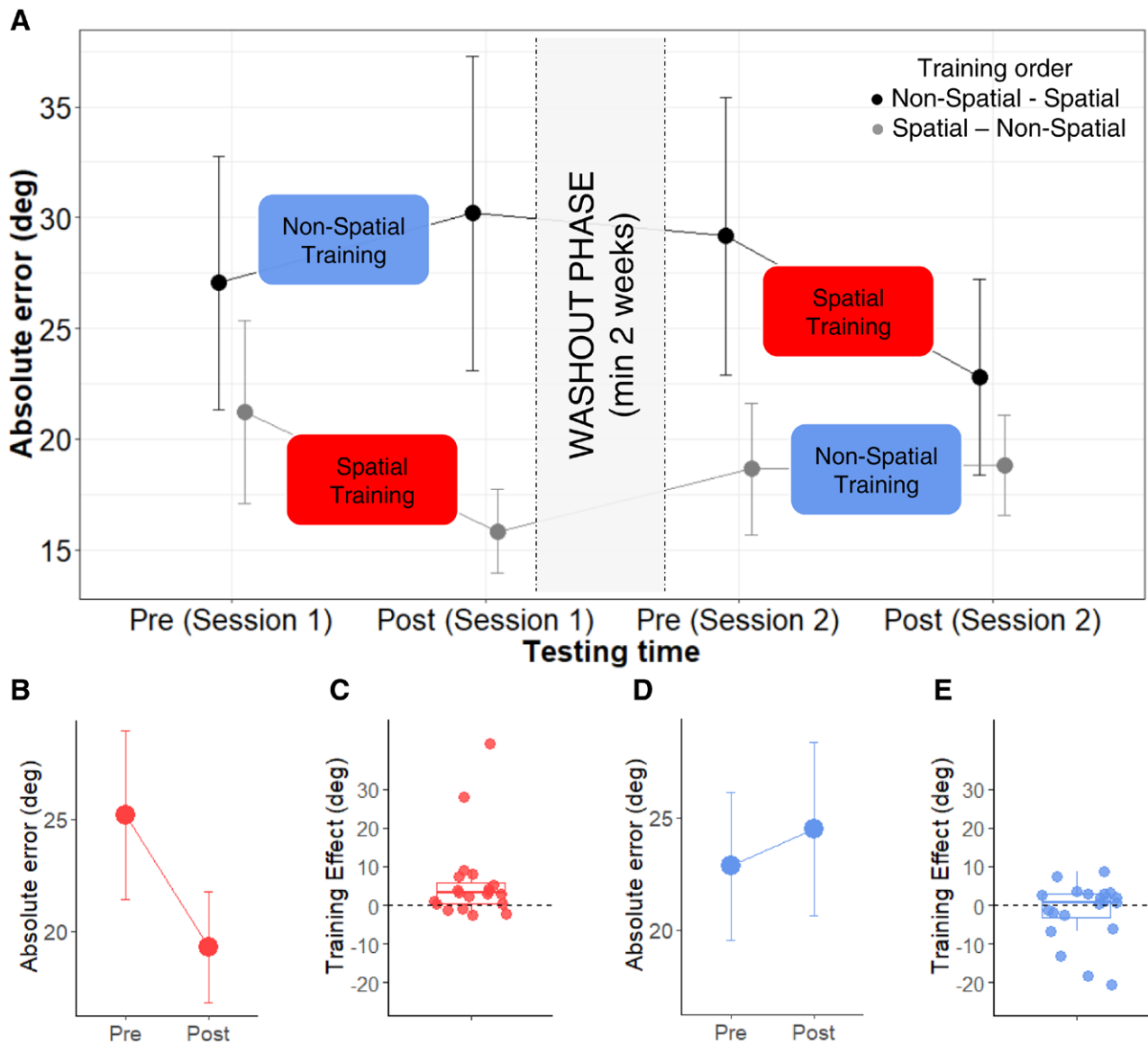


Fig. 3. Performance in head-pointing sound localization. A, Absolute localization across the four testing sessions of the experimental design, separately for participants who completed the Spatial training on session 1 (gray) or session 2 (black line). Error bars represent standard errors. B–D, Absolute errors (in degrees) as a function of phase before and after the spatial (red, B) and nonspatial (blue, D) training. C–E, Training effect (deg): the difference between pre and post training session (pre-post) in absolute error as function of training task (C, spatial in red and E, nonspatial in blue). Positive values in (C) and (E) represent improvements after the training.

Audiovisual Attention-Orienting Task

To examine the effects of our training protocols on the audiovisual attention-orienting task, we entered the response time (ms) into an LME analysis with testing phase (pre or post training), training type (spatial or nonspatial), and congruency between sound and visual target position (congruent or incongruent). Again, participant number and session (intercepts) were treated as random effects.

We found a main effect of congruency ($X^2(1) = 30.45$, $p < 0.001$) caused by faster visual discrimination responses when the auditory cue matched target side (congruent: $M = 291.2$ ms, $SD = 79.2$ ms) compared to when it occurred on the opposite side of space (incongruent trials: $M = 307.3$ ms, $SD = 79.4$ ms, Cohen's $d_{av} = 0.20$). We also found a 3-way interaction between testing phase, training type, and sound position ($X^2(1) = 4.90$, $p = 0.03$). After the nonspatial training, audiovisual attention-

orienting effect decreased (Pre: $M = 23.1$ ms, $SD = 17.8$ ms; Post: $M = 10.9$ ms, $SD = 20.5$ ms; $X^2(1) = 4.25$, $p = 0.04$, Cohen's $d_{av} = 0.64$), while it persisted after the spatial one (Pre: $M = 11.9$ ms, $SD = 17.5$ ms; Post: $M = 18.4$ ms, $SD = 19.6$ ms; $X^2(1) = 1.15$, $p = 0.28$, Cohen's $d_{av} = 0.35$).

DISCUSSION

In the present study, we examined if the ability of bilateral CI users to localize sounds in the horizontal plane improves as a result of training. Specifically, we implemented a spatial training based on reaching-to-sound, audiovisual feedback, and free-head movements. During this training, we found that sound localization in BCI users improved rapidly across trials. After training, sound localization ameliorated also in an auditory spatial task performed using a different motor response with respect

to the trained task (head-pointing instead of hand-reaching) and different sound-source positions. This generalization was significantly more pronounced compared to the one observed in a nonspatial control training task.

Our findings add to the clinical population of CI users training effects observed in animal models (Isaiah et al. 2014) and in normal-hearing human listeners (Strelnikov et al. 2011; Rabini et al. 2019; Steadman et al. 2019; Valzolgher et al. 2020b). Most notably, they go beyond early reports on training effects in CI users (Luntz et al. 2005) in two ways. First, they originate from a crossover experimental design, which allowed testing each participant in both training types, thus excluding potential confounding group differences when assessing training efficacy. Second, they show generalization of training effects to a different auditory spatial task. We additionally provided initial evidence that sound localization benefits partially persisted after the 2-week washout in participants who trained in the spatial task during the first session. This effect was, however, numerically small (3° on average); thus, it remains to be ascertained how long this improvement can really persist over time, and paraps how much this transfer to daily life.

Finally, our study also probed training generalization effects in a task that required *implicit* sound localization, that is, the audiovisual attention-orienting task. We found that lateralized auditory cues were effective in orienting BCIs' visual attention. This finding is novel and adds to previous work in unilateral CI users, in which this audiovisual attention-orienting ability was absent (Pavani et al. 2017). Yet, no change in this task was observed after the spatial training. In fact, we observed a reduction of the attention-orienting effect after the nonspatial training. While this finding was unexpected and it will not be discussed further, it may suggest that repeated processing of nonspatial (rather than spatial) sound attributes could impact the ability to exploit sound position when orienting visual attention, at least in this clinical population.

While the spatial training produced a clear within-subject improvement, the BCI users in our study were fitted with different CI processors. This implies that the auditory cues available to each participant likely differed based on the various device brand and settings. Most notably, AGC types (i.e., which modulates the acoustic signal in amplitude, to compensate the reduced dynamic range of stimulated hearing provided by the implant), vary as a function of the manufacturer. AGC may distort ILDs across frequencies (which are fundamental for sound localization in CI users; see Dorman et al. 2014) in a more or less aggressive way as a function of the processor (and manufacturer) type and of the settings made by the hearing care professional. In the present experiment, the lack of synchronization of AGCs between bilateral CI processors (Pastore et al. 2021; Dwyer et al. 2021), the different duration of sound stimuli as a function of response time (Boyle et al. 2009), and the possible AGC distortions as a function of head movements (Archer-Boyd & Carlyon 2019; see also Archer-Boyd & Carlyon 2021) could have substantially distorted the available ILDs. Yet, irrespective of the modulations introduced by the CI processor our spatial training was robust enough to allow generalization of training effects from the training task (in which sound duration continued until response, 9.6 sec on average), to a test task (in which the sound lasted only 3 sec).

Our instructions allowed free-head movement during sound, and our VR apparatus permitted continuous measuring

of head movements. Several findings arose from this approach. First, during the spatial training, we observed that the number of head rotations diminished for central sounds compared to peripheral, across trial repetitions. BCI users appear to have learned to distinguish if sound came from the center or the periphery and thus, when sound sources were central, they did not need to move their heads, corroborating the improvement revealed by reduction of absolute errors across trials. Second, during the test task, we observed a training related improvement also in terms of the first head movement-orienting response. These immediate and spontaneous orienting movements of the head can reflect early encoding of sounds positions (Muir & Field 1979) and can be considered an implicit index of participants' ability to identify sound direction (also described as saccade-like head-response by Isaiah et al. 2014). Immediate head-orienting responses were more accurate in the spatial compared to the nonspatial training. This finding supports the training benefit documented using absolute errors as a dependent variable. Finally, a third observation concerns the potential role of head movements as effective behaviors strategies when relearning sound localization abilities (Valzolgher et al. 2020c). We showed that head movements in BCI users changed across trial repetition during the spatial training. Although this was not correlated to changes in performance (see Materials in Supplemental Digital Content 2, <http://links.lww.com/EANDH/B43>), this preliminary observation suggests that BCI participants may also adapt their head-rotation behavior when relearning their spatial hearing skills, by making it more purposeful across trials. This finding adds to the literature showing that, compared to static listening conditions, moving the head while listening to sounds improves sound localization not only in normal-hearing participants (Pollack & Rose 1967; Perrett & Noble 1997a,b; Wightman & Kistler 1999; Vliegen et al. 2004; Brimijoin & Akeroyd 2012; Honda et al. 2013; McAnally & Martin 2014), but also in BCI users (Coudert et al. 2022; Mueller et al. 2014; Pastore et al. 2018). Future studies on spatial hearing relearning should examine further the spontaneous self-regulatory behaviors that participants implement during active listening, by adopting unconstrained head or even trunk movements and long-lasting sounds, as here.

Limitations

Two methodological issues should be considered. First, our training protocols are likely not equally complex. Although neither task was speeded, while BCI users were relatively fast in responding to amplitude modulations in the target sound (1.7 sec on average), in the sound localization training they listened to each sound for a longer period before responding (9.6 sec). This raises the possibility that the greater efficacy of the spatial training could also reflect, to some extent, longer exposure to the sound. Second, head-pointing to sounds entails a motor response (i.e., head-turning) that is typically coupled with manual reaching movements. This may imply a partial overlap between the motor responses we adopted for the training and testing tasks. Future studies could examine the effects of a training performed with frontal and near sounds (as here), on test stimuli coming in back and/or far space. Changes in the acoustic features of training and test stimuli could also be implemented (syllables or words, as in Zheng et al. 2015, instead of broad band sounds as used here).

CONCLUSION

Using a novel spatial training approach based on reaching-to-sounds, audiovisual feedback, and free-head movements during listening, we showed that training auditory skills in BCI users is possible. This improvement can be accounted for in terms of central cognitive processes and/or self-regulation strategies since the auditory input from the CI remained largely unchanged. Furthermore, it stresses the positive contribution that multisensory and motor processes can play in adapting to CI. While our approach to training proved successful, it clearly originated from an experimental more than a translational perspective. To assess the implications of our finding for clinical settings and real life, it would be important to expand our approach by testing the efficacy of the principles subtending our Spatial training by extending its duration and by testing further its generalization to a wider range of spatial locations and stimuli. Ultimately, the main aim remains to enhance spatial hearing in real life for CI users. Our findings can be seen as a first step towards expanding training approaches from speech perception alone to spatial hearing, after implantation. In this respect, they contribute to enrich the domain of actions that clinicians and patients themselves can undertake to improve their experience with the CIs.

ACKNOWLEDGMENTS

We are grateful to all the participants who accepted to participate in this study and to the three anonymous reviewers for comments on a previous version of the manuscript. We thank Jean-Louis Borach and Sonia Alouche for administrative support of this project and Eric Koun for technical support. We thank Giordana Torresani for her contributions to Figure 1.

C.V., F.P., and A.F. were supported by a grant of the Agence Nationale de la Recherche (ANR-16-CE17-0016, VIRTUALHEARING3D, France). C.V. was also supported by a grant of the Università Italo-Francesca/Università Franco-Italiana, the Zegna Founder's Scholarship and Associazione Amici di Claudio Demattè. F.P. and A.F. were also supported by a prize of the Foundation Medisite (France), by the Neurodis Foundation (France), by a grant from the Italian Ministry for Research and University (MUR, PRIN 20177894ZH to F.P.) and grants from IHU CaSaMe ANR-10-UBHU-0003 and ANR 2019CE37 Blind Touch (to A.F.).

C.V., F.P., and A.F. conceptualized and design the experiment and managed funding acquisition and project administration. C.V., J.G., S.B., S.G., and M.T. performed the experiment. C.V. and F.P. analyzed data and wrote the paper. G.V. and R.S. worked on software. All authors discussed the results and implications and commented on the manuscript. F. P., A.F., and E.T. supervised the project and managed funding.

The authors have no conflicts of interest to declare.

Address for correspondence: Chiara Valzolgher, Palazzo Fedrigotti - Corso Bettini 31, 38068 Rovereto (TN), Italy. E-mail: chiara.valzolgher@unitn.it

Received August 12, 2021; accepted May 31, 2022

REFERENCES

- Allen, M. C., Nikolopoulos, T. P., O'Donoghue, G. M. (1998). Speech intelligibility in children after cochlear implantation. *Am J Otol*, *19*, 742–746.
- Archer-Boyd, A. W., & Carlyon, R. P. (2019). Simulations of the effect of unlinked cochlear-implant automatic gain control and head movement on interaural level differences. *J Acoust Soc Am*, *145*, 1389.
- Archer-Boyd, A. W., & Carlyon, R. P. (2021). Further simulations of the effect of cochlear-implant pre-processing and head movement on interaural level differences. *J Acoust Soc Am*, *150*, 506.
- Aronoff, J. M., Yoon, Y. S., Freed, D. J., Vermiglio, A. J., Pal, I., Soli, S. D. (2010). The use of interaural time and level difference cues by bilateral cochlear implant users. *J Acoust Soc Am*, *127*, EL87–EL92.
- Asp, F., Mäki-Torkko, E., Karltorp, E., Harder, H., Hergils, L., Eskilsson, G., Stenfelt, S. (2012). Bilateral versus unilateral cochlear implants in children: Speech recognition, sound localization, and parental reports. *Int J Audiol*, *51*, 817–832.
- Bates, D., Maechler, M., Bolker, B. M., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *J Statist Softw*, *67*, 1–48.
- Brimjoin, W. O., & Akeroyd, M. A. (2012). The role of head movements and signal spectrum in an auditory front/back illusion. *Iperception*, *3*, 179–182.
- Boyer, E. O., Babayan, B. M., Bevilacqua, F., Noisternig, M., Warusfel, O., Roby-Brami, A., Hanneton, S., Viaud-Delmon, I. (2013). From ear to hand: The role of the auditory-motor loop in pointing to an auditory source. *Front Comput Neurosci*, *7*, 26.
- Boyle, P. J., Büchner, A., Stone, M. A., Lenarz, T., Moore, B. C. (2009). Comparison of dual-time-constant and fast-acting automatic gain control (AGC) systems in cochlear implants. *Int J Audiol*, *48*, 211–221.
- Calmels, M. N., Saliba, I., Wanna, G., Cochard, N., Fillaux, J., Deguine, O., Fraysse, B. (2004). Speech perception and speech intelligibility in children after cochlear implantation. *Int J Pediatr Otorhinolaryngol*, *68*, 347–351.
- Carlile, S. (2014). The plastic ear and perceptual relearning in auditory spatial perception. *Front Neurosci*, *8*, 237.
- Cohen, Y. E., & Andersen, R. A. (2002). A common reference frame for movement plans in the posterior parietal cortex. *Nat Rev Neurosci*, *3*, 553–562.
- Coudert, A., Gaveau, V., Gatel, J., Verdelet, G., Salemm, R., Farne, A., ... & Truy, E. (2022). Spatial Hearing Difficulties in Reaching Space in Bilateral Cochlear Implant Children Improve With Head Movements. *Ear Hear*, *43*, 192.
- Dehaene, S. (2020). *How We Learn: The New Science of Education and the Brain*. Penguin.
- Dwyer, R. T., Chen, C., Hehrmann, P., Dwyer, N. C., Gifford, R. H. (2021). Synchronized automatic gain control in bilateral cochlear implant recipients yields significant benefit in static and dynamic listening conditions. *Trends Hear*, *25*, 23312165211014139.
- Dorman, M. F., Loisselle, L., Stohl, J., Yost, W. A., Spahr, A., Brown, C., Cook, S. (2014). Interaural level differences and sound source localization for bilateral cochlear implant patients. *Ear Hear*, *35*, 633–640.
- Forli, F., Lazzarini, F., Fortunato, S., Bruschini, L., Berrettini, S. (2019). Cochlear implant in the elderly: Results in terms of speech perception and quality of life. *Audiol Neurootol*, *24*, 77–83.
- Fox, J., & Weisberg, S. (2020). Using Car and Effects Functions in Other Functions 3, pp. 1-5. <https://mran.revolutionanalytics.com/snapshot/2020-05-18/web/packages/car/vignettes/embedding.pdf>
- Gaveau, V., Coudert, A., Salemm, R., Koun, E., Desoche, C., Truy, E., Farne, A., Pavani, F. (2020). SPHERE: A novel approach to 3D and active sound localization. *bioRxiv*. Preprint posted online. March 20, 2020. doi: 10.1101/2020.03.19.998906
- Gifford, R. H., & Stecker, G. C. (2020). Binaural cue sensitivity in cochlear implant recipients with acoustic hearing preservation. *Hear Res*, *390*, 107929.
- Grantham, D. W., Ashmead, D. H., Ricketts, T. A., Labadie, R. F., Haynes, D. S. (2007). Horizontal-plane localization of noise and speech signals by postlingually deafened adults fitted with bilateral cochlear implants. *Ear Hear*, *28*, 524–541.
- Grieco-Calub, T. M., & Litovsky, R. Y. (2010). Sound localization skills in children who use bilateral cochlear implants and in children with normal acoustic hearing. *Ear Hear*, *31*, 645–656.
- Honda, A., Shibata, H., Hidaka, S., Gyoba, J., Iwaya, Y., Suzuki, Y. (2013). Effects of head movement and proprioceptive feedback in training of sound localization. *Iperception*, *4*, 253–264.
- Isaiah, A., Vongpaisal, T., King, A. J., Hartley, D. E. (2014). Multisensory training improves auditory spatial processing following bilateral cochlear implantation. *J Neurosci*, *34*, 11119–11130.
- Jones, H., Kan, A., Litovsky, R. Y. (2014). Comparing sound localization deficits in bilateral cochlear-implant users and vocoder simulations with normal-hearing listeners. *Trends Hear*, *18*, 2331216514554574.
- Keating, P., Dahmen, J. C., King, A. J. (2015). Complementary adaptive processes contribute to the developmental plasticity of spatial hearing. *Nat Neurosci*, *18*, 185–187.
- Kerber, S., & Seeber, B. U. (2012). Sound localization in noise by normal-hearing listeners and cochlear implant users. *Ear Hear*, *33*, 445–457.
- Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Front Psychol*, *4*, 863.

- Litovsky, R. Y., Parkinson, A., Arcaroli, J. (2009). Spatial hearing and speech intelligibility in bilateral cochlear implant users. *Ear Hear*, *30*, 419–431.
- Luntz, M., Brodsky, A., Watad, W., Weiss, H., Tamir, A., Pratt, H. (2005). Sound localization in patients with unilateral cochlear implants. *Cochlear Implants Int*, *6*, 1–9.
- McAnally, K. I., & Martin, R. L. (2014). Sound localization with head movement: Implications for 3-d audio displays. *Front Neurosci*, *8*, 210.
- Mendonça, C. (2014). A review on auditory space adaptations to altered head-related cues. *Front Neurosci*, *8*, 219.
- Moore, D. R., & Shannon, R. V. (2009). Beyond cochlear implants: Awakening the deafened brain. *Nat Neurosci*, *12*, 686–691.
- Moua, K., Kan, A., Jones, H. G., Misurelli, S. M., Litovsky, R. Y. (2019). Auditory motion tracking ability of adults with normal hearing and with bilateral cochlear implants. *J Acoust Soc Am*, *145*, 2498.
- Mueller, M. F., Meisenbacher, K., Lai, W. K., Dillier, N. (2014). Sound localization with bilateral cochlear implants in noise: How much do head movements contribute to localization? *Cochlear Implants Int*, *15*, 36–42.
- Muir, D., & Field, J. (1979). Newborn infants orient to sounds. *Child Dev*, *50*, 431–436.
- Pastore, M. T., Natale, S. J., Yost, W. A., Dorman, M. F. (2018). Head movements allow listeners bilaterally implanted with cochlear implants to resolve front-back confusions. *Ear Hear*, *39*, 1224–1231.
- Pastore, M. T., Pulling, K. R., Chen, C., Yost, W. A., Dorman, M. F. (2021). Effects of bilateral automatic gain control synchronization in cochlear implants with and without head movements: Sound source localization in the frontal hemifield. *J Speech Lang Hear Res*, *64*, 2811–2824.
- 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. *Hear Res*, *344*, 24–37.
- Perrett, S., & Noble, W. (1997a). The contribution of head motion cues to localization of low-pass noise. *Percept Psychophys*, *59*, 1018–1026.
- Perrett, S., & Noble, W. (1997b). The effect of head rotations on vertical plane sound localization. *J Acoust Soc Am*, *102*, 2325–2332.
- Pollack, I., & Rose, M. (1967). Effect of head movement on the localization of sounds in the equatorial plane. *Percept Psychophys*, *2*, 591–596.
- Rabini, G., Altobelli, E., Pavani, F. (2019). Interactions between egocentric and allocentric spatial coding of sounds revealed by a multisensory learning paradigm. *Sci Rep*, *9*, 7892.
- Roskoth-Kuhl, N., Buck, A. N., Li, K., Schnupp, J. W. (2021). Microsecond interaural time difference discrimination restored by cochlear implants after neonatal deafness. *Elife*, *10*, e59300.
- Seeber, B. U., & Fastl, H. (2008). Localization cues with bilateral cochlear implants. *J Acoust Soc Am*, *123*, 1030–1042.
- Senn, S. S. (2002). *Cross-Over Trials in Clinical Research* (Vol. 5). John Wiley & Sons.
- Smulders, Y. E., Van Zon, A., Stegeman, I., Rinia, A. B., Van Zanten, G. A., Stokroos, R. J., Hendrice, N., Free, R. H., Maat, B., Frijns, J. H. M., Briare, J. J., Mylanus, E. A. M., Huinck, W. J., Smit, A. L., Topsakal, V., Tange, R. A., Grolman, W. (2016). Comparison of bilateral and unilateral cochlear implantation in adults a randomized clinical trial. *JAMA Otolaryngol Head Neck Surg*, *142*, 249–256.
- Steadman, M. A., Kim, C., Lestang, J. H., Goodman, D. F. M., Picinali, L. (2019). Short-term effects of sound localization training in virtual reality. *Sci Rep*, *9*, 18284.
- Stickney, G. S., Zeng, F. G., Litovsky, R., Assmann, P. (2004). Cochlear implant speech recognition with speech maskers. *J Acoust Soc Am*, *116*, 1081–1091.
- Strelnikov, K., Rosito, M., Barone, P. (2011). Effect of audiovisual training on monaural spatial hearing in horizontal plane. *PLoS One*, *6*, e18344.
- Tyler, R. S., Witt, S. A., Dunn, C. C., Wang, W. (2010). Initial development of a spatially separated speech-in-noise and localization training program. *J Am Acad Audiol*, *21*, 390–403.
- Valzolgher, C., Alzhaler, M., Gessa, E., Todeschini, M., Nieto, P., Verdelet, G., Salemme, R., Gaveau, V., Marx, M., Truy, E., Barone, P., Farnè, A., & Pavani, F. (2020a). The impact of a visual spatial frame on real sound-source localization in virtual reality. *Curr Res Behav Sci*, *1*, 100003.
- Valzolgher, C., Campus, C., Rabini, G., Gori, M., Pavani, F. (2020b). Updating spatial hearing abilities through multisensory and motor cues. *Cognition*, *204*, 104409.
- Valzolgher, C., Verdelet, G., Salemme, R., Lombardi, L., Gaveau, V., Farnè, A., Pavani, F. (2020c). Reaching to sounds in virtual reality: A multi-sensory-motor approach to promote adaptation to altered auditory cues. *Neuropsychologia*, *149*, 107665.
- Valzolgher, C., Todeschini, M., Verdelet, G., Gatel, J., Salemme, R., Gaveau, V., Truy, E., Farnè, A., Pavani, F. (2022). Adapting to altered auditory cues: Generalization from manual reaching to head pointing. *PLoS One*, *17*, e0263509.
- van Hoesel, R. J. (2004). Exploring the benefits of bilateral cochlear implants. *Audiol Neurootol*, *9*, 234–246.
- Verdelet, G., Desoche, C., Volland, F., Farnè, A., Coudert, A., Hermann, R., Truy, E., Gaveau, V., Pavani, F., & Salemme, R. (2019). Assessing spatial and temporal reliability of the vive System as a tool for naturalistic behavioural research. In 2019 International Conference on 3D Immersion (IC3D) (pp. 1-8). IEEE, Brussels, Belgium. <https://doi.org/10.1109/IC3D48390.2019.8975994>
- Verhaert, N., Lazard, D. S., Gnansia, D., Bébér, J. P., Romanet, P., Meyer, B., Péan, V., Mollard, D., Truy, E. (2012). Speech performance and sound localization abilities in Neurelec Digisonic® SP binaural cochlear implant users. *Audiol Neurootol*, *17*, 256–266.
- Vliegen, J., Van Grootel, T. J., Van Opstal, A. J. (2004). Dynamic sound localization during rapid eye-head gaze shifts. *J Neurosci*, *24*, 9291–9302.
- Wightman, F. L., & Kistler, D. J. (1999). Resolution of front-back ambiguity in spatial hearing by listener and source movement. *J Acoust Soc Am*, *105*, 2841–2853.
- Wilson, B. S. (2017). The modern cochlear implant: A triumph of biomedical engineering and the first substantial restoration of human sense using a medical intervention. *IEEE Pulse*, *8*, 29–32.
- Wilson, B. S. (2019). The remarkable cochlear implant and possibilities for the next large step forward. *Acoust Today*, *15*, 53–61.
- Wilson, B. S., Finley, C. C., Lawson, D. T., Wolford, R. D., Eddington, D. K., Rabinowitz, W. M. (1991). Better speech recognition with cochlear implants. *Nature*, *352*, 236–238.
- Zheng, Y., Godar, S. P., Litovsky, R. Y. (2015). Development of sound localization strategies in children with bilateral cochlear implants. *PLoS One*, *10*, e0135790.

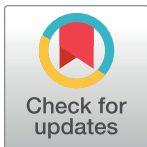
RESEARCH ARTICLE

Adapting to altered auditory cues: Generalization from manual reaching to head pointing

Chiara Valzolgher^{1,2*}, Michela Todeschini³, Gregoire Verdelet^{1,4}, Julie Gatel⁵, Romeo Saleme^{1,4}, Valerie Gaveau^{1,6}, Eric Truy⁵, Alessandro Farnè^{1,2,4}, Francesco Pavani^{1,2}

1 Integrative, Multisensory, Perception, Action and Cognition Team (IMPACT), Lyon Neuroscience Research Center, Lyon, France, **2** Center for Mind/Brain Sciences—CIMeC, University of Trento, Trento, Italy, **3** Department of Psychology and Cognitive Sciences (DiPSCo), University of Trento, Trento, Italy, **4** Neuroimmersion, Lyon Neuroscience Research Center, Lyon, France, **5** Hospices Civils de Lyon, Lyon, France, **6** University of Lyon 1, Villeurbanne, France

* chiara.valzolgher@unitn.it



OPEN ACCESS

Citation: Valzolgher C, Todeschini M, Verdelet G, Gatel J, Saleme R, Gaveau V, et al. (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>

Editor: Welber Marinovic, Curtin University, AUSTRALIA

Received: August 9, 2021

Accepted: January 21, 2022

Published: April 14, 2022

Peer Review History: PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: <https://doi.org/10.1371/journal.pone.0263509>

Copyright: © 2022 Valzolgher et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: Data and R codes can be retrieved from osf.io/dt76y.

Funding: The study was supported by a grant of the Agence Nationale de la Recherche (ANR-16-

Abstract

Localising sounds means having the ability to process auditory cues deriving from the interplay among sound waves, the head and the ears. When auditory cues change because of temporary or permanent hearing loss, sound localization becomes difficult and uncertain. The brain can adapt to altered auditory cues throughout life and multisensory training can promote the relearning of spatial hearing skills. Here, we study the training potentials of sound-oriented motor behaviour to test if a training based on manual actions toward sounds can learning effects that generalize to different auditory spatial tasks. We assessed spatial hearing relearning in normal hearing adults with a plugged ear by using visual virtual reality and body motion tracking. Participants performed two auditory tasks that entail explicit and implicit processing of sound position (head-pointing sound localization and audio-visual attention cueing, respectively), before and after having received a spatial training session in which they identified sound position by reaching to auditory sources nearby. Using a cross-over design, the effects of the above-mentioned spatial training were compared to a control condition involving the same physical stimuli, but different task demands (i.e., a non-spatial discrimination of amplitude modulations in the sound). According to our findings, spatial hearing in one-ear plugged participants improved more after reaching to sound trainings rather than in the control condition. Training by reaching also modified head-movement behaviour during listening. Crucially, the improvements observed during training generalize also to a different sound localization task, possibly as a consequence of acquired and novel head-movement strategies.

Introduction

Humans and other hearing species can localize sounds in space. Spatial hearing relies on the interpretation of binaural auditory cues, resulting from the different inputs reaching the two

CE17-0016, VIRTUALHEARING3D, France, <https://anr.fr/>) to F.P. and A.F. In addition, C.V. was supported by a grant of the Università Italo-Francese/Université Franco-Italienne (<https://www.universite-franco-italienne.org/>), the Zegna Founder's Scholarship (<https://www.zegnagroup.com/it/csr/founder-scholarship/>) and Associazione Amici di Claudio Demattè (<http://www.amicidematte.org/>). F.P. and A.F. were also supported 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) (<https://www.miur.gov.it/>). Finally, A.F. was supported by the IHU CaSaMe ANR-10-UBHU-0003 and ANR 2019CE37 Blind Touch (ANR: <https://anr.fr/>). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

ears, and monaural cues, resulting from the amplitude and spectral changes occurring in the single ear [1–3]. Although listeners experience their spatial hearing skills as constant and stable, temporary or permanent conditions can alter the auditory cues and affect this fundamental hearing ability. Examples include partial or complete hearing loss to one ear [4], age-related hearing loss [5], use of hearing aids [6] or use of cochlear implants [7]. Yet, in the last decades research has shown that the relearning of spatial hearing skills is possible, in humans [8–13] as well as in other animals [14].

Adaptation to new auditory cues can be achieved using multisensory cues to sound position (for reviews see: [8–10, 15, 16]). Training with audio-visual stimuli can be more effective than training with auditory information alone [11, 12, 17, 18]. In addition, adapting to new auditory cues may be easier through motor interactions with the sound sources, using tasks in which acoustic stimulation results from the subjects' own movements [13] or in which participants are encouraged to act towards sounds [19, 20]. For instance, participants could be asked to hit a moving sound presented in a virtual auditory space using a hand-held tool [21, 22], or to shoot audio-visual moving targets in virtual reality [23]. Other relevant examples include the work of Steadman and colleagues, in which participants were required to actively point their head toward sounds [24]. Taken together, these works suggest that motor interactions with sounds could promote adaptation to new auditory cues.

In a recent study [25], the above-mentioned issue was addressed by asking to one-ear-plugged normal hearing participants to identify the position of sounds by reaching or by naming the labels associated with each speaker. The manipulation was performed across groups, using virtual reality (VR). Seventeen virtual speakers were presented to participants in a virtual room; participants' hand movements and head rotations were monitored through kinematic tracking. Both groups received audio-visual feedback about their performance on a trial-by-trial basis. Importantly, participants in both groups were always allowed to move their head during listening, with sounds lasting until response (approximately 4 seconds). Results show a faster reduction in the number of errors in the reaching group than in the naming one. This suggests that reaching to sounds plays a specific role when listeners must adapt to new auditory cues. Moreover, the reaching group increased head exploration movements during listening, and these head-movements led to sound localization improvements. Specifically, the improvements determined by reaching to sounds were related to changes in the amount of space explored with the head. This suggests a potential role of head movements during listening in this adaptive behavior.

In this recent work [25] we also documented that reaching to sounds can ameliorate sound localization with one ear plugged on a trial-by-trial basis. However, it remains unclear if participants improved because of practice with the specific auditory task or if they learned new and effective ways to adapt to the altered auditory cues instead. If the latter answer is true, then sensorimotor training should improve sound source localization also when the task entails different sound source positions, and when it requires a response with a non-trained effector (i.e. the head rather than the hand). In the present study, this hypothesis was addressed directly by testing if performance improvement induced by reaching to sounds can extend (i.e., generalize) beyond the trained auditory task itself.

To this aim, we recruited a group of normal hearing participants and we temporarily degraded their spatial hearing by plugging one ear. We used the reaching task developed in our previous work [25] to train their sound localization. Crucially, before and after this training participants were also tested in two tasks aiming at revealing generalization effects: a head-pointing sound localization task [25] and an audio-visual attention cueing task [26]. The head-pointing localization task required to explicitly localize sounds and differed from training in terms of visual scenario (speaker position no longer visible), spatial position of the targets

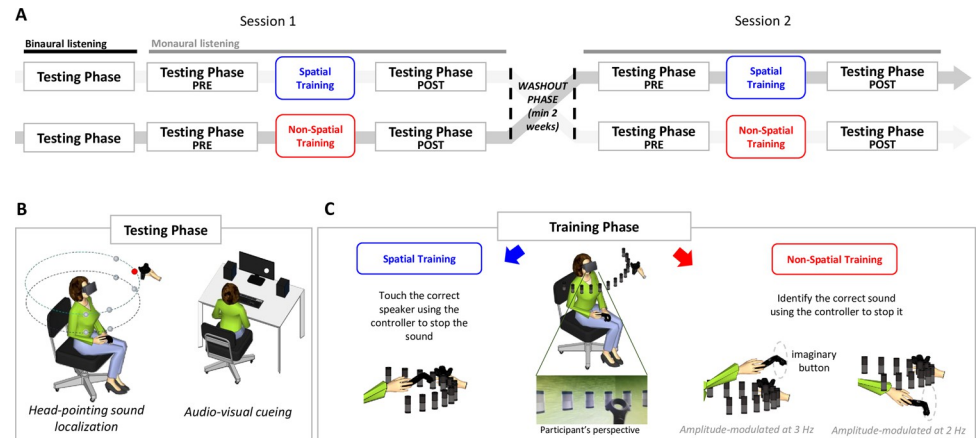


Fig 1. Experimental procedure and setting. (A) Experimental crossover protocol: all participants took part in 2 training sessions (session 1 and session 2). Each session was composed by two Testing phases (termed, 'pre' and 'post'), separated by a Training phase (Spatial or Non-Spatial, across sessions). All phases were performed in monaural listening (i.e. with one ear plugged), but an additional testing phase at the beginning of session 1 measured performance in binaural listening. (B) Testing Phases. To the left, a schematic representation of participant wearing the HMD and holding the controller used for validation during head-pointing sound localization. The eight spheres in front of the participant indicate the pre-determined speaker positions, the loudspeaker held by the experimenter's hand illustrates stimulation position in one example trial. To the right, a schematic representation of the setting for the audio-visual attention cueing task. (C) Training Phase. In the center, a schematic representation of participant wearing the HMD and holding the controller. Thirteen cylindrical visible speakers indicate the pre-determined positions during each of the training sessions. Below, a close-up of the scene as visible inside the HMD from participant's perspective: during the task participant saw the virtual room, the visible speakers and the controller they held in their hand. Cartoons on either side illustrate example of the response's movements performed during the Spatial (left) and Non-Spatial (right) training.

<https://doi.org/10.1371/journal.pone.0263509.g001>

(different azimuths and two different elevation of sound position) and from response demands (pointing with the head instead of reaching with the hand). The rationale for these changes was to minimize potential carry-over of mere sensorimotor adaptations acquired during the spatial training to the novel auditory task. The audio-visual attention cueing task was instead an implicit sound localization task: unlike head-pointing sound localization participants were never asked to explicitly indicate sound position and sounds only served as lateralized attention-orienting cues for the discrimination of visual targets (i.e., an audio-visual analogue to classic attention cueing paradigms) [26–29]. This second test was introduced to probe whether adaptation to new auditory cues could also impact audio-visual attention orienting, a skill that can be hampered by monaural listening [26].

Using a crossover experimental design (Fig 1), a training in which participants performed reaching movement to identify sound position (hereafter referred to as Spatial training) was compared to a control training in which participants performed comparable reaching movements to identify a non-spatial feature of the sound (from now on, the Non-Spatial training). During the Non-Spatial training participants were asked to discriminate between sounds with two different amplitude modulations rather than to focus on the spatial position of the sources. Each participant was tested in both training conditions in two successive sessions, separated by a wash-out period of 2-weeks minimum. While the two trainings used identical physical stimuli, the behavioral requests differed and entailed the processing of spatial vs. non-spatial features of sounds. These different task demands recruit substantially different cognitive and brain mechanisms, as revealed by converging evidence from neurophysiology [30], neuropsychology [31] and cognitive neuroscience [32–34]. Our key prediction was to observe generalization of training effects after the Spatial training more than after the Non-Spatial training paradigm.

Methods

Participants

Twenty participants (age: $M = 29.4$, $SD = 10.5$, 5 males, 19 right-handed) were recruited to participate in the study, carried out in the otolaryngology department of the civil Hospital Edouard Herriot (HEH) in Lyon (France). Previous findings from our laboratory [25] showed that sound localization improvements emerge following a training similar to the one proposed here (see ‘Spatial Training’ below), with an effect size (η^2) of 0.24 (which corresponds to Cohen’s d of 1.12) [35]. Using the G*Power, we calculated that to obtain a similar effect ($\alpha = 0.05$, Power = 0.99) the sample size required was at least 17 participants. Thus, we decided to include 20 participants.

All participants signed an informed consent before starting the experiment, which certified the ethical approval of the national ethics committee in France (Ile de France X, N° ID RCB 2019-A02293-54), and recorded in clinicaltrials.gov (NCT04183348). All participants had normal to corrected-to-normal vision and reported no motor or vestibular deficit as well as no history of neurological or psychiatric disorders. Hearing thresholds were measured for all participants using an audiometer (Equinox 2.0, Interacoustics), testing different frequencies (250, 500, 1000, 2000, 4000, 8000 Hz), on the right and left ear separately. All participants had an average threshold below 20 dB HL for both ears.

General structure of the experimental session

Participants were invited to participate in two experimental sessions, separated by at least 2 weeks. Each session comprised three phases: two testing phases and one training (Fig 1A). The two experimental sessions differed exclusively in the training phase task: firstly, participants were involved in the experimental training task (Spatial Training), and later in the control training task (Non-Spatial Training). In this way, all participants performed both training types in a within-subject design. Order of training type was counterbalanced across participants. The testing phases were identical in both experimental sessions: they included the head-pointing sound localization (conducted in VR) and the audio-visual attention cueing task (conducted outside VR).

Participants completed the entire experiment in monaural listening. This temporary auditory cue alteration was obtained by occluding the right ear of the participant with a plug (3M PP 01 002; attenuation values: high frequencies = 30 dB SPL; medium frequencies = 24 dB SPL; low frequencies = 22 dB SPL) and a monaural ear muffs (3M 1445 modified to cover only the right ear; attenuation values: high frequencies = 32 dB SPL; medium frequencies = 29 dB SPL; low frequencies = 23 dB SPL). At the beginning of their first session, participants also completed the testing phase without the ear plug (Fig 1A). This provided a baseline measure of their spatial hearing skills in binaural listening before exposure to monaural listening.

Apparatus

Virtual reality and kinematic tracking were implemented using the HTC Vive (Vive Enterprise). The system comprised one head-mounted display (HMD, resolution: 1080 x 1200 px, Field Of View (FOV): 110°, Refresh rate: 90 Hz), 2 hand-held controllers (one held by the experimenter to calibrate head center, and one held by participants to interact with the virtual environment during the training phase), 1 tracker mounted above the speaker to track its position in real time, and 2 lighthouse base-stations (Lighthouse V1.0) for scanning the position of the controller, trackers and the HMD. Tracking precision and accuracy of the HTC Vive System are adequate for behavioral research purposes [36]. Stimuli were controlled and delivered

using a LDLC ZALMAN PC (OS: Windows 10 (10.0.0) 64bit; Graphic card: NVIDIA GeForce GTX 1060 6GB; Processor: Intel Core i7-7700K, Quad-Core 4.2 GHz/4.5 GHz Turbo—Cache 8 Mo—TDP 95W) using Steam VR software and the development platform Unity3D (Unity Technologies, San Francisco, CA).

Real free-field auditory stimuli were delivered by a loudspeaker (JBL GO Portable from Harman International Industries, Northridge, California USA, 68.3 x 82.7 x 30.8 mm, Output Power 3.0W, Frequency response 180 Hz– 20 kHz, Signal-to-noise ration > 80 dB), with the HTC Vive tracker firmly attached to its top. During the entire experiment we tracked the loudspeaker position, as well as the position of the controller held in the participant's hand and the HMD, via a dead reckoning process using gyroscope and accelerometer, plus a correction signal from the lighthouse system every 8.333 milliseconds. Both tracking method allowed position tracking with a frequency sample of 250 Hz. The software is designed to guide the experimenter to align the real loudspeaker (i.e., the sound source) with a set of pre-determined positions defined in the 3D virtual environment in each trial. This method combining virtual reality and kinematic tracking to measure sound localization abilities has been developed in our laboratory [37] and has been already adopted in previous studies [25, 38].

The audio-visual attention cueing test was carried out without VR, with the participant sitting at a desk. The apparatus for this task included a separate PC, a DELL 24" monitor, a keyboard and two speakers, positioned at ear level on both sides of the screen, each located 20° to the left or to the right of central fixation (see Fig 1B). The height of the chair on which the participants sat was adjusted to favour the support of the head on the chin rest on the edge of the table, aligned with the centre of the monitor. The test consists of a visual discrimination task implemented with the program OpenSesame®.

Procedure and stimuli

Before starting the experiment, participants were informed about the use of the VR equipment. When engaged in the VR tasks participants sat on a rotating armless chair with no chin rest, placed in the center of the room. The room (3.6 m x 3.9 m, height 2.7 m) was quiet and treated with sound-proof panels to partially reduce echoes. The T60 reverberation of the room was 0.30–0.33 seconds, as measured using a Blue Solo 01bB phonometer. Instruction for each task were provided immediately before the task started.

Testing phase. *Head-pointing sound localization.* This part of the experiment was carried out entirely in VR. The participant was immersed in a virtual room with green walls, reproducing exactly the size and shape of the real room. The room was devoid of any objects, except for light source on the ceiling and a wooden door behind the subject. The task comprised 40 actual trials, plus 5 practice trials presented at the beginning of the task. At the beginning of each trial, participants were asked to direct their gaze in front of them to align their head with a white central fixation cross. As soon as the head correct position was reached, the fixation cross turned blue. This procedure ensured that initial posture was comparable across trials and participants, even in the absence of a chin-rest. In the meanwhile, the experimenter placed the speaker at one of the possible eight pre-determined positions. Eight predetermined positions were used throughout the experiment, resulting from the combination of 4 different azimuths in the frontal space in respect to participant's head position (-67.5° , -22.5° , 22.5° or 67.5°), 2 different elevations ($+5^\circ$ and -15°) and a single distance (55 cm) (Fig 1A). Each position was reached manually by the experimenter using visual indications provided on the dedicated instruction monitor. When the correct starting head posture was reached, and the loudspeaker was positioned correctly (i.e., within a sphere with diameter of 5 cm centred on the pre-determined location), the target sound was delivered.

The target sound consisted of 3 second white noise bursts, amplitude-modulated at 2.5 Hz, and delivered at about 65 dB SPL, as measured from the participant's head. During target sound delivery the fixation cross turned white. From that moment, participants were free to move their heads and rotate the chair they sat on to explore the surrounding space. The task consisted in localizing the exact source of the sound and to indicate it using the head direction as a pointer. This response was aided by the fact that the visible fixation cross in the HMD followed head direction displacements. It is noteworthy that, since the speaker was invisible in VR, participants did not have visual cues about sound source position. At the end of the 3 seconds of sound presentation, the central cross turned red, to indicate to the participants to validate their response (i.e., their current head direction) by pressing the button at the base of their hand-held controller.

Participants were informed that sounds could be delivered anywhere in the 3D around them and they did not have to judge the distance but only the elevation and the azimuth dimension of the sound space. Note that head and trunk movements remained unconstrained both during and after sound emission, allowing spontaneous active listening behaviour (e.g., orienting the head in the direction of the sound even during the sound emission). This task lasted approximately 10 minutes.

Audio-visual attention cueing task. This part of the experiment was carried out entirely outside VR. Participants sat at the experimental table, placed inside the same room in which the VR experiments are carried out. Unlike all the VR tasks, in this audio-visual cueing task participants rested their heads on the chinrest, hence no head-movements were ever allowed.

Each trial started with a white fixation cross appearing in the centre of the monitor and remaining visible until response. After a random delay (450–600 ms), an auditory stimulus (white noise, duration 100 ms) was emitted from one of the two loudspeakers. Loudness of the auditory stimulus was approximately 60 dB SPL, as measured from head position. After 100 ms from sound delivery, the visual target was presented. This consisted of a white filled circle (20 pixels radius, 0.5° of visual angle), appearing on a black background for 140 ms 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). In half of the trials, the visual target and the sound appeared on the same hemispace (congruent trials), whereas on the remaining half the visual target and the sound appeared on opposite hemispaces (incongruent trials).

Participants were asked to keep their gaze toward central fixation throughout the task and to indicate as quickly and accurately as possible the elevation of visual targets appearing one at a time on either side of fixation. Up/down responses were given using the up/down arrows keys on an external keyboard, using the index-middle fingers of the right hand. Participants had a timeout of 2000 ms to give their answer. The experiment started with 8 practice trials, following by 128 trials divided in 2 blocks randomly divided in congruent and incongruent audio-visual conditions and it lasted 10 minutes approximately (Pavani et al., 2017). Participants received feedback on accuracy (percentage of correct responses) and mean response time (in ms) only at the end of each block. Importantly, they were also explicitly informed that sounds were always entirely task-irrelevant.

Training phase. This part of the experiment was carried out entirely in VR. In each session, this phase was either a Spatial or a Non-Spatial training (see below). Both training tasks took place in the same virtual room used for the head-pointing sound localization task. The only difference was that during both training tasks, thirteen virtual speakers were visible in front of participants. They were arranged in a semicircle distributed in front of participants spanning between $\pm 72^\circ$ (12° between each of them). The distance to the participant was always 55 cm (Fig 1C). All sounds were presented just below ear level (-5° offset between

tracker and speaker center). Note that sound positions were different from head-pointing sound localization.

Target sounds were delivered from the same azimuth and elevation as the virtual speakers, with a small distance offset so that the real speaker was 5 cm further away from the virtual speaker, to avoid collisions between the controller and the real speaker during the reaching response (see below). The target stimulus always consisted of a white noise: half of the stimuli were amplitude-modulated at 2 Hz and the remaining half at 3 Hz, to create clearly distinguishable targets. Targets sounds were delivered from each of the virtual speaker in random order (12 repetition for each of the 13 speakers, resulting in 156 trials overall, divided into 3 blocks of 52 trials each).

Crucially, exactly the same stimuli were delivered in the two training tasks, thus making the auditory component of the two trainings identical in all respects. In addition, they both involved a similar motor response: a reaching movement. Participants were informed that at the beginning of each trial the controller had to be held in the right hand, at sternum level, and that once the response was completed it had to return to that starting position. However, for the Spatial training the reaching movements served to indicate the perceived sound position, whereas for the Non-Spatial training they served to indicate the perceived amplitude modulation in the target sound.

Spatial training. Participants had to identify the speaker from which the sound was coming by reaching it using the controller in the right hand (Fig 1C). To prevent participants from colliding with the speaker held in the experimenter's hand, the setting provides that the real speakers was moved back 5 centimeters in depth, which does not affect the directionality of the perceived sound. Participants received a brief vibration of the controller upon contact with the chosen speaker, irrespective of whether the response was correct or incorrect. Importantly, if they reached and touched the correct speaker, the sound stopped. On the contrary, if participants reached the wrong speaker they received a visual feedback: the correct speaker started to flash. Specifically, from the correct source location, a series of red concentric circles (1024x1024 px, 2 circles per second) expanded intermittently, irradiating the surrounding space. The rationale was to capture participants attention even if they were looking toward a different zone of the space, including the opposite hemisphere. The visual feedback and the sound stopped only when the subject reached the correct speaker position with the controller. This has two implications: first, a sense of agency was associated with the correct response; second, whenever the wrong speaker was originally selected, a combination of visual and auditory signals guided the participant to the correct sound source. The entire training lasted about 25 minutes.

Non-Spatial training. Participants had to identify whether the emitted sound amplitude was modulated at faster (3 Hz) or slower (2 Hz) rate (note that participants accustomed with the two auditory stimuli before starting the training). For fast amplitude-modulated sounds, participants directed the controller in front of them, above the row of speakers arranged at head height, aiming to touch an invisible virtual button above the central speaker. For fast amplitude-modulated sounds, participants had to reach an invisible virtual button below the same central speaker (Fig 1C). A vibration from the controller indicated that one of two buttons was correctly reached. As in the Spatial training feedback procedure, a visual feedback was delivered in case of erroneous responses. This was a series of red concentric circles that expanded intermittently from above or below the central speaker to indicate the correct position to reach. Recall that target sounds were nonetheless presented from different spatial positions during the training, although this spatial information was totally task-irrelevant. The entire training lasted about 15 minutes.

Analyses

Statistical analyses were run using R (version 1.0.143) (R Core Team, 2013). For the linear mixed-effect (LME) model, we used the R-packages *emmean*, *lme4*, *lmerTest* in R Studio [39, 40]. The R-package *car* was used to obtain deviance tables from the LME models. When appropriate, we calculated Cohen's d_{av} as index of effect size [41].

Performance. *Head-pointing sound localization.* To study performance in the head-pointing sound localization task, we focused on absolute error (i.e., the absolute deviation of the mean response from the source position) and signed error (i.e., the signed difference between the source and the response) in azimuth and elevation, separately. Signed error was negative or positive, to indicate an overall bias to respond. All variables were calculated for each individual trial. Trials in which the participant had problems with response validation (e.g., they responded before the end of sound emission), or in which the HMD signal were lost, were removed from the analyses (14 trials out of 4000 have been removed).

Audio-Visual Cueing task. To study performance in the Audio-Visual Cueing task we calculated the Cueing Effect (CE), expressed in milliseconds. The CE was calculated by subtracting the average reaction time (RTs) when the auditory cue and the visual target occur on the same side of space, from that obtained when the auditory cue and the visual target occur on opposite sides of space (27). The calculation was performed separately for each participant, phase and training. Trials with incorrect responses (i.e., wrong elevation judgement on the visual target) were excluded from the analyses (369 trials out of 12800).

Training task. To study changes in sound localization performance during the Spatial Training, we examined the absolute error (degrees) and the signed error (degrees) during the reaching to sound task. As target position only changed in the horizontal dimension, all errors were computed only in azimuth. The percentage of correct answers was considered to study the performance during the Non-Spatial training.

Spontaneous head-movements. Head movements were measured in all VR tasks (i.e., head-pointing to sounds, reaching to sounds during the Spatial Training, reaching to indicate amplitude-modulation differences in the Non-Spatial training).

The tangential velocity of the head on the x, y, z axis (expressed in degrees of rotation) using two-points central difference derivate algorithm [42] with 5 points for the half-window (smoothing purpose) was calculated to study head movements. The onset and the end of the movements were computed automatically using a velocity-based threshold ($10^\circ/\text{s}$) [25]. Each head-movement was checked manually by visualizing the spatial rotation changes of the head and its speed using a custom-made tool box in MATLAB R2018b. The rationale was to eliminate trials in case the HMD data were lost.

We focused the analysis on three dependent variables: the number of head-movements, the head-rotation extent around the vertical axis and the head-rotation bias (i.e., center of gravity of head-rotation). To compute the number of head-movements, all the detected peaks of velocity in the head trace were considered, yet all movements smaller than 2° degrees to avoid noise were removed (i.e. excluding movements which are not indicators of spontaneous head intentional movements and may reflect micro postural movements not related to the task). To calculate head-rotation extent we sum the absolute value of the rightward and leftward head-rotation extremity. For instance, if the head rotated 20° to the right and 40° to the left, the head-rotation extent was calculated as the sum of the two, hence 60° . To calculate head-rotation bias, we sum values of the rightward and leftward head-rotation extremity. For instance, if the head rotated 20° to the right and 40° to the left (left is expressed with negative sign: -40°), the head-rotation bias was calculated as the signed sum of the two, therefore -20° .

Data and R codes can be retrieved from osf.io/dt76y. This study was not preregistered.

Results

Does monaural listening increase sound localization errors?

To confirm the immediate effects of monaural plugging before training, we studied absolute and signed errors using separate linear mixed-effects (LME) models with LISTENING (binaural or monaural) and AZIMUTH as fixed effects, and PARTICIPANT (intercept and slope) as random effect. The main effect of LISTENING (absolute error: $\chi^2(1) = 882.70$, $p < 0.001$; signed error: $\chi^2(1) = 785.57$, $p < 0.001$) and the interaction between LISTENING and AZIMUTH (absolute error: $\chi^2(1) = 91.26$, $p < 0.001$; signed error: $\chi^2(1) = 57.39$, $p < 0.001$) reached significance. Compared to binaural listening, errors increased after ear-plugging (mean \pm SD; absolute error: binaural = $4.3^\circ \pm 4.8^\circ$, monaural = $20.9^\circ \pm 11.3^\circ$, Cohen's $d_{av} = 2.06$; signed error: binaural = $-0.2^\circ \pm 2.2^\circ$, monaural = $-17.4^\circ \pm 14.2^\circ$, Cohen's $d_{av} = 2.14$). This was particularly evident for targets delivered more toward the plugged side (see plots of absolute and signed error in Fig 2A and 2B).

Monoaural plugging affected sound localization also in elevation (not shown in Fig 2; but see S1 Fig). We entered absolute and signed errors in elevation in separate LME analyses with LISTENING and ELEVATION as fixed effects, and PARTICIPANT (intercept) as random effects. We

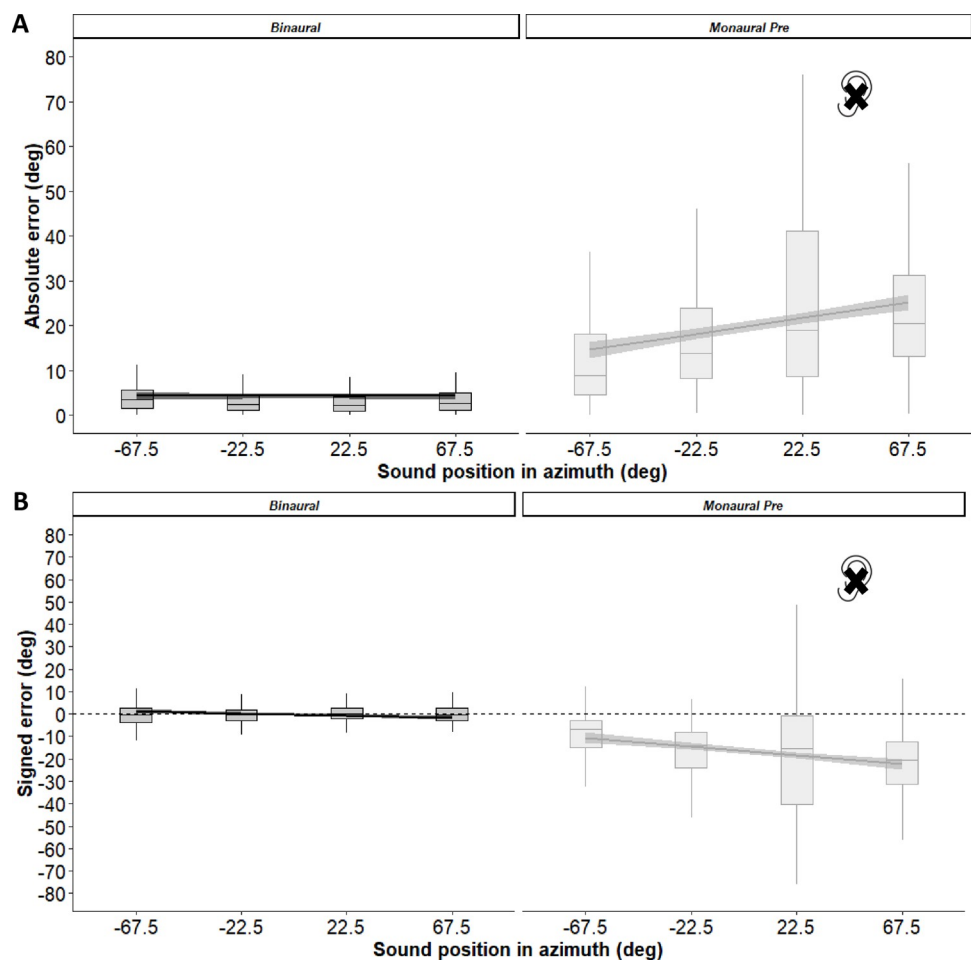


Fig 2. Head-pointing sound localization in session 1 before training. (A) Absolute error and (B) signed error, as a function of sound position in azimuth and listening condition. Dashed lines indicate the 95% confidence intervals of the overall linear regression. Sound positions have been grouped as a function of azimuth, irrespective of elevation.

<https://doi.org/10.1371/journal.pone.0263509.g002>

found a main effect of LISTENING (absolute error: $\chi^2(1) = 13.33$, $p < 0.001$; signed error: $\chi^2(1) = 104.6$, $p < 0.001$). Localization errors in elevation changed after ear-plugging compared to binaural listening (mean \pm SD; absolute error: binaural = 10.6 ± 6.0 , monaural = 9.4 ± 2.7 , Cohen's $d_{av} = 0.28$; signed error: binaural = 9.8 ± 6.6 , monaural = 5.3 ± 4.3 , Cohen's $d_{av} = 0.83$).

Does spatial training reduce sound localization errors?

Next, we turned to examine if the spatial training was effective to reduce sound localization errors caused by monaural listening. Specifically, we studied if participants adapted to monaural listening across successive trials. To this aim, we entered absolute and signed errors into separate LME models with TRIAL NUMBER as fixed effect (25). We also included PARTICIPANT (intercept and slope) and SESSION (intercept) as random effects in the model, with the latter factor added to account for the variability related to the session in which the Spatial training was completed (first or second). The analysis on absolute error revealed a main effect of TRIAL NUMBER, $\chi^2(1) = 4.11$, $p = 0.04$. As shown in Fig 3A (upper panel), the absolute error reduced across trials. The analysis on signed error revealed no main effect of TRIAL NUMBER emerged ($\chi^2(1) = 0.77$, $p = 0.38$), but it is noteworthy that that leftward bias in sound localization decreased numerically, approaching zero (see Fig 3B).

The Non-spatial task was completed with high accuracy by all participants (percent accurate: mean \pm SD = 99.0 ± 1.1) already from trial 1 and throughout the training session (Fig 3C, lower panel). During the Non-Spatial training participants were also faster in completing the trial compared to the Spatial training (Non-Spatial: mean \pm SD = 1.61 ± 0.34 seconds; Spatial training: mean \pm SD = 4.34 ± 1.40 ; $t(19) = 7.96$, $p < .001$ on paired t-test; Cohen's $d_{av} = 3.14$).

Does Spatial training effects generalize to the head-pointing sound localization task?

Having documented that the spatial training improved sound localization, we tested our key hypothesis about generalization of training effects to other sound localization tasks. Fig 4A shows the progression of absolute localization errors in head-pointing sound localization across the two sessions of the experiment, separately for the testing sessions before training (Pre) and after training (Post). Participants who underwent the Spatial training on session 1 and those who underwent the same Spatial training on session 2 (i.e., started with the Non-spatial training instead) are indicated by separate lines (dashed vs. continuous line, respectively). Three aspects are clearly visible in Fig 4A: (1) both trainings improved performance; (2) the Spatial training improved performance to a greater extent compared to the Non-Spatial training; (3) the interval between the two testing sessions (session 1 and session 2) made the two groups again comparable in the pre-training session of session 2, suggesting a partial wash-out of training effects.

To directly compare the effects of training type on head-pointing sound localization, we entered absolute and signed errors in separate LME models with PHASE (Pre or Post Training), TRAINING (Spatial or Non-spatial) and AZIMUTH as fixed effects. As before, we included PARTICIPANT (intercept and slope) and testing SESSION (intercept) as random effects in the model.

The results of these analyses are shown in Fig 4B and 4C. We found a main effect of PHASE, caused by performance improvements after both training types (absolute error: $\chi^2(1) = 76.45$, $p < 0.001$; signed error: $\chi^2(1) = 58.07$, $p < 0.001$), and for absolute error only an interaction between PHASE and AZIMUTH ($\chi^2(1) = 7.78$, $p = 0.005$). Critically, the 2-way interaction between PHASE and TRAINING also reached significance (absolute error: $\chi^2(1) = 19.18$, $p < 0.001$; signed error: $\chi^2(1) = 27.26$, $p < 0.001$; for the complete summary of results of the LME analyses see

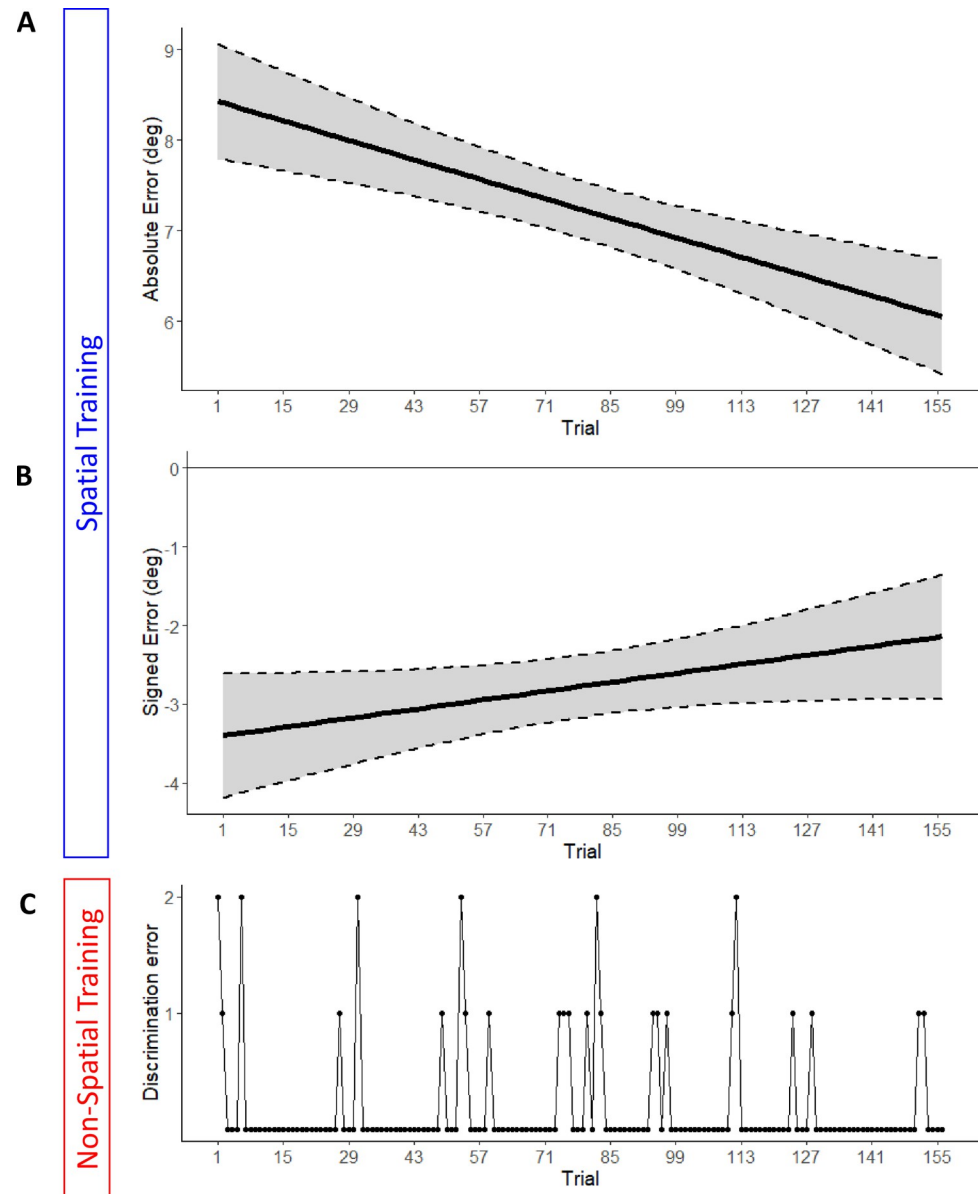


Fig 3. Performance during Spatial and Non-Spatial training. Reduction of absolute (A) and signed (B) error as a function of trial in the Spatial training. (C) Cumulative discrimination error across participants (i.e., number of participants who made a mistake in the trial), shown as a function of trials in the Non-Spatial training. Negative values of signed error indicate a bias toward the unplugged side. Dashed lines indicate the 95% confidence intervals of the overall linear regression.

<https://doi.org/10.1371/journal.pone.0263509.g003>

S1 Table. Before training, localization errors were comparable across training type (absolute error: $t = 0.38$, $p = 0.70$; signed error: $t = 0.23$, $p = 0.82$). After training, errors decreased more substantially after the Spatial compared to the Non-Spatial training (absolute error: $t = 5.81$, $p < 0.001$; signed error: $t = 7.61$, $p < 0.001$). For signed error, also the 3-way interaction reached significance ($\chi^2(1) = 7.18$, $p = 0.007$): the impact of azimuth sound position on localization responses (i.e., larger leftward biases for sounds more toward the plugged side) decreased after the Spatial training ($\chi^2(1) = 23.26$, $p < 0.001$) but not after the Non-spatial training ($\chi^2(1) = 0.24$, $p = 0.62$). Means with standard deviations for each condition are shown in [Table 1](#),

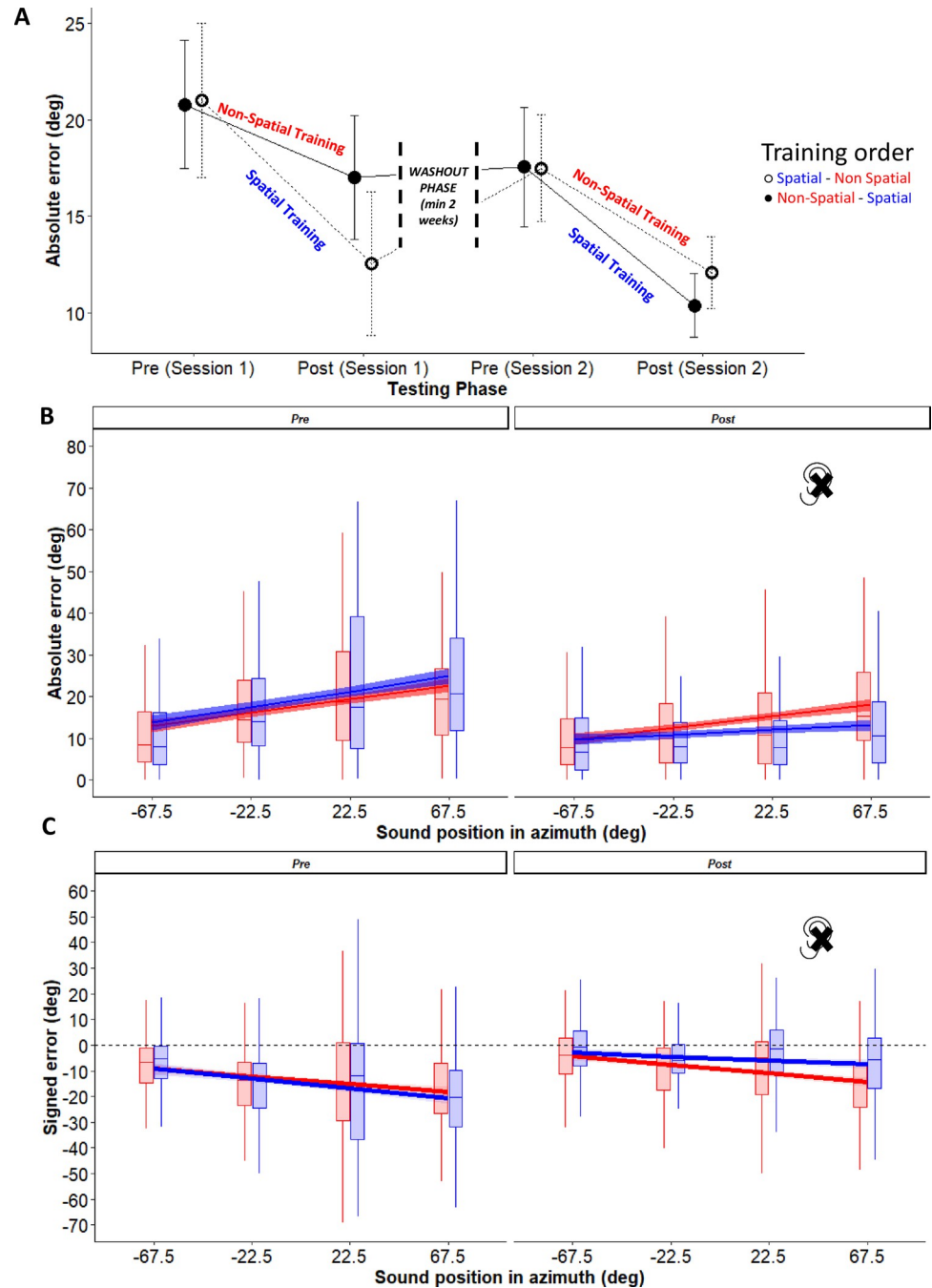


Fig 4. Performance in head-pointing sound localization. Top: (A) Progression of absolute localization across the four testing sessions of the experimental design, separately for participants who completed the Spatial training on session 1 (open circles and dashed line) or session 2 (filled circles and continuous line). Bottom: Absolute (B) and signed (C) errors (in degrees) in monaural listening are plotted as a function of sound position in azimuth, separately for each training type (blue: Spatial training; red: Non-spatial training). For the signed error negative values indicate a bias toward the unplugged side. Dashed lines indicate the 95% confidence intervals of the overall linear regression. Sound positions have been grouped as a function of azimuth and irrespective of elevation.

<https://doi.org/10.1371/journal.pone.0263509.g004>

Table 1. Mean \pm SD and Cohen's d_{av} for absolute and signed errors in azimuth and elevation as a function of PHASE (pre or post training) and training (Spatial or Non-Spatial during monaural listening). Results are pooled irrespective of the order in which the two training tasks were executed.

		Spatial			Non-Spatial		
		Pre	Post	Cohen's d_{av}	Pre	Post	Cohen's d_{av}
Azimuth	Absolute error	19.3° \pm 11.1°	11.5° \pm 8.9°	0.78	19.1° \pm 9.6°	14.5° \pm 8.4°	0.51
	Signed error	-15.1° \pm 14.8°	-5.2° \pm 11.3°	0.75	-15.3° \pm 12.2°	-10.2° \pm 10.2°	0.46
Elevation	Absolute error	9.2° \pm 2.3°	9.5° \pm 3.3°	0.11	8.6° \pm 2.9°	9.2° \pm 3.3°	0.19
	Signed error	5.5° \pm 3.5°	6.5° \pm 5.4°	0.22	3.9° \pm 4.1°	6.0° \pm 4.2°	0.51

<https://doi.org/10.1371/journal.pone.0263509.t001>

together with the corresponding Cohen's d_{av} values (see [S2 Table](#) for each of the 8 sound positions).

To measure the wash-out effect, we entered absolute error during the post testing phase of session 1 and the pre testing phase of session 2 into a LME model. While errors of participants who performed the Non-Spatial training in the first session did not change after the wash-out ($t = 0.72$, $p = 0.47$), the errors of participants who performed the Spatial training in the first session increased after the wash-out ($t = 5.82$, $p < 0.001$, $X^2(1) = 12.96$, $p < 0.001$).

Localization errors in elevation remained unchanged, irrespective of training type. We noted only an upward bias in the post training session compared to the pre-training one (see mean \pm SD in [Table 1](#) and full description of the analyses in [S1](#) and [S2 Tables](#)).

Do improvements during the spatial training predict generalization effects?

We asked if improvements during the spatial training ([Fig 3](#)) predicted the observed generalization effects, as measured in head-pointing sound localization ([Fig 4](#)). To this aim, we correlated performance in the two tasks. As indicator of improvement in the Spatial training, we used individual slope coefficients obtained from the LME model on absolute error (see [Fig 3A](#)). The higher the slope coefficient, the more the participant improved in performance during the Spatial training (for clarity, we expressed improvements as positive numbers by multiplying each slope by -1). As indicator of improvements in head-pointing sound localization, we calculated the error difference (z-normalised) before and after spatial training, separately for absolute and signed error (again, to express improvements in signed error as positive numbers we multiplied the decreasing bias by -1). The higher the error difference, the larger the training generalization effect.

A correlation between the two measures emerged. The more a participant improved during Spatial training, the greater the reduction in absolute ($R = -0.77$, $p < 0.001$) and signed error ($R = 0.57$, $p = 0.009$) during head-pointing sound localization (see [Fig 5](#)).

Does Spatial training change spontaneous head-rotation behavior while listening?

As anticipated in the Introduction, in our previous work [[25](#)] we found that reaching to sounds increased head exploration movements during listening, and these head-movements correlated with improvements in sound localization. In this section, we examined if our training procedures also changed spontaneous head-movements during listening (from sound onset until the first response, but excluding the audio-visual feedback phase). To this aim, we examined three variables: (1) number of head-movements; (2) head-rotation extent; (3) head-rotation bias (positive values indicate rightward head-rotations). Recall that head-rotation extent and bias refer only to movements around the vertical axis (see [Methods](#)).

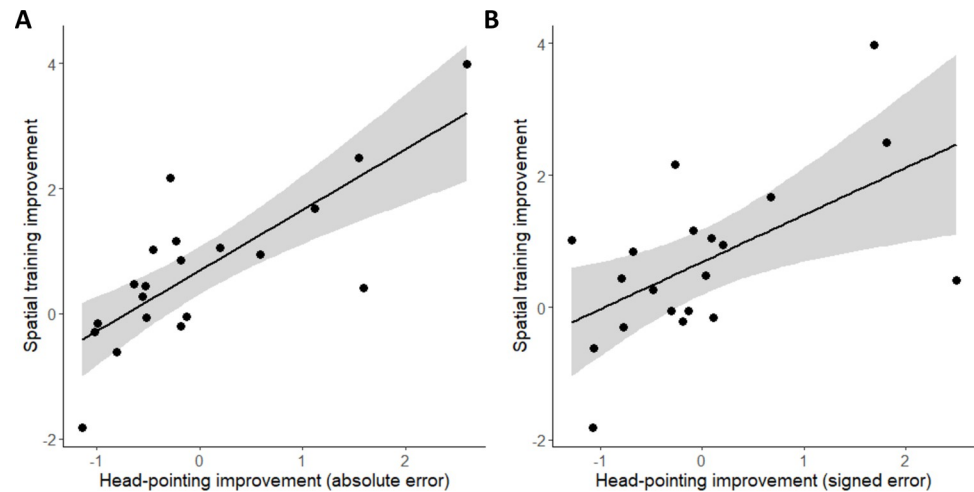


Fig 5. Correlation between the improvement in spatial training and the improvement in head-pointing sound localization, computed on absolute (A) or signed (B) errors.

<https://doi.org/10.1371/journal.pone.0263509.g005>

Overall, participants made more head-movements in the Spatial (2.2 ± 0.8) compared to the Non-Spatial training (1.1 ± 0.4 , $t(19) = 5.94$, $p < .001$ on paired t-test, Cohen's $d_{av} = 1.83$). Furthermore, head-rotation extent was larger during the Spatial ($64.3^\circ \pm 25.9^\circ$) compared to the Non-Spatial training ($4.8^\circ \pm 1.8^\circ$, $t(19) = 10.17$, $p < 0.001$ on paired t-test, Cohen's $d_{av} = 4.29$). Finally, the head-rotation bias revealed that during the Spatial training participants rotated their head more toward the right ($21.3^\circ \pm 25.2^\circ$) as compared to the Non-Spatial training ($2.1^\circ \pm 4.5^\circ$, $t(19) = 3.21$, $p = 0.005$ on paired t-test, Cohen's $d_{av} = 1.29$). These findings indicate that spontaneous head-movements during listening were elicited more in the Spatial compared to the Non-Spatial task.

To study if and how spontaneous head-movements behaviour evolved during the Spatial training (i.e., reaching to sounds) (25), we entered the variables describing head-movements into separate LME models using TRIAL NUMBER and SIDE as fixed effect, and PARTICIPANT (intercept and slope) and SESSION (intercept) as random effects. The main effects of TRIAL NUMBER and SIDE emerged for head-rotation bias (TRIAL NUMBER: $\chi^2(1) = 8.01$, $p = 0.005$; SIDE: $\chi^2(1) = 7671.55$, $p < 0.0001$) and head-rotation extent (TRIAL NUMBER: $\chi^2(1) = 3.82$, $p = 0.05$; SIDE: $\chi^2(1) = 292.83$, $p < 0.0001$), but not for number of head movements (TRIAL NUMBER: $\chi^2(1) = 0.003$, $p = 0.96$; SIDE: $\chi^2(1) = 1.27$, $p = 0.26$). As shown in Fig 6A, participants increased their head rotation bias toward the plugged (right) side (positive) as a function of trial repetition. This head-orienting behaviour is compatible with participants exposing their unplugged ears toward the sound. For head-rotation extent, also the 2-way interaction reached significance (TRIAL NUMBER \times SIDE: $\chi^2(1) = 3.81$, $p < 0.05$, Fig 6B). Finally, changes in head-rotation extent and head-rotation bias as a function of trials were correlated with one another ($R = 0.47$, $p = 0.04$). In sum, during Spatial training phase, participants oriented more their heads toward the plugged (right) side as training progressed. Moreover, they increased the space explored with the head, particularly when sounds were delivered on the plugged side (Fig 6A and 6B).

One final aspect worth noting concerns the relative timing of spontaneous head-movements and required hand-responses (i.e., reaching to sound) during the Spatial training. While the first head movement started on average 0.97 seconds (SD = 0.27) after sound emission, the hand-held controller touched the speaker on average 3.78 seconds (SD = 1.33) after sound emission. This indicates that head-movements implemented during listening preceded hand-

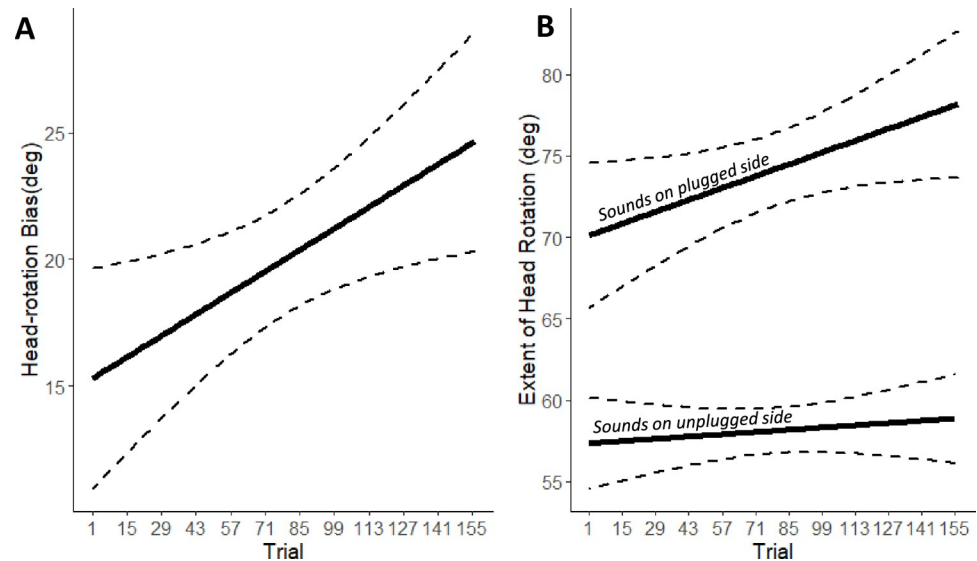


Fig 6. Changes in head behavior during the spatial training. Head-rotation bias as a function of trial (A) and extent of head-rotation as a function of trial and separately for sound delivered to the left or to the right of participant's midline (B). Dashed lines indicate the 95% confidence intervals of the overall linear regression.

<https://doi.org/10.1371/journal.pone.0263509.g006>

reaching movements required to provide the response, hence were not the mere consequence of hand-head coordination towards the target sound.

Does training effects change immediate head-orienting to sounds?

In the head-pointing sound localization task, head-movements were functional to the response, hence they cannot be considered fully spontaneous (unlike in the training tasks). Yet, the first head-movement (which occurred around 0.99 seconds after the beginning of the sound) could be taken to reflect the immediate and spontaneous orienting response toward the sound. Here, we focused on these first movements to study if training effects can be detected also in immediate head-orienting to sounds. Specifically, we examined the horizontal direction of the first head-movement and its onset (i.e., Reaction Time, RT, in seconds).

An LME model on the horizontal direction of the first head-movement (with *PHASE*, *TRAINING* and *SIDE* as fixed effects and *PARTICIPANT* as random effect) revealed that first head movements were overall directed toward the correct hemispace (right: 46.2 ± 26.9 ; left: -42.0 ± 26.6 ; main effect of *SIDE*, $X^2(1) = 1270$, $p < 0.001$). Moreover, there were more rightward oriented responses after training (6.5 ± 25.4) compared to before training (-2.3 ± 26.0 ; *PHASE*, $X^2(1) = 8.82$, $p = 0.003$). This indicates that head-orienting to sounds was more biased towards the right (plugged) side after training, but irrespective of training type.

Most interestingly, a similar LME model on first head-movements RT (analyzed only for the trials in which the head-movement direction was correct, 97.1%) showed a 3-way interaction between *PHASE*, *TRAINING* and *SIDE* ($X^2(1) = 9.33$, $p = 0.003$). This interaction is illustrated in Fig 7. Before training, sounds delivered from the plugged (right) side resulted in slower first head-movements compared to sounds delivered from the left side, for both the Spatial (1.01 ± 0.29 vs. 0.99 ± 0.34 ; $p < 0.001$) and the Non-Spatial (1.07 ± 0.33 vs. 1.02 ± 0.32 ; $p = 0.02$) training. This difference in RT latency for the plugged side reversed selectively after the Spatial training (left: 0.96 ± 0.28 ; right: 0.92 ± 0.28 ; $p = 0.05$), whereas it persisted after the Non-Spatial training (left: 1.02 ± 0.32 ; right: 1.03 ± 0.28 ; $p = 0.009$). This finding provides further

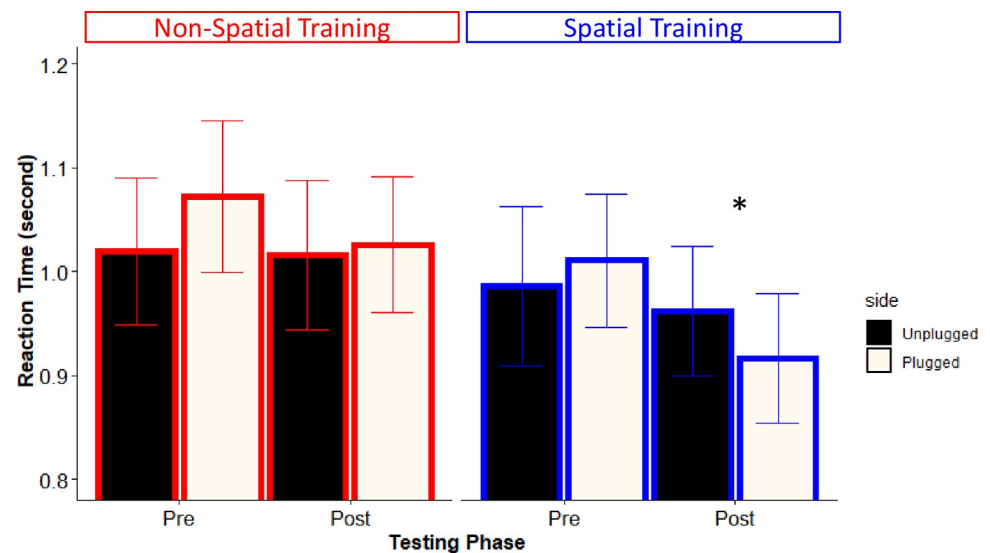


Fig 7. Reaction Time, RT, in seconds as a function of training type (Non-Spatial or Spatial) phase (pre and post training), separately for sound delivered to the left (black) or to the right (white) of participant's midline. Error bars indicate the standard errors. In all comparisons RTs for right targets are significantly slower than those for left targets, except in the post phase of the Spatial training. In the latter case the RT pattern is significantly reversed (marked by an asterisk).

<https://doi.org/10.1371/journal.pone.0263509.g007>

support to the notion that participants developed a bias in orienting their heads toward the right (plugged) hemispace, specifically after the Spatial Training.

Does training induce changes in audio-visual attention orienting?

Finally, we turned to investigate the effects of training in the audio-visual attention cueing task to examine if generalization effects emerged also for this implicit spatial hearing task. To assess the immediate effects of monaural plugging on the audio-visual cueing task, we compared attention Cueing Effects (CE, in milliseconds) in binaural and monaural listening during the first session. The larger the CE the more participants exploited the lateralised sound as a cue for visual attention, i.e., they implicitly took advantage of its spatial position. We entered CE values in a LME model with the LISTENING (binaural or monaural) as fixed effect and PARTICIPANT (intercept) as random effects. As expected from previous works [26], we found a main effect of LISTENING ($X^2(1) = 12.24, p < 0.001$) revealing that the CE decreased after monaural plugging (binaural = 20.2 ± 11.0 ms, monaural = 6.8 ± 13.1 ms, Cohen's $d_{av} = 1.11$).

To study the effects of our training protocols, we entered the CE values in monaural listening conditions into a LME model with PHASE (Pre or Post training), TRAINING (Spatial or Non-Spatial) as fixed effects, and PARTICIPANT and SESSION (intercepts) as random effects. No main effect or interaction emerged from this LME model (all $p_s > 0.49$). As shown in Fig 8, CE remained unchanged between pre and post testing phases, in both the Spatial (pre = 6.9 ± 9.5 ms, post = 9.1 ± 14.6 ms) and Non-Spatial training (pre = 6.0 ± 16.1 ms, post = 3.9 ± 15.9 ms). This indicates that our spatial training affected performance during the training task itself, during head-pointing to sounds (both in terms of localization errors and first head-movements), but did not improve the implicit processing of sound side in relation to the visual stimulus.

Discussion

Generalization is essential when assessing the potentials of any training procedure [10]. Here we tested if a training based on sound-oriented actions can result in adaptations that extend

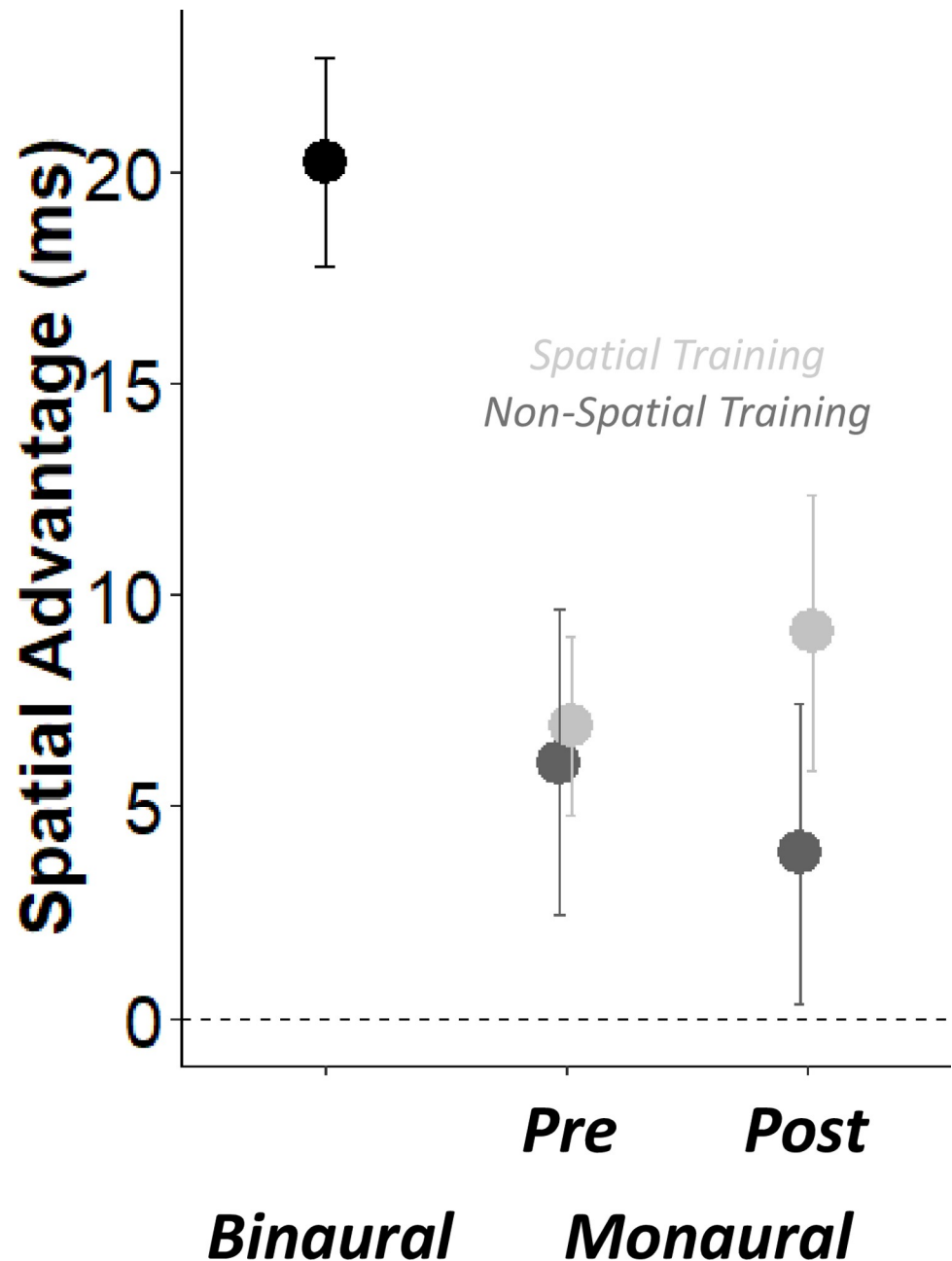


Fig 8. Audio-Visual Cueing task. Audio visual cueing (ms) as a function of PHASE and TRAINING (Spatial in light grey and Non-Spatial in dark grey).

<https://doi.org/10.1371/journal.pone.0263509.g008>

(i.e., generalize) to other untrained sound localization tasks. Using a crossover experimental design, we examined the effects of a Spatial training (indicate sound position through reaching) on two tasks performed before and after the training session: a head-pointing to sound task and an audio-visual attention cueing task. As a control, we used a Non-Spatial training (indicate sound type through reaching). Three main findings emerged. First, we confirmed that listeners using an ear plug can rapidly improve sound localization while performing a Spatial training based on reaching to sounds [25]. Second, we found that the improvements

induced by the Spatial training, generalize to head-pointing sound localization. Instead, the audio-visual attention cueing task remained unaffected. Third, we documented changes in head-movement behaviour during the Spatial training, and we provide initial evidence that such head-movement adjustments can transfer to an untrained auditory spatial task. In the next paragraphs, we discuss each of the main findings in turn.

Reaching to sound is an effective training strategy

In this study, we observed generalization of training effects, particularly after the Spatial training. These results confirm and extend previous findings documenting the benefits of training procedures based on audio-visual feedback [11, 12] and motor interactions with sounds [13]. In these previous works, however, the efficacy of training protocols was always examined using between-subject experimental designs that grouped participants as a function of the proposed training type. By contrast, in the present work, we successfully used a crossover design to compare generalization effects of two training procedures within-subjects. This design gave us the opportunity to minimize any potential intergroup differences. Furthermore, such a crossover design makes this paradigm particularly suited for clinical applications, because it permits to involve the totality of participants in the experimental training without the need of testing a control group (which could not benefit from training).

Our Spatial training approach took advantage of the benefits of audio-visual feedback related to sound position [as in 11 and 12]. In addition, it was characterised by the presence of goal-directed actions toward each sound. We built on works in which interactions with sound sources occurred using reaching movements, as in some pioneers VR approaches [21, 22] and more recent studies [23]. A motor approach based on hand movements in the space of the sound sources has also been adopted by Valzolgher and colleagues [13], in a study that demonstrated the effect of kinesthetic cues when re-learning to localize sounds with one ear plugged, as here. In the above-mentioned previous work, however, the action was not sound-oriented and it was largely repetitive: participants were instructed to move their arm repeatedly over the speakers while wearing a sound-emitting bracelet attached to their wrist. Conversely, in the Spatial training proposed in the current study, participants performed a reaching to sounds action. Our working hypothesis was that reaching to sound could enhance spatial coding of sound position by favouring the coordination of different effectors (eyes, head, hand) into a common reference frame [25, 43, 44]. In addition, reaching to sounds could help directing attention toward the position occupied by the sound source and make the task more engaging, which is a fundamental feature in any learning procedure [45].

While the task requirement in our Spatial training was to reach each sound source with the hand, one critical feature was that participants were free to make spontaneous head-movements while listening to the sounds. The rationale for introducing this spontaneous behaviour in our training was based on two notions. First, head movements imply continuous updates in the acoustic cues reaching the ears, which may reduce ambiguities in the altered auditory cues and, eventually, favour re-learning of sound-space correspondences [46–48]. Second, by moving the head listeners can discover and implement self-regulation strategies which may prove adaptive when coping with the altered auditory cues. In the present study, for instance, a clear strategy emerged: the participant moved the head to expose the unplugged ear to the sounds. Steadman and colleagues [24] also tested three training approaches based on head pointing to sound and audio-visual feedback. Interestingly, their training based on active listening, in which head movements were permitted during sound playback, resulted in greater improvements in localization accuracy, compared to the two other training they proposed (i.e., a non-gamified and gamified version of the training in which the head remained still). In their study,

the time for exploring the auditory scene with head movements was limited to 1.6 seconds, whereas in our experimental training listening occurred without time restriction (with an average sound duration of 4.6 seconds). This difference in sound duration may have led to greater exposure to changes in the auditory cues related to spontaneous head-movements, thus fostering more strategic behaviors to identify sound sources. In future studies, it would be interesting to understand which are the most relevant specific aspect(s) of sound involved motor interactions in our Spatial paradigm.

Although generalisation effects were greater after the Spatial than after the Non-spatial training, performance in the head-pointing task improved significantly after both training procedures. This indicates that participants spontaneously adapted to the altered listening situation, even without feedbacks about the spatial position of sounds (recall that the audio-visual feedback in the Non-Spatial training only entailed the categorical discrimination between the two sound types). Improvements in the absence of relevant training or even no training at all (i.e., simple test-retest) have already been documented in literature [11, 49]. In the present paradigm several methodological aspects could have favoured this spontaneous relearning. First, sounds were emitted from 8 different positions during the head-pointing task and participants may have learned them between the pre- and the post-training sessions, although no visual cues about sound location were ever provided. Second, participants may have become more familiar with the task and with the altered auditory experienced. The performance improvements observed irrespective of training type, however, may be qualitatively different compared to the ones observed after the Spatial training. Indications in this direction may emerge from the differential effects of the wash-out phase as a function of the training performed in the first session (see Fig 4A). The wash-out phase was effective for participants who performed the Spatial training during the first session, whereas it did not change performance for participants who performed the Non-Spatial training. This observation may distinguish between two types of improvement: a more stable one, common to both groups and possibly related to the generic experience with the auditory alteration and the experimental setup, and more a contingent one, proper to the features of the Spatial training.

An aspect worth discussing concerns the lack of training effects in the vertical dimension. In our study the ear plug effect affected sound localization specifically in the horizontal dimension, and altered monaural cues (required to localize in elevation) only at the plugged ear. In addition, the Spatial training protocol provided a rich visual feedback only for the horizontal position of sounds, whereas the only information about sound elevation was related to the continuously visible array of speaker that showed the veridical sound height in both training types. Finally, the difference in elevation (10°) in the head-pointing task was introduced primarily as a way to increase uncertainty about sound position and to differentiate between sound positions in the trained and untrained auditory spatial tasks. Note that the reduced range of elevation (20 degrees) made any potential improvement difficult to see, compared to the larger range of azimuths. It would be interesting to examine if our Spatial training could also prove effective to improve sound position in elevation when hearing is altered using ear-moulds [as in 50] or by degraded spectral content [as in 49].

One word of caution concerning our methodology relates to the different length of the two training protocols. While participants were fast in identifying the amplitude modulation rate of target sounds (mean trial duration in the Non-Spatial training was 1.6 seconds), identification of their spatial source required a longer listening time (mean trial duration in the Spatial training was 4.3 seconds). This implies that two factors could have contributed to the Spatial training efficacy we have documented: the spatial task requirements and the longer stimulus duration. Future studies could overcome this methodological limitation by creating tasks of comparable difficulty, or by imposing a fix sound duration in each trial. Note however that

extending trial duration in the Non-Spatial task could allow participants to process both the sound's amplitude modulation rate and the sound's spatial position, thus making task-relevant and task-irrelevant sound feature more difficult to disentangle.

A possible role for head-movements in adapting to altered spatial hearing

Although studies regarding the contribution of head-movements to spatial hearing have been advocated since the 1940s, only few works have examined how head-movements could change while adapting to altered auditory cues. Our VR approach to spatial hearing made it possible to study head movements throughout the experiment and we always presented sounds with long durations (i.e., until response, during the training phase; or 3 seconds long, during the head-pointing task) precisely to examine the contribution of head-movements during training and generalization. Our work adds to the limited body of work that focused on spontaneous head-movements during sound localization relearning [46–48, 51] by contributing two main findings.

The first finding emerged during the Spatial training phase. While adapting to the altered auditory cues, participants increased head-rotation extent to explore a wider portion of space during listening. Furthermore, they turned their heads more to the right as training trials progressed, as revealed by the increasing head rotation bias. These two head-orienting behaviours can be conceived as spontaneous adaptation strategies for coping with the altered binaural cues, which change spatial hearing specifically in the horizontal dimension. In particular, the rightward bias in head-orienting allowed listeners to direct their unplugged (left) ear toward the sound sources, even when sounds originated from the plugged side. The second finding emerged in the post-training phase. Even in the untrained task (i.e., head-pointing to sounds), we observed again a bias in exploring the right (plugged) side of space. After training, participants directed their first head-movement (within the first second from sound onset) more to the right. Moreover, they triggered this rightward movement faster specifically after the Spatial training. This suggests that the head rightward bias implemented during the Spatial training may have transferred, at least to some extent, to the untrained task.

One important aspect to notice is that head movements during the Spatial training were not only functional to orient the head to the target selected for reaching. Actually, most head-movements preceded reaching, hence they were spontaneously implemented by participants. Although 61.5% of trials fell within 10 degrees from the speaker reached with the hand, in the remaining proportion of trials the head was not aligned with the hand. In fact, the misalignment between the hand and the head was 9.7 degrees (SD = 8.7) when the hand reached the target speaker. Moreover, during the Spatial training, participants made 2.2 movements (on average) in each trial, and the timing of head-movements and reaching actions differed. The first head movement during the Spatial training started around 1 second after the beginning of the sound, whereas the first hand-contact with the speaker in each trial occurred around 4 seconds after the beginning of the sound. Most notably, the rightward bias of the head we described above indicates that head movements were not only performed to coordinate head-pointing with the hand-reaching response. Taken together, these findings indicate that participants explored the auditory scene with their head before taking the decision to act on the sound source. One important implication of this finding is that head-movements during the Spatial training likely served two purposes: exploration of the acoustic scene and support for the reaching response phase. The latter coupling between head-movements and reaching response reveals a partial overlap between the motor response required for the untrained task (head-pointing to sound task) and the Spatial training task (head movements that accompanied the hand reaching response). Future work should address generalization also using

auditory tasks that do not entail any such overlap in the motor response between trained and untrained task (e.g., assessing generalization through minimum audible-angle tasks, as in [13]).

To what extent spontaneous head-movements during sound contributed to the exploitation of auditory cues remain to be ascertained. As suggested above, one possibility is that head movements dynamically changed the auditory cues to location, thereby reducing ambiguity in the incoming signals. This better sampling of the auditory cues under the altered listening condition could have then favoured the creation of new correspondences between cues and space. A second possibility, however, is that head-orienting did not serve the purpose to reduce ambiguity, but allowed to better exploit monaural signals at the unplugged ear. Specifically, participants may have learned to use sound intensity at the unplugged ear as a proxy of sound position, searching head-position at which sound intensity reached a peak and then encoding the position in space in memory. Our findings are suggestive of this second head-orienting strategy, because participants clearly developed a bias in head-orienting toward the plugged side (which implies exposing more the unplugged ear to the frontal sounds). Yet, we cannot exclude that by implementing this strategy they also acquired richer information about the auditory cues and reduced their ambiguity.

A better understanding of the relations between head-movements and the exploitation of auditory cues would be important to characterize in which way the spontaneous head-orienting behaviors interact with the two cognitive mechanisms that have been typically associated with sound localization relearning, namely ‘cue-reweighting’ and ‘cue-remapping’. Cue-reweighting, indicates the mechanism by which the brain changes the relative weight attributed to each auditory cue (ILD, ITD and monaural), giving greater prominence to the reliable auditory cues at the expenses of the more impoverished ones [52]. For instance, when one ear is plugged the brain can learn to weight monaural cues more than the altered binaural cues, when localizing sounds in azimuth. Instead, cue-remapping [53] refers to the ability to create new correspondences between auditory cues and spatial positions through experience. Although all these mechanisms could have played a role in the observed sound localization improvements, we cannot examine exactly the extent of their relative contribution primarily because of our use of naturalistic stimuli in the free-field and the adoption of head-movements during listening in most tasks. Future studies should address in which way allowing spontaneous head-movements during training could impact on cue-reweighting and cue-remapping mechanisms, for instance by pairing our active Spatial training task with more controlled auditory inputs to the ears and tasks performed with the head static.

No generalization of training to the audio-visual attention cueing task

In the present work we also attempted to examine if training acoustic space perception could have effects beyond explicit sound localization alone. To this aim, we tested generalization effects in an audio-visual attention cueing task, in which sound localisation skills are functional to capture attention to the portion of space in which a visual target appears. However, our findings did not reveal any generalization to this task.

While negative findings must be interpreted with great caution, one methodological difference between the audio-visual attention cueing task and the head-pointing task is worth noting. Head movements were entirely prevented by a chin-rest in the attention cueing task, whereas they were fully allowed in the head-pointing task. This aspect was not the only difference between the two tasks: they were also conducted outside vs. inside VR, they entailed indirect vs. direct used of auditory spatial information, respectively. Yet, when considered together with the observations reported above on the prominence of head-movements during training

and the possible role of this behaviour in the generalization effects, restrained head-movements may have been a critical factor limiting generalization of training to our attention task.

Conclusions

Our findings show that reaching to sounds improves spatial hearing in participants using an ear-plug, more than a control condition using the same stimuli but non-spatial task demands. Training by reaching also modified head-movement behaviour. Crucially, the improvements observed during training generalized to a different sound localization task, possibly as a consequence of acquired and novel head-movement strategies. These findings extend the current perspectives about the mechanisms by which humans can adapt to altered auditory cues by showing the important role of active interactions with sound sources (reaching and head-movements) during relearning. Note that these interactions are prevented when participants use a chin-rest, and may be limited when sounds are delivered for a very short time and beyond reaching. In this respect, our results point to a methodological shift towards more active listening scenarios, which could better approximate the naturalistic experience with sounds. They also have implications for clinical and applied settings because they demonstrate that implementing multisensory-motor approaches to train acoustic space perception can be done effectively. Future studies could extend this approach to hearing-impaired population with spatial hearing difficulties. For example, they could test the effectiveness of reaching-to-sound training protocols by measuring the performance of deaf patients with unilateral hearing loss or people who wear cochlear implants or hearing aids [54].

Supporting information

S1 Fig. Head-pointing sound localization in elevation. Density plot of responses during Session 1, as a function of sound position in elevation. The center of the speaker was located at -15 or +5 with respect to the ear level (shown by dashed lines), depending on the trial. This asymmetrical arrangement of the speakers resulted from the fact that the software for placing the speakers at pre-determined positions was based on the position of the VIVE tracker above the speaker, rather than the position of the actual speaker.

(PDF)

S1 Table. Results of the LME analyses. LME analyses of absolute and signed errors in azimuth and elevation.

(PDF)

S2 Table. Absolute and signed errors in azimuth. Mean \pm SD for absolute and signed errors in azimuth as a function of 8 target positions and PHASE (pre or post training) and training (Spatial or Non-Spatial during monaural listening).

(PDF)

Acknowledgments

We are grateful to all the participants in the study and to Dr. Stefania Benetti for comments on a previous version of the manuscript. We thank Isabelle Viaud-Delmon and one anonymous reviewer for their comments and suggestions on a previous version of the manuscript. We thank Jean-Louis Borach and Sonia Alouche for administrative support of this project and Eric Koun for technical support. We thank Giordana Torresani for her contributions to [Fig 1](#) and Angela Valenti for improving English in a previous version of the manuscript.

Author Contributions

Conceptualization: Chiara Valzolgher, Alessandro Farnè, Francesco Pavani.

Data curation: Chiara Valzolgher, Julie Gatel.

Formal analysis: Chiara Valzolgher.

Funding acquisition: Chiara Valzolgher, Eric Truy, Alessandro Farnè, Francesco Pavani.

Investigation: Chiara Valzolgher, Michela Todeschini, Valerie Gaveau.

Methodology: Chiara Valzolgher, Alessandro Farnè, Francesco Pavani.

Project administration: Chiara Valzolgher, Julie Gatel, Eric Truy, Alessandro Farnè, Francesco Pavani.

Resources: Gregoire Verdet, Julie Gatel, Romeo Salemme, Eric Truy.

Software: Gregoire Verdet, Romeo Salemme.

Supervision: Alessandro Farnè, Francesco Pavani.

Visualization: Chiara Valzolgher, Francesco Pavani.

Writing – original draft: Chiara Valzolgher, Francesco Pavani.

Writing – review & editing: Chiara Valzolgher, Valerie Gaveau, Alessandro Farnè, Francesco Pavani.

References

1. Middlebrooks JC. Sound localization. *Handb Clin Neurol*. 2015; 129:99–116. <https://doi.org/10.1016/B978-0-444-62630-1.00006-8> PMID: 25726265.
2. Middlebrooks JC, Green DM. Sound localization by human listeners. *Annu Rev Psychol*. 1991; 42:135–59. <https://doi.org/10.1146/annurev.ps.42.020191.001031> PMID: 2018391
3. Risoud M, Hanson JN, Gauvrit F, Renard C, Lemesre PE, Bonne NX, et al. Sound source localization. *Eur Ann Otorhinolaryngol Head Neck Dis*. 2018 Aug; 135(4):259–264. <https://doi.org/10.1016/j.anorl.2018.04.009> Epub 2018 May 3. PMID: 29731298.
4. Kumpik DP, King AJ. A review of the effects of unilateral hearing loss on spatial hearing. *Hear Res*. 2019 Feb; 372:17–28. <https://doi.org/10.1016/j.heares.2018.08.003> Epub 2018 Aug 11. PMID: 30143248; PMCID: PMC6341410.
5. Dobrev MS O'Neill WE, Paige GD. Influence of aging on human sound localization. *J Neurophysiol*. 2011 May; 105(5):2471–86. <https://doi.org/10.1152/jn.00951.2010> Epub 2011 Mar 2. PMID: 21368004; PMCID: PMC3094163.
6. Van den Bogaert T, Klasen TJ, Moonen M, Van Deun L, Wouters J. Horizontal localization with bilateral hearing aids: without is better than with. *J Acoust Soc Am*. 2006 Jan; 119(1):515–26. <https://doi.org/10.1121/1.2139653> PMID: 16454305.
7. Akeroyd MA. An overview of the major phenomena of the localization of sound sources by normal-hearing, hearing-impaired, and aided listeners. *Trends Hear*. 2014 Dec 9;18:2331216514560442. <https://doi.org/10.1177/2331216514560442> PMID: 25492094; PMCID: PMC4271773.
8. Carlile S. The plastic ear and perceptual relearning in auditory spatial perception. *Front Neurosci*. 2014 Aug 6;8:237. <https://doi.org/10.3389/fnins.2014.00237> PMID: 25147497; PMCID: PMC4123622.
9. Keating P, King AJ. Sound localization in a changing world. *Curr Opin Neurobiol*. 2015 Dec; 35:35–43. <https://doi.org/10.1016/j.conb.2015.06.005> Epub 2015 Jun 28. PMID: 26126152.
10. Mendonça C. A review on auditory space adaptations to altered head-related cues. *Front Neurosci*. 2014 Jul 25; 8:219. <https://doi.org/10.3389/fnins.2014.00219> PMID: 25120422; PMCID: PMC4110508.
11. Rabini G, Altobelli E, Pavani F. Interactions between egocentric and allocentric spatial coding of sounds revealed by a multisensory learning paradigm. *Sci Rep*. 2019 May 27; 9(1):7892. <https://doi.org/10.1038/s41598-019-44267-3> PMID: 31133688; PMCID: PMC6536515.

12. Strelnikov K, Rosito M, Barone P. Effect of audiovisual training on monaural spatial hearing in horizontal plane. *PLoS One*. 2011 Mar 29; 6(3):e18344. <https://doi.org/10.1371/journal.pone.0018344> PMID: [21479241](https://pubmed.ncbi.nlm.nih.gov/21479241/); PMCID: PMC3066228.
13. Valzolgher C, Campus C, Rabini G, Gori M, Pavani F. Updating spatial hearing abilities through multi-sensory and motor cues. *Cognition*. 2020 Nov; 204:104409. <https://doi.org/10.1016/j.cognition.2020.104409> Epub 2020 Jul 24. PMID: [32717425](https://pubmed.ncbi.nlm.nih.gov/32717425/).
14. Kacelnik O, Nodal FR, Parsons CH, King AJ. Training-induced plasticity of auditory localization in adult mammals. *PLoS Biol*. 2006 Apr; 4(4):e71. <https://doi.org/10.1371/journal.pbio.0040071> Epub 2006 Mar 7. PMID: [16509769](https://pubmed.ncbi.nlm.nih.gov/16509769/); PMCID: PMC1393755.
15. Irving S, Moore DR. Training sound localization in normal hearing listeners with and without a unilateral ear plug. *Hear Res*. 2011 Oct; 280(1–2):100–8. <https://doi.org/10.1016/j.heares.2011.04.020> Epub 2011 May 27. PMID: [21640176](https://pubmed.ncbi.nlm.nih.gov/21640176/).
16. Knudsen EI, Knudsen PF. Vision guides the adjustment of auditory localization in young barn owls. *Science*. 1985 Nov 1; 230(4725):545–8. <https://doi.org/10.1126/science.4048948> PMID: [4048948](https://pubmed.ncbi.nlm.nih.gov/4048948/)
17. 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*. 2013 Sep; 134(3):2148–59. <https://doi.org/10.1121/1.4816543> PMID: [23967945](https://pubmed.ncbi.nlm.nih.gov/23967945/).
18. Zahorik P, Bangayan P, Sundareswaran V, Wang K, Tam C. Perceptual recalibration in human sound localization: learning to remediate front-back reversals. *J Acoust Soc Am*. 2006 Jul; 120(1):343–59. <https://doi.org/10.1121/1.2208429> PMID: [16875231](https://pubmed.ncbi.nlm.nih.gov/16875231/).
19. Parseihian G, Katz BF. Rapid head-related transfer function adaptation using a virtual auditory environment. *J Acoust Soc Am*. 2012 Apr; 131(4):2948–57. <https://doi.org/10.1121/1.3687448> PMID: [22501072](https://pubmed.ncbi.nlm.nih.gov/22501072/).
20. Mendonça C, Campos G, Dias P, Santos JA. Learning auditory space: generalization and long-term effects. *PLoS One*. 2013 Oct 22; 8(10):e77900. <https://doi.org/10.1371/journal.pone.0077900> PMID: [24167588](https://pubmed.ncbi.nlm.nih.gov/24167588/); PMCID: PMC3805533.
21. Honda A, Shibata H, Gyoba J, Saitou K, Iwaya Y, Suzuki Y. Transfer effects on sound localization performances from playing a virtual three-dimensional auditory game. *Appl Acoust*. 2007; 68(8):885–96.
22. Ohuchi M, Iwaya Y, Suzuki Y, Munekata T. Training effect of a virtual auditory game on sound localization ability of the visually impaired. *Georgia Institute of Technology*. 2005.
23. Poirier-Quinot D, Katz B. Assessing the Impact of Head-Related Transfer Function Individualization on Task Performance: Case of a Virtual Reality Shooter Game. *J Audio Eng Soc*. 2020; 68(4):248–60.
24. Steadman MA, Kim C, Lestang JH, Goodman DFM, Picinali L. Short-term effects of sound localization training in virtual reality. *Sci Rep*. 2019 Dec 4; 9(1):18284. <https://doi.org/10.1038/s41598-019-54811-w> PMID: [31798004](https://pubmed.ncbi.nlm.nih.gov/31798004/); PMCID: PMC6893038.
25. Valzolgher C, Verdelet G, Salemm R, Lombardi L, Gaveau V, Farné A, et al. Reaching to sounds in virtual reality: A multisensory-motor approach to promote adaptation to altered auditory cues. *Neuropsychologia*. 2020 Dec; 149:107665. <https://doi.org/10.1016/j.neuropsychologia.2020.107665> Epub 2020 Oct 29. PMID: [33130161](https://pubmed.ncbi.nlm.nih.gov/33130161/).
26. Pavani F, Venturini M, Baruffaldi F, Artesini L, Bonfioli F, Frau GN, et al. Spatial and non-spatial multi-sensory cueing in unilateral cochlear implant users. *Hear Res*. 2017 Feb; 344:24–37. <https://doi.org/10.1016/j.heares.2016.10.025> Epub 2016 Oct 31. PMID: [27810286](https://pubmed.ncbi.nlm.nih.gov/27810286/).
27. Spence C, Driver J. Audiovisual links in endogenous covert spatial attention. *J Exp Psychol Hum Percept Perform*. 1996 Aug; 22(4):1005–30. <https://doi.org/10.1037//0096-1523.22.4.1005> PMID: [8756965](https://pubmed.ncbi.nlm.nih.gov/8756965/)
28. Koelewijn T, Bronkhorst A, Theeuwes J. Competition between auditory and visual spatial cues during visual task performance. *Exp Brain Res*. 2009 Jun; 195(4):593–602. <https://doi.org/10.1007/s00221-009-1829-y> Epub 2009 May 13. PMID: [19436999](https://pubmed.ncbi.nlm.nih.gov/19436999/).
29. Posner MI. Orienting of attention. *Q J Exp Psychol*. 1980 Feb; 32(1):3–25. <https://doi.org/10.1080/00335558008248231> PMID: [7367577](https://pubmed.ncbi.nlm.nih.gov/7367577/)
30. Rauschecker JP. Where, When, and How: Are they all sensorimotor? Towards a unified view of the dorsal pathway in vision and audition. *Cortex*. 2018 Jan; 98:262–268. <https://doi.org/10.1016/j.cortex.2017.10.020> Epub 2017 Nov 3. PMID: [29183630](https://pubmed.ncbi.nlm.nih.gov/29183630/); PMCID: PMC5771843.
31. Clarke S, Bellmann A, Meuli RA, Assal G, Steck AJ. Auditory agnosia and auditory spatial deficits following left hemispheric lesions: evidence for distinct processing pathways. *Neuropsychologia*. 2000; 38(6):797–807. [https://doi.org/10.1016/s0028-3932\(99\)00141-4](https://doi.org/10.1016/s0028-3932(99)00141-4) PMID: [10689055](https://pubmed.ncbi.nlm.nih.gov/10689055/).
32. Noyce AL, Cestero N, Michalka SW, Shinn-Cunningham BG, Somers DC. Sensory-Biased and Multiple-Demand Processing in Human Lateral Frontal Cortex. *J Neurosci*. 2017 Sep 6; 37(36):8755–8766. <https://doi.org/10.1523/JNEUROSCI.0660-17.2017> Epub 2017 Aug 8. PMID: [28821668](https://pubmed.ncbi.nlm.nih.gov/28821668/); PMCID: PMC5588466.

33. Alain C, Arnott SR, Hevenor S, Graham S, Grady CL. "What" and "where" in the human auditory system. *Proc Natl Acad Sci U S A*. 2001 Oct 9; 98(21):12301–6. <https://doi.org/10.1073/pnas.211209098> Epub 2001 Sep 25. PMID: [11572938](https://pubmed.ncbi.nlm.nih.gov/11572938/); PMCID: PMC59809.
34. Ahveninen J, Jääskeläinen IP, Raij T, Bonmassar G, Devore S, Hämäläinen M, et al. Task-modulated "what" and "where" pathways in human auditory cortex. *Proc Natl Acad Sci U S A*. 2006 Sep 26; 103(39):14608–13. <https://doi.org/10.1073/pnas.0510480103> Epub 2006 Sep 18. PMID: [16983092](https://pubmed.ncbi.nlm.nih.gov/16983092/); PMCID: PMC1600007.
35. Brysbaert M, Stevens M. Power Analysis and Effect Size in Mixed Effects Models: A Tutorial. *J Cogn*. 2018 Jan 12; 1(1):9. <https://doi.org/10.5334/joc.10> PMID: [31517183](https://pubmed.ncbi.nlm.nih.gov/31517183/); PMCID: PMC6646942.
36. Verdelet G, Saleme R, Desoche C, Volland F, Farne A, Coudert A, et al. Assessing spatial and temporal reliability of the vive System as a tool for naturalistic behavioural research. In: 2019 International Conference on 3D Immersion (IC3D). IEEE, 2019. p. 1–8.
37. Gaveau V., Coudert A., Saleme R., Koun E., Desoche C., Truy E., et al. (2020). SPHERE: A novel approach to 3D and active sound localization. *bioRxiv*.
38. Valzolgher C, Alzhaler M, Gessa E, Todeschini M, Nieto P, Verdelet G, et al. The impact of a visual spatial frame on real sound-source localization in virtual reality. *Curr Res Behav Sci*. 2020; 1:100003.
39. Bates D, Mächler M, Bolker BM, Walker SC. Fitting linear mixed-effects models using lme4. *J Stat Softw*. 2015; 67(1).
40. Fox J, Weisberg S. Using car and effects Functions in Other Functions. 2021.
41. Lakens D. Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. *Front Psychol*. 2013 Nov 26; 4:863. <https://doi.org/10.3389/fpsyg.2013.00863> PMID: [24324449](https://pubmed.ncbi.nlm.nih.gov/24324449/); PMCID: PMC3840331.
42. Bahill AT, McDonald JD. Frequency limitations and optimal step size for the two-point central difference derivative algorithm with applications to human eye movement data. *IEEE Trans Biomed Eng*. 1983 Mar; 30(3):191–4. <https://doi.org/10.1109/tbme.1983.325108> PMID: [6862495](https://pubmed.ncbi.nlm.nih.gov/6862495/).
43. Cohen YE, Andersen RA. A common reference frame for movement plans in the posterior parietal cortex. *Nat Rev Neurosci*. 2002 Jul; 3(7):553–62. <https://doi.org/10.1038/nrn873> PMID: [12094211](https://pubmed.ncbi.nlm.nih.gov/12094211/).
44. Boyer EO, Babayan BM, Bevilacqua F, Noisternig M, Warusfel O, Roby-Brami A, et al. From ear to hand: the role of the auditory-motor loop in pointing to an auditory source. *Front Comput Neurosci*. 2013 Apr 22; 7:26. <https://doi.org/10.3389/fncom.2013.00026> PMID: [23626532](https://pubmed.ncbi.nlm.nih.gov/23626532/); PMCID: PMC3631711.
45. Dehaene S. *How We Learn: Why Brains Learn Better Than Any Machine*. for Now. Viking; 2020.
46. Brimijoin WO, Akeroyd MA. The role of head movements and signal spectrum in an auditory front/back illusion. *Iperception*. 2012; 3(3):179–82. <https://doi.org/10.1068/i7173sas> Epub 2012 Mar 22. PMID: [23145279](https://pubmed.ncbi.nlm.nih.gov/23145279/); PMCID: PMC3485843.
47. Pastore MT, Natale SJ, Yost WA, Dorman MF. Head Movements Allow Listeners Bilaterally Implanted With Cochlear Implants to Resolve Front-Back Confusions. *Ear Hear*. 2018 Nov/Dec; 39(6):1224–1231. <https://doi.org/10.1097/AUD.0000000000000581> PMID: [29664750](https://pubmed.ncbi.nlm.nih.gov/29664750/); PMCID: PMC6191386.
48. Wallach H. The role of head movements and vestibular and visual cues in sound localization. *J Exp Psychol*. 1940 Oct; 27(4):339–68.
49. Zonooz B, Arani E, Opstal AJV. Learning to localise weakly-informative sound spectra with and without feedback. *Sci Rep*. 2018 Dec 18; 8(1):17933. <https://doi.org/10.1038/s41598-018-36422-z> PMID: [30560940](https://pubmed.ncbi.nlm.nih.gov/30560940/); PMCID: PMC6298951.
50. Hofman PM, Van Riswick JG, Van Opstal AJ. Relearning sound localization with new ears. *Nat Neurosci*. 1998 Sep; 1(5):417–21. <https://doi.org/10.1038/1633> PMID: [10196533](https://pubmed.ncbi.nlm.nih.gov/10196533/).
51. Coudert A, Gaveau V, Gatel J, Verdelet G, Saleme R, Farne A, et al. Spatial Hearing Difficulties in Reaching Space in Bilateral Cochlear Implant Children Improve With Head Movements. *Ear Hear*. 2021 Jul 1; 43(1):192–205. <https://doi.org/10.1097/AUD.0000000000001090> PMID: [34225320](https://pubmed.ncbi.nlm.nih.gov/34225320/); PMCID: PMC8694251.
52. Keating P, King AJ. Developmental plasticity of spatial hearing following asymmetric hearing loss: context-dependent cue integration and its clinical implications. *Front Syst Neurosci*. 2013 Dec 27; 7:123. <https://doi.org/10.3389/fnsys.2013.00123> PMID: [24409125](https://pubmed.ncbi.nlm.nih.gov/24409125/); PMCID: PMC3873525.
53. Keating P, Dahmen JC, King AJ. Complementary adaptive processes contribute to the developmental plasticity of spatial hearing. *Nat Neurosci*. 2015 Feb; 18(2):185–7. <https://doi.org/10.1038/nn.3914> Epub 2015 Jan 12. PMID: [25581359](https://pubmed.ncbi.nlm.nih.gov/25581359/); PMCID: PMC4338598.
54. Litovsky RY, Parkinson A, Arcaroli J. Spatial hearing and speech intelligibility in bilateral cochlear implant users. *Ear Hear*. 2009 Aug; 30(4):419–31. <https://doi.org/10.1097/AUD.0b013e3181a165be> PMID: [19455039](https://pubmed.ncbi.nlm.nih.gov/19455039/); PMCID: PMC2873678.

Spatial Hearing Difficulties in Reaching Space in Bilateral Cochlear Implant Children Improve With Head Movements

Aurélie Coudert,^{1,2,3,4} Valérie Gaveau,^{1,4} Julie Gatel,³ Grégoire Verdelet,^{1,5}
Romeo Salemme,^{1,5} Alessandro Farne,^{1,4,5,6} Francesco Pavani,^{1,6,7} and Eric Truy^{1,2,3,4}

Objectives: The aim of this study was to assess three-dimensional (3D) spatial hearing abilities in reaching space of children and adolescents fitted with bilateral cochlear implants (BCI). The study also investigated the impact of spontaneous head movements on sound localization abilities.

Design: BCI children (N = 18, aged between 8 and 17) and age-matched normal-hearing (NH) controls (N = 18) took part in the study. Tests were performed using immersive virtual reality equipment that allowed control over visual information and initial eye position, as well as real-time 3D motion tracking of head and hand position with subcentimeter accuracy. The experiment exploited these technical features to achieve trial-by-trial exact positioning in head-centered coordinates of a single loudspeaker used for real, near-field sound delivery, which was reproducible across trials and participants. Using this novel approach, broadband sounds were delivered at different azimuths within the participants' arm length, in front and back space, at two different distances from their heads. Continuous head-monitoring allowed us to compare two listening conditions: "head immobile" (no head movements allowed) and "head moving" (spontaneous head movements allowed). Sound localization performance was assessed by computing the mean 3D error (i.e. the difference in space between the X-Y-Z position of the loudspeaker and the participant's final hand position used to indicate the localization of the sound's source), as well as the percentage of front-back and left-right confusions in azimuth, and the discriminability between two nearby distances. Several clinical factors (i.e. age at test, interimplant interval, and duration of binaural experience) were also correlated with the mean 3D error. Finally, the Speech Spatial and Qualities of Hearing Scale was administered to BCI participants and their parents.

Results: Although BCI participants distinguished well between left and right sound sources, near-field spatial hearing remained challenging, particularly under the "head immobile" condition. Without visual priors of the sound position, response accuracy was lower than that of their NH peers, as evidenced by the mean 3D error (BCI: 55 cm, NH: 24 cm, $p = 0.008$). The BCI group mainly pointed along the interaural

axis, corresponding to the position of their CI microphones. This led to important front-back confusions (44.6%). Distance discrimination also remained challenging for BCI users, mostly due to sound compression applied by their processor. Notably, BCI users benefitted from head movements under the "head moving" condition, with a significant decrease of the 3D error when pointing to front targets ($p < 0.001$). Interimplant interval was correlated with 3D error ($p < 0.001$), whereas no correlation with self-assessment of spatial hearing difficulties emerged ($p = 0.9$).

Conclusions: In reaching space, BCI children and adolescents are able to extract enough auditory cues to discriminate sound side. However, without any visual cues or spontaneous head movements during sound emission, their localization abilities are substantially impaired for front-back and distance discrimination. Exploring the environment with head movements was a valuable strategy for improving sound localization within individuals with different clinical backgrounds. These novel findings could prompt new perspectives to better understand sound localization maturation in BCI children, and more broadly in patients with hearing loss.

Key words: Children, Cochlear implant, Head movements, Sound localization, Spatial hearing, Virtual reality.

Abbreviations: 3D = three-dimensional; AGC = automatic gain control; BCI = bilateral cochlear implant; CI = cochlear implant; DRR = direct-to-reverberant energy ratio; HMD = head-mounted display; ILD = interaural-level differences; ITD = interaural time differences; NH = normal hearing; RMS = root-mean-square; SSQ = Speech, Spatial, and Qualities of Hearing Scale.

(Ear & Hearing 2022;43:192–205)

INTRODUCTION

Spatial hearing is fundamental for our interactions with the physical and social environment: it allows detection of events beyond our visual field, efficient re-orienting of multisensory attention (Pavani et al. 2017) and auditory scene analysis (Kerber & Seeber 2012; Shinn-Cunningham et al. 2017). Spatial hearing is three-dimensional, as it allows the estimation of sound directionality (i.e., its position in azimuth and elevation), and distance. It relies on the correct interpretation of the auditory cues reaching the ears, notably the binaural cues supported by interaural-level differences (ILD) and interaural time differences (ITD); and spectral cues filtered by each pinna (e.g., Brungart 1999; Middlebrooks 2015). It is well established that accuracy of the auditory system is better in azimuth than distance (e.g. Middlebrooks & Green 1991).

The type of stimulus (notably its spectral characteristics) and the acoustic environment (reverberant or anechoic) are known factors that can influence sound localization (e.g., Brungart 1999; Brungart et al. 1999). However, the distance between the listener and the sounds also play an important role. The availability of

¹Integrative Multisensory Perception Action & Cognition Team—ImpAct, Lyon Neuroscience Research Center, Lyon, France; ²Department of Pediatric Otolaryngology—Head & Neck Surgery, Femme Mere Enfant Hospital, Hospices Civils de Lyon, Lyon, France; ³Department of Otolaryngology—Head & Neck Surgery, Edouard Herriot Hospital, Hospices Civils de Lyon, Lyon, France; ⁴University of Lyon 1, Lyon, France; ⁵Hospices Civils de Lyon, Neuro-immersion Platform, Lyon, France; ⁶Center for Mind/Brain Sciences (CIMeC), University of Trento, Rovereto, Italy; and ⁷Department of Psychology and Cognitive Sciences, University of Trento, Rovereto, Italy. Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and text of this article on the journal's Web site (www.ear-hearing.com).

Copyright © 2021 The Authors. Ear & Hearing is published on behalf of the American Auditory Society, by Wolters Kluwer Health, Inc. This is an open-access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

auditory cues differs between the near-field (less than 1 m from the head) and the far-field (beyond 1 m). Sound directionality mainly relies on ITD in near and far spaces (Brungart 1999). Sound distance perception is specifically based on low-frequency ILDs (below 3 kHz) in the near-field, whereas other factors (e.g., room reverberation and spectral characteristics of the sound) are involved in the far-field (Middlebrooks & Green 1991; Brungart 1999). The closer the sound to the listener (less than 50 cm), the larger the low-frequency ILD is. That allows better accuracy in distance perception in the near-field compared to the far-field (Brungart & Rabinowitz 1999; Kolarik et al. 2016). Moreover, accuracy of distance perception is better for lateral than frontal sound sources (Brungart 1999; Kopčo & Shinn-Cunningham 2011).

Although spatial hearing critically relies on binaural and monaural cues (Middlebrooks 2015), interventions for auditory restoration may fail to fully preserve these auditory cues. For instance, it is well known that cochlear implant (CI) users mostly rely on ILD cues for sound localization, ITD cues are not fully available or are incorrectly processed by patients (Seeber et al. 2004, van Hoesel 2004). Likewise, while many technological advances have been made in recent years to promote speech understanding in noise, these same advances can have detrimental consequences on sound localization performance. One notable example is the automatic gain control (AGC) implemented in the CI processor which compresses the large intensity range of our auditory environment (i.e., from 0 to 120 dB HL) into a smaller intensity range (i.e., 10 to 20 dB HL) to comply with the more restricted electrical dynamics imposed by the CI (Dillon 2001). Hence, the AGC amplifies soft sounds, to make them better perceived by a listener, and it compresses loud sounds, to preserve the auditory system from discomfort and injury, aiming at a mean comfortable level (called the C-SPL) of 65 dB (Stöbich et al. 1999; Khing et al. 2013). This technology is valuable for speech understanding, but combines with other factors to increase distortions of ILDs, which are one of the few auditory cues available for CI users to localize sound sources (Archer-Boyd & Carlyon 2019).

These concerns are particularly relevant when one considers the developing auditory system of infants and children. The benefits of bilateral CIs (BCIs) in children with profound deafness have been documented in relation to several everyday life contexts (for review see Gordon et al. 2017), such as speech understanding in noisy environments (Misurelli & Litovsky 2012), opportunities for integration in mainstream schools (Gordon & Papsin 2009; Van Wieringen & Wouters 2015; Choi et al. 2020), and quality of life (Galvin & Mok 2016). To date, however, spatial hearing skills of BCI children and behavioral strategies to help promote these skills remain largely underinvestigated. The few studies that examined spatial hearing in this population reported substantial variability in children's abilities. Some documented performance at levels similar to normal-hearing (NH) listeners (Van Deun et al. 2010; Zheng et al. 2015; Killan et al. 2018), others reported chance level performance (Litovsky et al. 2004; Grieco-Calub & Litovsky 2010; Choi et al. 2017).

In addition to the limited number of pediatric studies, current approaches to spatial hearing limit our knowledge of localization capabilities in children. First, sound localization has only been examined in relation to responses in azimuth, without considering the two other dimensions of space (elevation and depth). Second, sampling of the auditory space has always been

constrained to the portion in front of the participant (Grieco-Calub & Litovsky 2010; Zheng et al. 2015; Choi et al. 2017), thereby neglecting back space. Sound perception in this region is fundamental because it is outside the visual field; hence, its monitoring relies crucially on the auditory modality alone. Third, all of the studies conducted so far examined BCI children in contexts in which the speaker array was visible throughout the experiment (see Table 1 for literature review). Any visual cue about the apparatus, even when seemingly uninformative, provides crucial priors to sound position (Da Silva 1985; Loomis et al. 1998). Visual priors are fully exploited during sound localization, as revealed by the study of visual dominance on spatial hearing in the context of audiovisual mismatch between perceptual cues (e.g., Kumpik et al. 2019). This implies that current data measured in laboratory may not fully reflect the actual performance of children in their daily life.

In addition to all these methodological aspects, two key features may limit our broader understanding of the auditory spatial skills of BCI children. First, previous studies with BCI children typically prevented head movements while listening to the sounds (see Table 1). Although most investigators did not use a chin rest, children were always instructed to face forward during sound delivery and refrain from moving their head. In some studies, the experimenter remained next to the child during the sound localization task to validate this constraint (Van Deun et al. 2010; Choi et al. 2017). Yet, it has long been acknowledged that head movements play a key role in sound localization (Wallach 1940). Humans use head movements to focus on the speaker in complex hearing situations (Wightman & Kistler 1999). In addition, head rotations and head translations (Wallach 1940; Perrett & Noble 1997) lead to changes in binaural cues which prove useful for resolving front-back confusion in people with NH (Brimijoin et al. 2010) and even more so in people with hearing impairment (Brimijoin et al. 2012). The impact of spontaneous head movements on auditory spatial performance has been highlighted in CI adults (Mueller et al. 2014; Pastore et al. 2018; Fischer et al. 2020), but it remains entirely overlooked in pediatric populations. This implies that current data measured in laboratory may be a limited approximation of the actual sound localization ability of children in their daily life, where head movements during sounds occur spontaneously.

The second key feature that remained overlooked is the study of near-by regions of auditory space. In the present study, we examined spatial hearing abilities in the near-field, referred to hereafter as "reaching space." As a matter of fact, while many studies have been conducted in the far-field in the last decades, the reaching space has been largely overlooked in children. This portion of space is particularly relevant for social interactions, where fast motor responses are needed in case of an approaching auditory object (e.g., a bee), when reaching toward a sound source (e.g., a musical toy) or when orienting towards a nearby talker (Kolarik et al. 2016). Finally, a recent study (Valzolgher et al. 2020a) has shown that the possibility to interact with near-field sounds by reaching to the sound sources directly can promote head movements during listening in case of simulated hearing impairment (monaural ear-plugging) and result in faster adaptation altered auditory cues.

In the present study, we examined (1) spatial hearing abilities in the near-field in BCI and NH children, considering front and back space, and different sound distances and (2) the impact of spontaneous versus restrained head movements during sound

TABLE 1. Literature review on sound localization in bilateral cochlear implant children

		Killan et al. (2018)	Choi et al. (2017)	Zheng et al. (2015)	Grieco-Calub and Litovsky (2010)	Van Deun et al. (2010)	Litovsky et al. (2004)
Population	Number of BCI children	10	13	19	21	30	3
	Age at test (yrs)	5–18	7–18	4–9	5–14	4–15	8–12
	Binaural experience (mo)	12	>12	13–51	3–28	12–44	3
	Simultaneous/sequential CI	All simultaneous	Unknown	3 simultaneous, 16 sequential	All sequential	All sequential	All sequential
Experience	Loudspeaker positions in azimuth	5 positions (–60° to 60°)	13 positions (–90° to 90°)	15 positions (–70° to 70°)	15 positions (–70° to 70°)	9 positions (–60° to 60°)	15 positions (–70° to 70°)
	Type of sound stimulus	Short sentence (3 words)	Bisyllabic word (“ja-yeon”)	Bisyllabic word (25 different)	Bisyllabic word (“baseball”)	1-sec bell-ring	Pink noise
	Head movements free	Yes	No	No	No	No	No
	Speaker visible	Yes	Yes	Yes	Yes	Yes	Yes
	Feedback	No	No	Yes	Yes	Yes	Yes
	Mean RMS error (degrees)	16.2 (range 11–22.6)	39.4 (range 30.6–50.5)	28.5 (range 13.8–47.6)	(range 19–56)	(range 13–63)	55 (range unknown)

BCI, bilateral cochlear implant; CI, cochlear implantation; RMS error, root mean square error.

emission on performance. To these aims and in the attempt to overcome existing limitation in the testing conditions, we developed a novel approach to near-field sound localization testing (European patent n°WO2017203028A1), based on virtual reality and real-time motion tracking that implements emission of real sounds. This allows for the study of spatial hearing with (1) very limited constraints on sound source locations and responses; (2) control of all available visual cues; (3) recorded pointing responses in 3D space; and (4) continuous recording of head movements. Within this virtual reality approach, the latter feature ensures reproducibility of sound source positioning across trials and participants, and it enables active listening strategies during and after sound delivery.

Performance of BCI users was compared to that of age-matched NH controls. Based on previous studies on sound localization and the known consequences of early auditory deprivation on the development of binaural processing (Litovsky 2015), we predicted that BCI children would perform worse than NH peers. Specifically, we expected errors to emerge both in azimuth and distance since CI processing limits ILD extraction which is one of the key auditory cues for sound localization in this population. Finally, we aimed to investigate the impact of head movements on sound localization performance by comparing a condition where the head was immobile during sound emission and another one where head movements were allowed. Based on previous results on hearing impaired adults (Mueller et al. 2014; Pastore et al. 2018), we predicted that BCI children would benefit from this active listening condition, especially when resolving front–back confusion.

MATERIALS AND METHODS

This clinical prospective study was approved by the Ethical Committee (Ile de France II, N° 18.09.19.37537 RIPH2), and recorded in clinicaltrials.gov (NCT03738592).

Participants

Eighteen children and adolescents aged between 8.3 and 16.7 years (mean age \pm SD: 12.1 \pm 2.7 years) fitted with two

CIs were recruited from a referral center for pediatric cochlear implantation. Inclusion criteria included age at testing over 8 years old, a minimum of 2 years of binaural experience to avoid large setting variations of their device, no areflexia and no attention disorders. Binaural experience ranged from 24 to 169 months (mean 67.3 \pm 32.4 months). BCI users were all bilateral CI daily users and had excellent monosyllabic word recognition performance (mean with left CI: 96.8 \pm 3.3%, and with right CI: 95.6 \pm 4.6%). Additional information about demographics and device settings (i.e., implants, sound processors, programming parameters, and sound coding strategies) are summarized in Table 2. Importantly, children wore their own CI processors (i.e., all were new generation processors including AGC, and two microphones, one at the front, and one at the rear). To test children with the settings, they were most familiar with in their daily life no parameters adjustments were made before testing. All but two children used the omnidirectional mode with their CIs. As we only used one sound source without back noise, this resulted in all BCI children being stimulated in the omnidirectional mode.

Eighteen age-matched NH controls were also recruited through advertisement to take part in the study (mean age \pm SD: 12.3 \pm 2.5 years; range 8.7–17 years). None of them had any history of hearing loss, middle ear problems, oculomotor or neurological disorders.

Apparatus and Stimuli

All tests were performed in a reverberant room (6 m \times 3 m, reverberation time RT_{60} : 0.36 s) which belongs to the Neuro-immersion research facility using a new virtual reality and motion tracking system (Fig. 1A). This is comprised of a head-mounted display (HMD) (HTC VIVE System, resolution: 1080 \times 1200 px, Field of View: 110°, Refresh rate: 90 Hz) and two tracked VIVE devices (one placed above a loudspeaker, the other on the pointer held by the participant). Tracking accuracy of the HTC VIVE System is adequate for behavioral research purposes (Verdelet et al. 2019). Specifically, the HTC VIVE has subcentimeter accuracy (9.0 mm when trackers are static; 9.4 mm when trackers are dynamic). Eye-tracking technology

TABLE 2. Demographics of bilateral cochlear implant children and device information

BCI group	Sex	Age at testing (mo)	Etiology	Onset of profound deafness	First cochlear implant				Second cochlear implant				Duration of binaural experience (mo)	
					Age at CI1 (mo)	Internal Part, Processor, Ear side	Strategy	Microphone	Age at CI2 (mo)	Internal Part, Processor, Ear side	Strategy	Microphone		Interimplant interval (mo)
B01	M	136	Genetic	Prelingual	28	Digisonic SP EVO, SAPHYR SP, L	Crystallis XDP	Omnidirectional	91	Digisonic SP EVO, SAPHYR SP, R	Crystallis XDP	Omnidirectional	63	44
B02	F	133	Genetic (connexin 26)	Perlingual	40	Nucleus CI512, CP950, R	ACE	Omnidirectional	68	Freedom CI24RE, CP950, L	ACE	Omnidirectional	28	64
B03	F	146	Genetic (connexin 26)	Prelingual	13	Freedom CI24RE, CP1000, R	ACE	Omnidirectional	74	Freedom CI24RE, CP1000, L	ACE	Omnidirectional	61	72
B04	M	178	Meningitis	Prelingual	8	Nucleus CI24R, CP950, L	ACE	Omnidirectional	9	Nucleus CI24R, CP950, R	ACE	Omnidirectional	1	169
B05	F	139	Unknown	Perlingual	90	Nucleus CI422, CP910, R	ACE	Directional	107	Nucleus CI422, CP910, L	ACE	Directional	17	32
B06	F	158	Unknown	Prelingual	34	HiRes90K, Naida, L	HiRes Optima-S	Omnidirectional	108	HiRes90K Advantage, Naida, R	HiRes Optima-S	Omnidirectional	74	49
B07	M	162	Meningitis	Postlingual	52	Digisonic SP, SAPHYR SP, L	MPIS XDP	Omnidirectional	52	Digisonic SP, SAPHYR SP, R	MPIS XDP	Omnidirectional	0	110
B08	M	110	Unknown	PRELINGUAL	25	Freedom CI24RE, CP1000, R	ACE	Directional	45	Freedom CI24RE, CP1000, L	ACE	Directional	20	64
B09	M	207	Unknown	Prelingual	38	Nucleus CI24R, CP910, R	ACE	Omnidirectional	140	Freedom CI24RE, CP810, L	ACE	Omnidirectional	102	67
B10	F	104	Cytomegalovirus	Prelingual	19	Freedom CI24RE, CP1000, L	ACE	Omnidirectional	31	Freedom CI24RE, CP1000, R	ACE	Omnidirectional	12	73
B11	F	154	Unknown	Prelingual	48	Nucleus CI512, CP950, L	ACE	Omnidirectional	91	Freedom CI24RE, CP950, R	ACE	Omnidirectional	43	62
B12	M	122	Unknown	Prelingual	15	Nucleus CI512, CP1000, R	ACE	Omnidirectional	48	Freedom CI24RE, CP1000, L	ACE	Omnidirectional	33	74
B13	M	106	Unknown	Prelingual	75	HiRes90K Advantage, Naida, L	HiRes Optima-S	Omnidirectional	81	HiRes90K Advantage, Naida, R	HiRes Optima-S	Omnidirectional	6	24
B14	F	162	Unknown	Perlingual	65	Digisonic SP EVO, SAPHYR SP, L	Crystallis XDP	Omnidirectional	106	Digisonic SP EVO, SAPHYR SP, R	Crystallis XDP	Omnidirectional	41	55
B15	M	200	Genetic (connexin 26)	Prelingual	28	Nucleus CI24R, CP910, R	ACE	Omnidirectional	127	Freedom CI24RE, CP1000, L	ACE	Omnidirectional	99	72
B16	M	117	Unknown	Perlingual	55	Concerto FlexSoft, OPUS 2, L	FS4-p	Omnidirectional	84	Synchro Standard, SONNET, R	FS4-p	Omnidirectional	29	32
B17	F	174	Genetic	Prelingual	17	Freedom CI24RE, CP1000, R	ACE	Omnidirectional	101	Freedom CI24RE, CP1000, L	ACE	Omnidirectional	84	73
B18	M	100	Genetic (connexin 26)	Prelingual	12	Freedom CI24RE, CP1000, R	ACE	Omnidirectional	23	Freedom CI24RE, CP1000, L	ACE	Omnidirectional	11	75

BCI, bilateral cochlear implant; CI1, first cochlear implantation; CI2, second cochlear implantation; L, left; R, right.

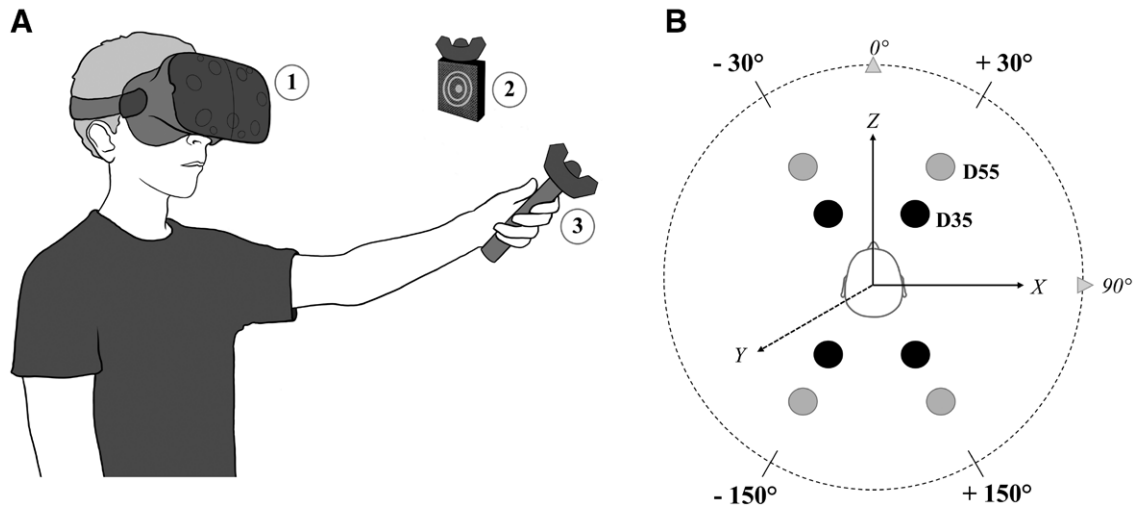


Fig. 1. Experimental setup. A, Apparatus based on the virtual reality system, comprising (1) a head-mounted display (HTC VIVE), (2) a VIVE tracker mounted on a loudspeaker, and (3) another tracker mounted on a hand-held pointer. Head and trackers positions were recorded in real time by two cameras, and defined in a head-centered system. B, Sound localization setup. Black and gray circles indicate two target distances in reaching space, at 35 cm (D35) and 55 cm (D55). Three axes were defined according to the reference frame (i.e., participant head-centered): X, azimuth; Y, elevation; and Z, distance.

(SensoriMotoric Instruments, Berlin, Germany; www.smivision.com; 60 Hz frequency and 0.5 degrees spatial precision) was added to the HMD to allow monitoring of initial eye position throughout testing. The whole setup was controlled using a custom program made with Unity (2017.4.10f1) which enabled: (1) control of the HMD visual display; (2) accurate positioning of the loudspeaker at the beginning of each trial; (3) control over sound delivery; and (4) recording of the exact position of all tracked elements (loudspeaker, participant's head with the HMD, participant's hand, and cyclopean gaze position) in three-dimensional space (azimuth, elevation, and distance). The position of all tracked elements (i.e., the HMD and VIVE trackers) was expressed in a head-centered reference frame. The head-centered position was calculated by collecting the 3D position of both ears (using the VIVE controller), and averaging these positions to obtain the origin of the head-centered system.

Only one loudspeaker (mini speaker model JBL GO Portable from HARMAN International Industries, Northridge, CA; 68.3 × 82.7 × 30.8 mm, Output Power 3.0 W; frequency response: 180 Hz–20 kHz) was used to sample the auditory space around the head. With the tracker fixed on it, the loudspeaker position was tracked and controlled in real time. For each trial, the experimenter held the loudspeaker in his hand and moved it to the desired position (predetermined and defined in head-centered coordinates). The loudspeaker coordinates were indicated by two main cues delivered to the experimenter: (1) visual cues (on a computer screen for azimuth and distance position) and (2) an echo radar signal (for elevation, given by an in-ear headphone nonaudible to the participant). Hence, this enabled the choice of whatever position in space without any physical constraints on loudspeaker placement.

Motor responses were obtained using a hand-held pointer (i.e., a wand with a tracker fixed on it). Participants were informed that they had to place the pointer in the exact position of the perceived sound. This allowed us to record children's responses in 3D, considering azimuth, elevation and distance perception.

The sound stimulus was 3 seconds of white noise, modulated in amplitude at 2.5 Hz frequency (modulation depth at 80%), and delivered at 70 dB SPL when measured one meter from the participant's head (corresponding to 76.6 dB SPL at 35 cm, and 73 dB SPL at 55 cm). The background noise measured at the beginning of the experiment was 33.7 dB SPL.

Unbeknownst to the participant, 8 predetermined sound positions within reaching space were used so that they could reach each sound source without leaving their chair. The loudspeaker could be located at +30°, +150°, -30°, and -150° in azimuth (see Fig. 1B positive values indicating right space and negative values left space) with respect to the participant's straight ahead. Two distances were evaluated in the near-field (reaching space) for each azimuthal position: D35, at 35 cm (13 3/4") from the participant's head, and D55 at 55 cm (21 5/8"). Elevation remained constant throughout the study at ear level but children were unaware of this.

Procedure

Before the experimental session, we collected self-report data from BCI participants on their perceived spatial hearing abilities in daily life. We administered an adapted version of the Speech, Spatial, and Qualities of Hearing Scale (SSQ) to BCI children (SSQ-Child, Galvin & Noble 2013) and their parents (SSQ-Parents, Galvin & Noble 2013). Each SSQ questionnaire was divided into three subscales: (1) speech perception; (2) spatial hearing; and (3) other qualities of hearing. Questionnaire data were collected face to face with children and over the phone with parents, three times 1 week apart. Hence, an observation period of 1 week was allowed for parents to fill out each subscale after paying attention to their child's behavior in the less common situations of daily life.

Before testing, all participants were introduced to the apparatus and the sound localization task by watching a short video. The video informed them that they would be sitting on a rotating chair which allowed them to easily access all sound sources around them, and they would localize a sound source delivered in the near-field within their reachable space. At the end of each

trial (sound emission), the experimenter rapidly removed the loudspeaker to avoid any collision during the pointing phase. For this reason, children were instructed to wait until the sound ended before pointing.

Upon entering the experimental room, the children saw the environment and objects present in the physical space, but received no prior information about the loudspeaker positions used in the study. They were invited to wear the HMD and they underwent a 5-minute training session to ensure that the instructions in the video had been fully understood. This training included different elevations (high and low), distances (near the head and arm outstretched), and azimuthal positions (left/right, and front/back). Participants were encouraged to point either with the left or right hand in order to comfortably reach the sound's perceived position. Participants clearly saw that there were no visual clues in the HMD to help them guess the exact location of sound targets. If no discomfort was reported, the experimental session then began.

Two experimenters were present in the testing room: the first one placed the loudspeaker in the position indicated by the computer and the second one provided information or explanations to the children during the test if necessary. The noises produced by the first experimenter while placing the loudspeaker were not informative (for details, see Supplemental Digital Content 1, <http://links.lww.com/EANDH/A846>, which shows a control experiment).

Each testing session included two listening conditions: head immobile (HI) and head moving (HM) during sound delivery. Under the “head immobile” (HI) condition, participants were instructed to keep their head fixed during sound emission and were only allowed head movements during the response phase. Under the “HM” condition, participants were encouraged, but not forced, to move their head both during sound emission and during the response phase. Overall, the experiment lasted 20 minutes and consisted of 96 trials: 12 repetitions for each sound position, equally distributed among the two listening conditions. An ABBA counterbalancing scheme was used to control for the effects of listening condition order (HM-HI-HI-HM or HI-HM-HM-HI), each block comprising 24 trials. For each group (BCI and NH), half of the participants followed the first order, whereas the other followed the second one.

To ensure reproducibility of trials, several controls were used, specifically for head alignment since no chin rest was used. Hence, at the beginning of each trial, three concomitant conditions had to be validated to trigger sound emission: (1) the head and (2) eyes were aligned with the participant's midsagittal plane, (3) the loudspeaker fell within the predetermined 3D location (see video and comments in Supplemental Digital Content 2, <http://links.lww.com/EANDH/A847> and Supplemental Digital Content 3, <http://links.lww.com/EANDH/A846>, which show the two listening conditions).

To summarize, in this experiment we explored the ability of children to discriminate different sound positions in azimuth and in distance and their answers were evaluated by manual pointing in 3D.

Data Analysis

Kinematic analyses on head and hand movements were performed with a custom-written MATLAB program. This allowed us to reject trials where participants anticipated hand movements during sound emission or performed head movements

during the HI condition. As these analyses were performed after the experiment, rejected trials were not replaced.

Statistical analyses and data visualization were performed using the R-studio environment (www.rstudio.com). First, we separately analyzed sound performance in the azimuthal plane and in distance. For azimuth, the percentage of front-back and left-right confusions was assessed as a function of listening condition for each participant. We also calculated the mean absolute error in degrees for each participant as a function of the real sound position (i.e., error between the sound source and the hand pointing response). For distance, we assessed discriminability between two nearby distances, D35 and D55 within each group.

As we recorded a motor response in 3D, we then computed a variable called the 3D error, corresponding to the difference in space between the X-Y-Z position of the loudspeaker and the participant's final hand position used to indicate the localization of the sound's source. This was adapted from the system introduced by Rakerd and Hartmann (1986) (see also Grantham et al. 2008). The 3D error was calculated for each participant, collapsing across all sound positions, according to the formula: $3D\ error = \sqrt{\underline{C}^2 + \underline{s}^2}$, where \underline{C} is the mean of the vector norm between the sound source and the hand pointing position, and \underline{s} is the standard deviation. The 3D error mean and standard deviation were computed irrespective of sound position. The main advantage of the 3D error is that it combines response errors across the three spatial dimensions, and also takes into account each participant's response variability.

Finally, to evaluate the improvement in sound localization by head motion in the HM condition we calculated an index of Listening Improvement “*I*” for each participant with the following formula: $I = (HI\ 3D\ error - HM\ 3D\ error) / HI\ 3D\ error$. The HI 3D error corresponded to the 3D error under the HI condition, and the HM 3D error to the 3D error under the HM condition. The index “*I*” was normalized for each participant which allowed us to have an objective measure of each participant's improvement independently of their baseline performance under the HI condition. We also correlated “*I*” with the percentage of trials where at least one movement was performed during sound emission under the HM condition.

RESULTS

Overall, 2.6 % of trials were rejected in the NH group and 5.2% in the BCI group. Trials were rejected either for head movements during sound emission under the HI condition (2.8% of trials for NH children and 4% for BCI children) or for anticipated hand movement responses in either of the listening conditions (2.4% of trials for NH children and 6.4% for BCI children). This indicates that all participants easily complied with the instructions and easily adapted to our novel methodology for spatial hearing measurement.

Condition HI: HI During Sound Emission

First, we compared sound localization performance between groups with the H during sound emission. Figure 2 shows mean responses for NH and BCI children. Figure 2A shows sound discrimination as a function of side (i.e., left or right space) and distance (i.e., 35 or 55 cm). Almost all trials BCI children were as good as NH children in discriminating sounds from left and right space, confusing stimulus side in only 7.9% of cases

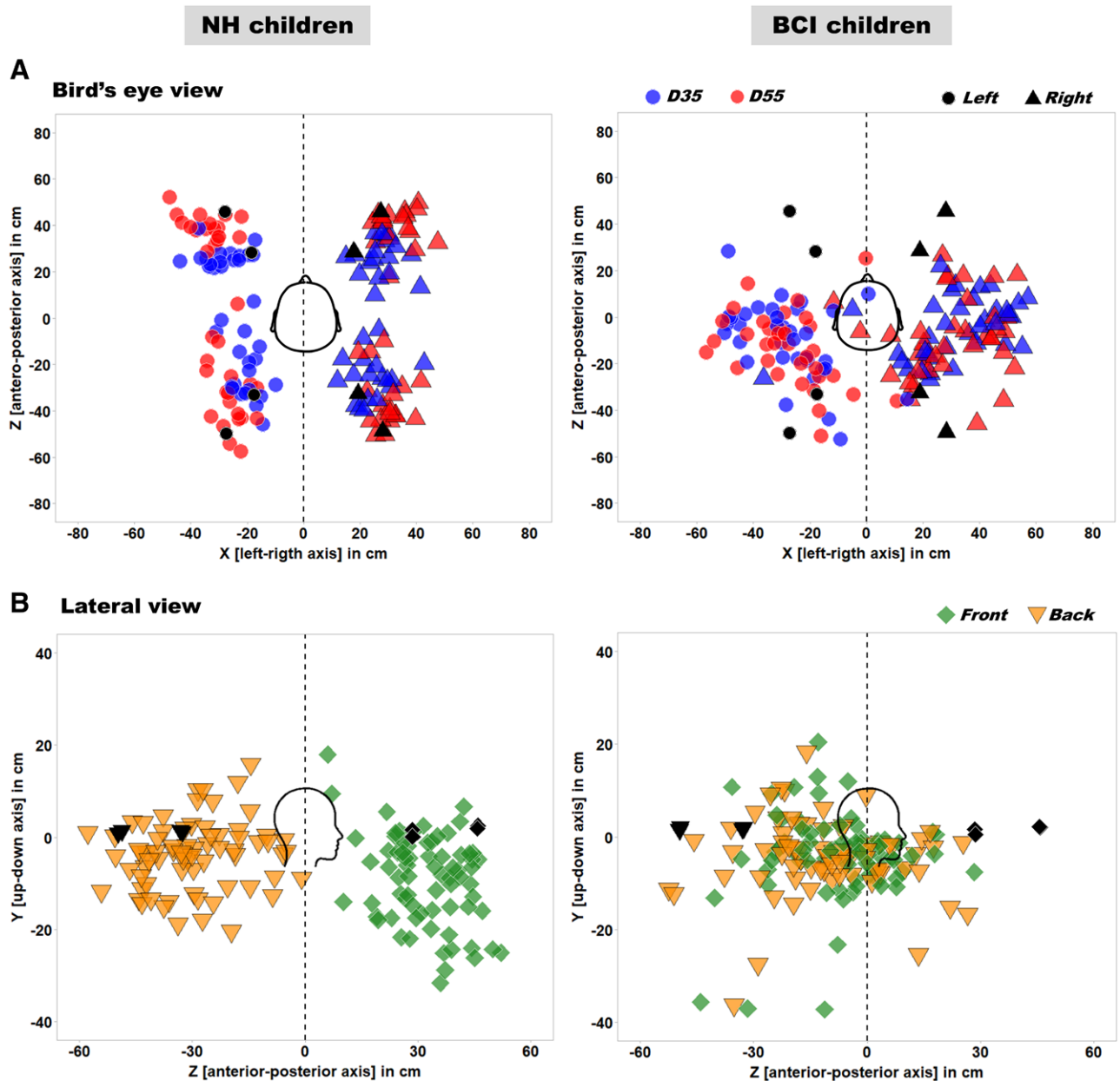


Fig. 2. Three-dimensional sound localization performance of normal-hearing (NH) and bilateral cochlear implant (BCI) children under the head immobile condition. Black symbols represent the sound sources and colored dots correspond to the mean response of each participant per target. A, Bird's eye view showing hand responses as a function of stimulation side (circles for left sounds and triangles for right sounds) and distances (blue and red for 35 and 55 cm sound sources, respectively). B, Lateral view showing hand responses as a function of front stimulation (green diamonds) and back stimulation (yellow triangles).

(compared to 0.5% for NH peers). This figure also shows that their hand pointing was more lateralized, without a clear distinction between the four stimulation quadrants as was the case for the NH children. The separation of the blue (35 cm) and red (55 cm) points for NH but not for BCI children demonstrates that sound discrimination in distance is possible for NH children both in front and behind, whereas it seems to be more difficult for BCI children.

Figure 2B shows front-back discrimination. This figure shows that NH children easily segregated front sources from back sound sources but that this ability was degraded in BCI children, who mainly pointed towards the interaural axis next to their CIs. This led to substantial front-back confusions, 44.6%

(range 31.7–62.5), whereas NH children confused front and back space in only 2.1% of cases (range 0–16.7; Chi-squared test, $p < 0.001$). Front-back localization accuracy in BCI children was close to chance performance.

To characterize sound localization performance as a function of group and sound position we first focused on separate space dimensions. The mean absolute error in azimuth was overall higher for BCI compared to NH participants (BCI: $58.8^\circ \pm 8.4^\circ$; NH: $16.2^\circ \pm 6.5^\circ$; main effect of group: $F(1, 34) = 298.6$, $p < 0.001$). Figure 3 shows sound localization as a function of sound distance in each group. Sounds located at 55 cm from the center of the head were perceived on average at 47.5 ± 13.2 cm by the BCI group and at 50.9 ± 7.1 cm by the NH group. The

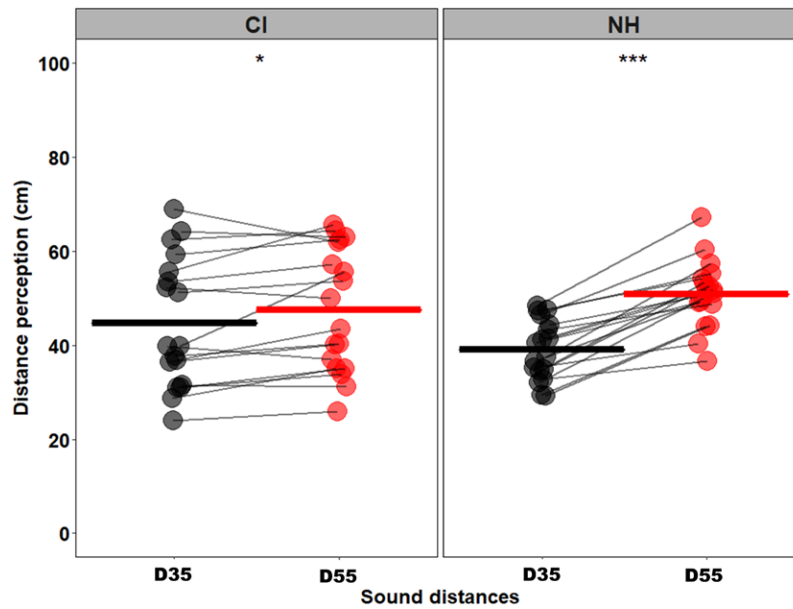


Fig. 3. Sound distance perception in normal-hearing (NH) and bilateral cochlear implant (BCI) children under the head immobile condition. Thick lines represent the mean response distances for each group for D35 (black lines: i.e., sound sources at 35 cm), and D55 (red lines: i.e., sound sources at 55 cm). Thin black lines join black and red dots for each participant. Asterisks indicate significant differences (paired t-test, * $p < 0.05$; *** $p < 0.001$).

closest sounds (located at 35 cm) were perceived on average at 44.6 ± 13.9 cm by the BCI group and at 39 ± 6.3 cm by the NH group. Notably, both groups were able to perceive a difference between D35 and D55 stimulations (on paired t-test: BCI: $t(17) = -2.417, p = 0.027$; NH: $t(17) = -9.471, p < 0.001$), but this was clearer in NH children.

We next examined the 3D error which took into account all three spatial dimensions. Table 3 summarizes the 3D error as a function of group (NH or BCI), listening condition (HI or

HM), front–back position, and stimulation side (left or right). Individual data show substantial variability across participants, mainly in the BCI group (see Tables in Supplemental Digital Contents 4 and 5, <http://links.lww.com/EANDH/A846>). Under the HI condition, the 3D error was higher in BCI (mean \pm SD: 55 ± 13.3 cm) compared to NH participants (24.2 ± 5.6 cm). An ANOVA with group and the front–back position as variables, revealed a significant two-way interaction ($F(1, 34) = 8.04, p = 0.008$) caused by larger 3D errors for front than back targets in the BCI group ($p = 0.0002$, Bonferroni corrected) but not in the NH group ($p = 0.868$, Bonferroni corrected). This reflects the large number of front-to-back confusions in the BCI group reported above. A similar analysis with group and stimulation side as variables revealed only the main effect of group ($F(1, 34) = 0.86, p = 0.359$) caused by larger errors overall for BCI compared to NH participants.

TABLE 3. Sound localization performance in NH and BCI children

	3D error (cm)	
	BCI	NH
	Mean (SD)	Mean (SD)
HI		
Front		
Left	59.8 (15.2)	23.2 (7.4)
Right	60 (14.5)	24.5 (8.5)
Back		
Left	46.8 (17.2)	24 (7.1)
Right	50.1 (19.3)	22.8 (8.3)
Overall	55 (13.3)	24.2 (5.6)
HM		
Front		
Left	48.2 (17.2)	23.8 (6.2)
Right	46.7 (14.7)	25.6 (7.6)
Back		
Left	44.4 (18.5)	23 (7.1)
Right	42.2 (16.2)	21.3 (6.8)
Overall	46.4 (13.1)	23.7 (5.6)

Performance errors are expressed in 3D (3D error), which represents the absolute and variable errors for the three spatial dimensions in each group. The mean errors were segregated with respect to listening condition, HI and HM, front and backspace, and stimulation side (left and right). These values were then combined to create an overall error for each listening condition.

BCI, bilateral cochlear implant; HI, head immobile; HM, head moving; NH, normal hearing.

Condition HM: Head Movements During Sound Emission

Under the HM condition, children were free and encouraged to move their head during sound emission. We evaluated the impact of head movements on localization performance in the same way as for the HI condition. Figure 4A and C shows mean responses for BCI children as a function of side, distance, and front–back sound sources. Even when head movements were allowed sound discrimination remained difficult, D35 was perceived on average at 45.4 ± 13.9 cm, and D55 at 48.7 ± 14.1 cm, but discrimination of front–back sources improved. This is illustrated in Figure 4D, which shows a significant decrease in front–back confusions, from 44.6% to 25.1%, under HI and HM conditions (McNemar test, $p < 0.001$). The individual data show that performance improved in 12 out of 18 children (67%) and remained stable for the other six. Notably, there was a significant increase in left–right confusions from 8.3% to 15.2% (McNemar test, $p < 0.001$; see Fig. 4B), mainly in six children (33%) for

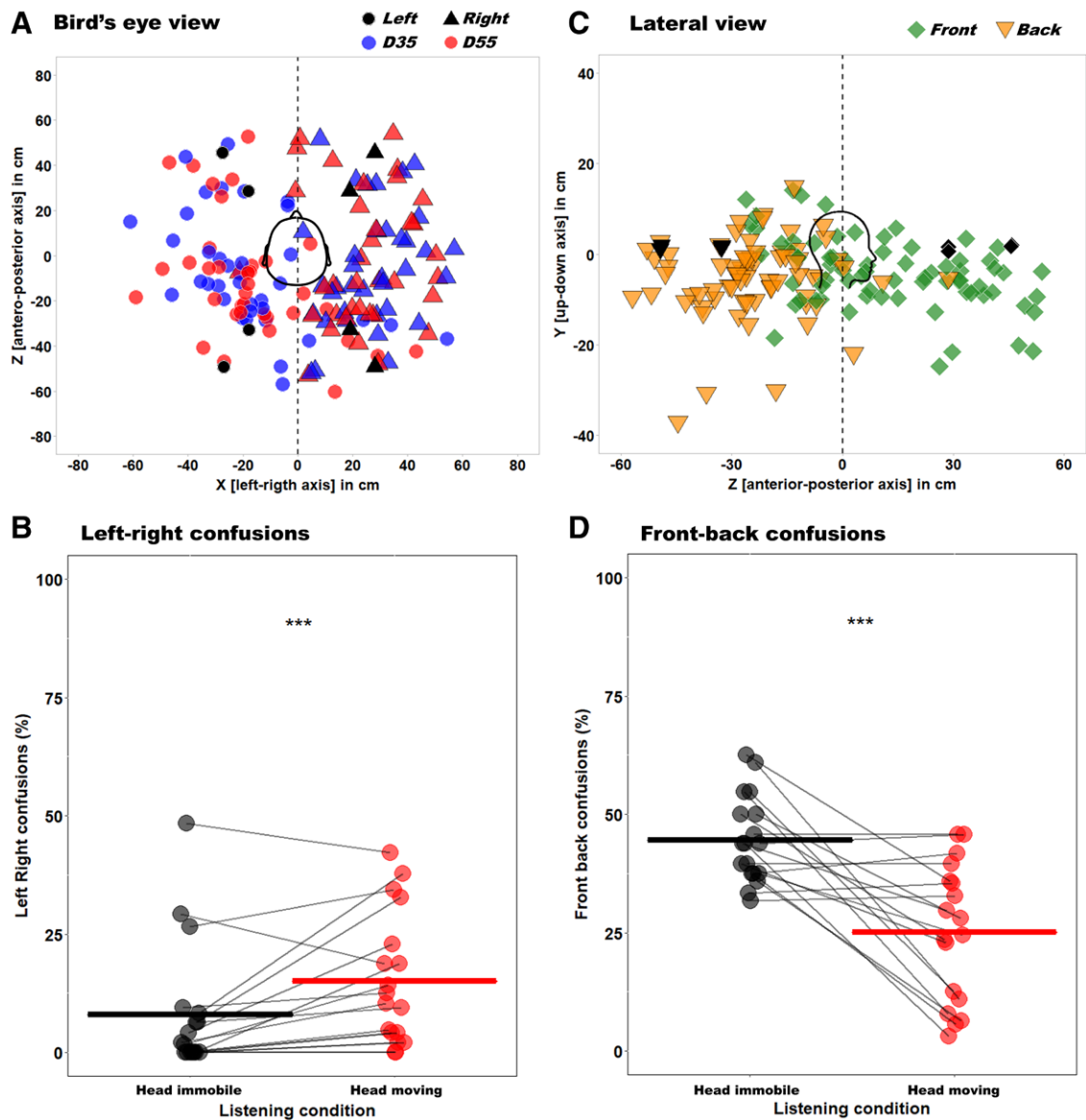


Fig. 4. Three-dimensional sound localization performance of children fitted with bilateral cochlear implant during the head moving condition. A, Bird's eye view showing hand responses as a function of stimulation side (circles for left sounds and triangles for right sounds) and distances (blue and red for 35 and 55 cm sound sources, respectively). Black symbols represent the sound sources and colored dots correspond to the mean response of each participant per target. B, Left-right confusions as a function of listening condition. Thick black lines represent the mean percentage of confusions when head movements were forbidden, and the thick red line when head movements were free during sound emission. Thin black lines join black and red dots for each BCI participant. C, Lateral view showing hand responses for front-back stimulations. D, Front-back confusions as a function of listening condition. Asterisks indicate significant differences (McNemar test, $***p < 0.001$). BCI, bilateral cochlear implant; NH, normal hearing.

back sound sources (Fig. 4A). Visualization of pointing data for NH participants are available in Supplemental Digital Content 6, <http://links.lww.com/EANDH/A846>; no significant changes occurred in this group.

Figure 5A illustrates 3D error changes for each participant under HI and HM listening conditions. The 3D error significantly decreased in the BCI group by 8.6 cm (paired t-test, $t(17) = 3.41$, $p = 0.003$) compared to the NH group who only decreased by 0.5 cm (paired t-test, $t(17) = 0.84$, $p = 0.412$). A mixed ANOVA on the 3D error, with group, listening condition, and front-back position as variables revealed a significant three-way interaction ($F(1, 34) = 4.89$, $p = 0.034$). The significant improvement caused by active listening concerned specifically the BCI group

when pointing to front targets ($p < 0.001$, Bonferroni corrected) but not when pointing to back targets ($p = 0.15$, Bonferroni corrected). No significant improvement emerged for the NH group, either for front targets ($p = 0.98$, Bonferroni corrected) or back targets ($p = 0.95$, Bonferroni corrected). When participants were free to move their heads during sound delivery, we noticed that head behavior differed across trials and participants. Hence, we assessed if the percentage of trials with at least one head movement during sound emission correlated with a change in localization performance. This change was computed as a Listening Improvement index, considering individual performance under the HI condition (see Methods section for details). We found a strong positive correlation between the Listening Improvement

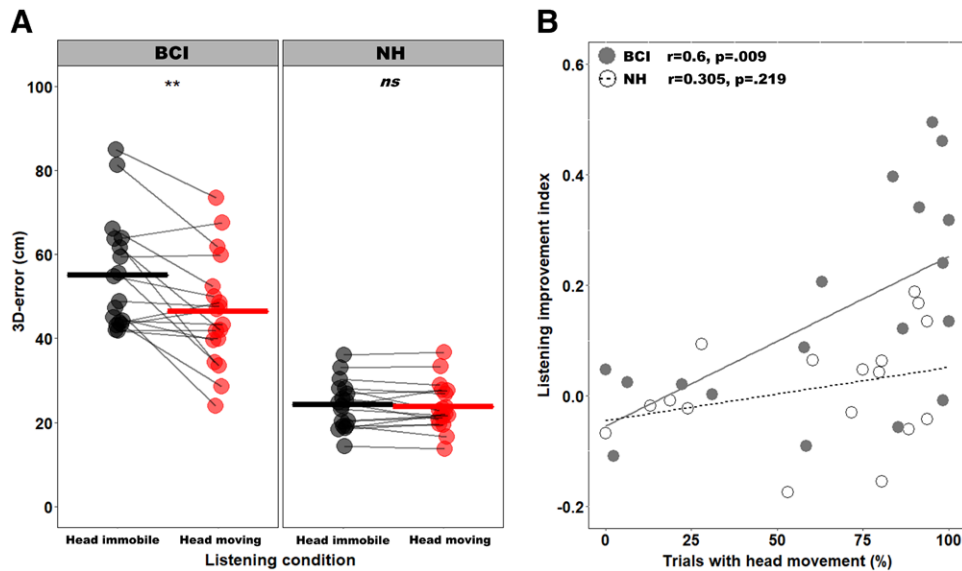


Fig. 5. Effect of head motion on spatial performance. A, Three-dimensional 3D error in both groups (BCI and NH) as a function of listening condition. Thick lines represent the mean 3D error within each group during HI listening (black) and HM listening (red) condition. Thin black lines join black and red dots for each participant. Asterisks indicate significant differences (paired t-test, $**p < 0.01$). B, Listening improvement index as a function of the percentage of trials with at least one head movement during sound emission. BCI, bilateral cochlear implant; NH, normal hearing.

index and the percentage of trials with head movements for the BCI group (Pearson correlation, $r = 0.6, p = 0.009$; see Fig. 5B), whereas no correlation emerged for the NH group (Pearson correlation, $r = 0.305, p = 0.219$). This suggests that exploring with the head had a positive effect on localization performance of BCI children. This effect was not found in the NH group, most likely because performance under the HI condition was already at ceiling for this group.

Clinical Predictors for Sound Localization Performance

We first explored the effect of clinical predictors on sound localization performance in the BCI group in the HI listening condition, for comparison with previous work that also adopted a similar listening condition (Killan et al. 2019). Specifically, we focused on age at test, interimplant interval, and duration of binaural experience (i.e., the time between the second CI and age at test). The 3D error was correlated with interimplant interval (Pearson correlation, $r = 0.721, p < 0.001$; see Fig. 6), but also with age at test (Pearson correlation, $r = 0.693, p = 0.001$). As these two clinical variables were also correlated with one another (Pearson correlation, $r = 0.659, p = 0.003$), we ran a partial correlation to test whether the relation between the 3D error and interimplant delay held up when controlling for the effect of age. The partial correlation remained significant ($p = 0.047$), reinforcing the conclusion that the greater the interimplant delay, the greater the 3D error is. We also ran a partial correlation between the 3D error and age at test, after controlling for interimplant delay. The partial correlation was not significant ($p = 0.096$). There was no correlation between the 3D error and the duration of binaural experience (Pearson correlation, $r = 0.041, p = 0.872$). Noticeably in the NH group, we found a significant negative correlation between age at test and 3D error (Pearson correlation, $r = -0.541, p = 0.02$). The older the NH children, the lower the 3D error is.

Finally, we examined if the results of the SSQ questionnaire correlated with the 3D error. No significant correlation between the mean score of subscale B (specific to spatial hearing) emerged, for children (Pearson correlation, $r = 0.04, p = 0.874$) or parents ($r = -0.12, p = 0.644$). Comparing self-assessment of all items by children to their parents, we found a positive correlation ($r = 0.53, p = 0.024$) with a mean score slightly higher for children (7.2 points; range 5.8–8.6) than parents (6.7 points; range 4.7–8.3).

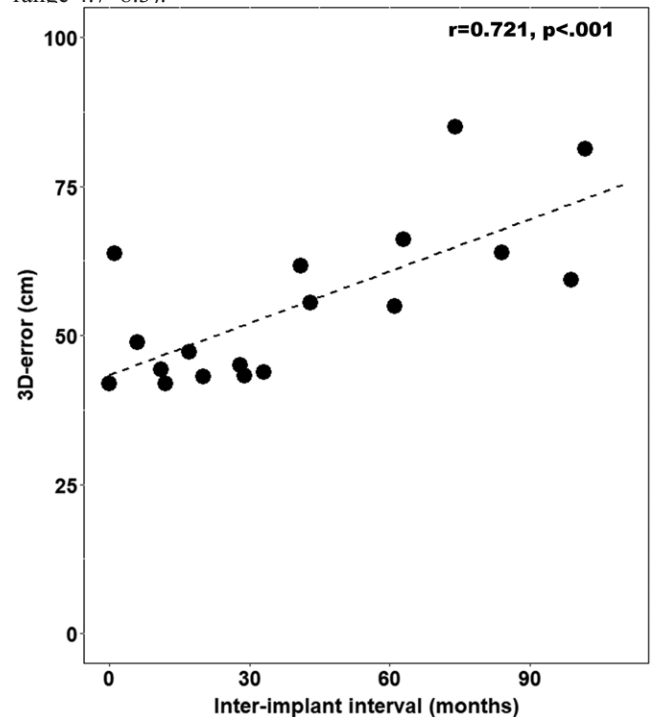


Fig. 6. Three-dimensional error (in centimeters) under the head immobile condition as a function of interimplant interval (in months).

DISCUSSION

The present study examined sound localization abilities of young BCI users (aged 8–17 years old) compared to NH peers in reaching space. This is a portion of space in which many social interactions occur in daily life, and in which we interact with sound sources through avoidance and reaching movements (e.g., Valzolgher et al., 2020a). To this aim we used a portable virtual reality and kinematic tracking system which allowed us to deliver real sounds within reaching distance from the listener, at reproducible head-centered coordinates. This same equipment also allowed us to pursue our additional aim, the investigation of the role of spontaneous head movements in spatial hearing abilities during sound emission.

Spatial Hearing Abilities for BCI Users

In our sound localization experiment, sound sources were presented laterally, in front and back space. Without free head exploration during sound presentation, BCI children performed at chance in front/back discrimination, pointing consistently along the interaural axis. This phenomenon has already been observed in adult CI users (Mueller et al. 2014; Pastore et al. 2018), and it could largely be attributed to the ambiguity introduced by the omnidirectional mode of the CI microphone settings which merges front and back sound sources (Fischer et al. 2020).

Until now, studies on sound localization in CI children only reported absolute azimuthal errors in front space, and these did not exceed 39.4° (Zheng et al. 2015; Choi et al. 2017; Killan et al. 2018). This is substantially lower than the azimuth error we observed here (i.e., mean absolute error of 59°). This difference could reflect the greater uncertainty about sound position that participants may have experienced in our study, since they were informed that sounds could be delivered anywhere around them within reaching distance. In addition, they gave their responses in 3D (i.e., by indicating a point in space with a hand-held pointer) and they did not receive any feedback after their answer. Most importantly, no visual cues were available to support sound localization, unlike previous reports which used a touch screen for response validation and all loudspeaker positions were visible (Grieco-Calub & Litovsky 2010; Zheng et al. 2015; Killan et al. 2018).

We also explored sound perception in depth by favoring accessibility of auditory cues to BCI children. We opted for a lateral arrangement of sounds, as this gives an advantage for distance judgements compared to medial sources (Kopčo & Shinn-Cunningham 2011). Moreover, in the near-field, distance perception mostly relies on low-frequency ILD (Brungart 1999; Seeber et al. 2004), which are partially restored binaural cues in BCI users. Despite these aspects, compared to NH children, BCI children had substantial difficulties in discriminating sound sources spaced 20 cm in depth. They perceived a mean difference of 3 cm between the two sources versus the 11 cm observed for NH children. Recent findings suggest that the AGC currently implemented in new generation CI processors are mostly unsynchronized and degrades ILDs by applying a broadband compression above loudness threshold (Dorman et al. 2014). Depending on the position of the sound source and its loudness, the compression can apply for the CI closest to the sound source, and not for the contralateral one (because the sound level becomes below the compression threshold with the

head shadow effect). Moreover, the AGC can increase the low-frequency components of the CI close to the sound source, thus leading to inverted ILDs (Dorman et al. 2014; Archer-Boyd & Carlyon 2019). These situations could lead to instances of maximal distortions of ILDs. In the present study, it is possible that BCI children were able to extract ILD cues to perceive small variations between close distances but these cues were occasionally too distorted to provide veridical information for each distance. Interestingly, we also noticed substantial variability in distance judgments in both groups, suggesting a differential maturation between children that continues at least until adolescence.

In sum, these findings highlight limitations of sound localization abilities of BCI children in azimuth and distance. However, it is likely that children are less negatively affected than expected, possibly because in their daily life they rely on multisensory perception. In recent reviews, several authors have noted that vision plays a key role in calibrating spatial hearing abilities (Tabry et al. 2013; Valzoghger et al. 2020b), especially during the first 10 years of development (King 2009). This is particularly important for azimuthal and distance perception of a stimulus (Zahorik 2001; Calcagno et al. 2012), and to a lesser extent for vertical judgments (Shelton & Searle 1980). To have a perception of the environment that is as accurate and consistent as possible, individuals with a hearing impairment tend to use vision to compensate for spatial hearing difficulties (King 2009). The children's reports support this assumption, since none of the sound localization difficulties demonstrated here emerged in the self-reported experiences of BCI users or their care-givers. Indeed, neither BCI children nor their parents reported major spatial hearing impairments in daily life on the SSQ questionnaire (Galvin & Noble 2013).

An Important Role of Head Movements for BCI Children

The pioneering works by Wallach in NH adults has shown that head movements, a natural orienting behavior (Kim et al. 2013), are helpful to disambiguate front and back sound sources, especially if they are close to the midline (Wallach 1940). Head motion creates important changes in binaural cues (ILD and ITD) by increasing information about level and time differences perceived by each ear. Until now, it was unclear if BCI children could extract relevant auditory information from head movements, or instead they might be disturbed by them. A study conducted by Mueller et al. (2014) in seven BCI adult users highlighted that head movements did not give them any benefits to sound localization accuracy when short speech sentences (less than 1 second) were delivered in a background noise of 60 dB SPL. However, a significant decrease of front-back confusions was observed for longer sentences (2 and 4.5 seconds) but without improvement of the absolute angular accuracy.

In the present study, BCI children significantly decreased their overall 3D error (from 55 to 46 cm) and their front-back confusions (from 44.6% to 25.1%) when they were allowed to perform spontaneous head movements during a 3-second sound emission. This suggests that despite their limited access to auditory cues, BCI children are able to extract and interpret binaural dynamic differences induced by head movements. These results are in line with a recent study in which a similar improvement of front-back confusions (from 41.9% to 6.7%) was demonstrated

in BCI adults (Pastore et al. 2018). Head behavior in BCI children was unrelated to their age at test. Instead, the higher the percentage of trials on which BCI children spontaneously moved their head during sound delivery, the greater their improvement in overall performance. For NH children, no benefit of head motion emerged, probably because they were already performing at ceiling with their head still in our experimental conditions. In further studies, it would be interesting to repeat the test in a more complex auditory environment (e.g., with competing background noise) to uncover the extent of head motion benefits in this pediatric population.

Individual Variability

The study of sound localization abilities in children is relatively recent, with less than two decades of research conducted in NH children of the age range we tested here (Litovsky et al. 2004). For this reason, several questions remain open, including the degree of maturation of the brain circuits involved in this specific task in NH children. In the present study, we found a negative correlation between the 3D error and chronological age at test in NH children. This suggests that the multiple sensory and cognitive components contributing to spatial hearing—from auditory processing to mapping sounds in spatial coordinates—undergo maturation processes that continue until adolescence, and may thus be more complex than expected based on the results of previous studies (Litovsky 2011; Kühnle et al. 2013; Freigang et al. 2015; Litovsky 2015).

For BCI children, clinical factors are also likely to contribute to explaining performance and interindividual variability. We found a correlation between interimplant interval and 3D error under the HI condition, which remained significant after controlling for chronological age at test. Since 2012, bilateral cochlear implantation is standard practice in France (Simon et al. 2019), thus only the youngest children in our cohort benefited from a sequential cochlear implantation with a short delay. This observation is in keeping with a recent study (Killan et al. 2019) which showed that a longer interimplant delay was a poor prognostic for spatial hearing abilities, mainly due to asymmetric brain processing after a 24-month interval (Gordon et al. 2008, 2013; Kral et al. 2019). This could explain why we expected (based on, e.g., Zheng et al. 2015; Killan et al. 2019) but did not observe a significant improvement in spatial hearing abilities with increased binaural experience. It is possible that this correlation is more sensitive in the first 2 years after cochlear implantation, when localization performance improves rapidly and parallels experience-induced plasticity (Kral & Tillein 2006; Gordon et al. 2011). In sum, variability of BCI performance is more complex than expected. Future studies might leverage longitudinal approaches to gain deeper insights into the roles of these and other individual differences.

Perspectives

Our new approach for measuring spatial hearing in the near-field has been validated in a virtual reality platform easy to access in clinical practice (HTC VIVE; Verdelet et al. 2019; Valzolgher et al. 2020a,b). These first results give new information about children's performance, both when they have typical hearing, and when they are deafness but fitted with BCI. However, to approximate even further the children's everyday environments, it would be useful to add complexity to our

sound localization task, with background noise or changes in sound elevation. This could provide further insights into the understanding of how BCI children extract dynamic monaural cues. In addition, it could help to clarify our preliminary results on the progressive and non-linear maturation of spatial hearing in each space dimension for NH children. We hypothesize that sound discrimination in distance could mature later than azimuthal judgement. Owing to the different auditory cues at play to render distance perception in the far compared to the near-field, it will be important to extend research to the far-field to evaluate whether and to what extent spatial hearing maturation in distance differs from the near-field. Finally, spontaneous head movements significantly improved sound localization performance in BCI children. While it is still unclear to what extent BCI children use this valuable strategy in their daily environment, it seems essential to pay more attention to this orienting behavior in future studies of clinical assessment or spatial hearing rehabilitation. For example, in line with speech therapy rehabilitation, it could be helpful to train the most severely impaired BCI users to localize various sounds in a controlled environment.

CONCLUSION

BCI children and adolescents display important spatial localization deficits compared to their NH peers. These deficits hamper sound localization in all spatial dimensions, but go undetected on self-report questionnaires. All BCI children had some localization skills (i.e., left/right discrimination) but in the context of other spatial hearing difficulties related to front-back confusions and distance perception, which likely resulted from the reduction in auditory cues resulting from the implant settings (e.g., omnidirectional mode) and sound adjustments. Notably, BCI children improved under conditions of free head exploration during sound emission. This suggests that head movements could represent a rehabilitation entry strategy to help BCI users when faced with complex auditory scenes in daily life.

ACKNOWLEDGMENTS

All the authors significantly contributed to this work. A.C., V.G., A.F., and F.P. designed the experiment, A.C. and V.G. performed experiments, A.C., V.G., and F.P. analyzed the data and wrote the paper. E.T. provided cochlear implant participants, while J.G. oversaw participants' inclusions and help in the ethical procedures required for the study. G.V. and R.S. implemented the experiments. All authors discussed the results, reviewed the manuscript and approved the final version of the revised manuscript.

V.G., R.S., A.F., and F.P. filed a patenting procedure for the system reported in this study, patent pending. The other authors have no conflicts of interest to disclose.

The authors are grateful to all the children who took part in this study and their families. We thank students of the Institute of Rehabilitation Sciences and Technology (ISTR, Lyon) who helped for data analysis, and Karen Reilly for her precious English proofreading of the manuscript.

A.C. was supported by a grant of the Hospices Civils de Lyon; J.G., G.V., F.P., V.G., E.T., and A.F. were supported by a grant of the Agence Nationale de la Recherche (ANR-16-CE17-0016, VIRTUALHEARING3D, France). F.P. and A.F. were supported by a prize of the Fondation Medisite (France) and the Fondation Neurodis (France). The study was supported by the IHU CeSaMe ANR-10-IBHU-0003, and it was performed within the framework of the LABEX CORTEX (ANR-11-LABX-0042) of Université de Lyon. We thank the administrative staff of the IMPACT team for their administrative and informatics support.

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

Received April 25, 2020; accepted May 16, 2021; published online ahead of print July 1, 2021.

REFERENCES

- Archer-Boyd, A. W., & Carlyon, R. P. (2019). Simulations of the effect of unlinked cochlear-implant automatic gain control and head movement on interaural level differences. *J Acoust Soc Am*, *145*, 1389.
- Brimijoin, W. O., McShefferty, D., Akeroyd, M. A. (2010). Auditory and visual orienting responses in listeners with and without hearing-impairment. *J Acoust Soc Am*, *127*, 3678–3688.
- Brimijoin, W. O., McShefferty, D., Akeroyd, M. A. (2012). Undirected head movements of listeners with asymmetrical hearing impairment during a speech-in-noise task. *Hear Res*, *283*, 162–168.
- Brungart, D. S. (1999). Auditory localization of nearby sources. III. Stimulus effects. *J Acoust Soc Am*, *106*, 3589–3602.
- Brungart, D. S., & Rabinowitz, W. M. (1999). Auditory localization of nearby sources. Head-related transfer functions. *J Acoust Soc Am*, *106*(3 Pt 1), 1465–1479.
- Brungart, D. S., Durlach, N. I., Rabinowitz, W. M. (1999). Auditory localization of nearby sources. II. Localization of a broadband source. *J Acoust Soc Am*, *106*, 1956–1968.
- Calcagno, E. R., Abregú, E. L., Eguía, M. C., Vergara, R. (2012). The role of vision in auditory distance perception. *Perception*, *41*, 175–192.
- Choi, J. E., Moon, I. J., Kim, E. Y., Park, H. S., Kim, B. K., Chung, W. H., Cho, Y. S., Brown, C. J., Hong, S. H. (2017). Sound localization and speech perception in noise of pediatric cochlear implant recipients: Bimodal fitting versus bilateral cochlear implants. *Ear Hear*, *38*, 426–440.
- Choi, J. E., Hong, S. H., Moon, I. J. (2020). Academic performance, communication, and psychosocial development of prelingual deaf children with cochlear implants in mainstream schools. *J Audiol Otol*, *24*, 61–70.
- Da Silva, J. A. (1985). Scales for perceived egocentric distance in a large open field: Comparison of three psychophysical methods. *Am J Psychol*, *98*, 119–144.
- Dillon, H. (2001). *Hearing Aids*. Boomerang press.
- Dorman, M. F., Loisele, L., Stohl, J., Yost, W. A., Spahr, A., Brown, C., Cook, S. (2014). Interaural level differences and sound source localization for bilateral cochlear implant patients. *Ear Hear*, *35*, 633–640.
- Fischer, T., Schmid, C., Kompis, M., Mantokoudis, G., Caversaccio, M., Wimmer, W. (2020). Pinna-imitating microphone directionality improves sound localization and discrimination in bilateral cochlear implant users. *Ear Hear*, *42*, 214–222.
- Freigang, C., Richter, N., Rübsamen, R., Ludwig, A. A. (2015). Age-related changes in sound localisation ability. *Cell Tissue Res*, *361*, 371–386.
- Galvin, K. L., & Noble, W. (2013). Adaptation of the speech, spatial, and qualities of hearing scale for use with children, parents, and teachers. *Cochlear Implants Int*, *14*, 135–141.
- Galvin, K. L., & Mok, M. (2016). Everyday listening performance of children before and after receiving a second cochlear implant: Results using the parent version of the speech, spatial, and qualities of hearing scale. *Ear Hear*, *37*, 93–102.
- Gordon, K. A., Valero, J., van Hoesel, R., Papsin, B. C. (2008). Abnormal timing delays in auditory brainstem responses evoked by bilateral cochlear implant use in children. *Otol Neurotol*, *29*, 193–198.
- Gordon, K. A., & Papsin, B. C. (2009). Benefits of short interimplant delays in children receiving bilateral cochlear implants. *Otol Neurotol*, *30*, 319–331.
- Gordon, K. A., Jiواني, S., Papsin, B. C. (2011). What is the optimal timing for bilateral cochlear implantation in children? *Cochlear Implants Int*, *12*(Suppl 2), S8–14.
- Gordon, K. A., Wong, D. D. E., Papsin, B. C. (2013). Bilateral input protects the cortex from unilaterally-driven reorganization in children who are deaf. *Brain*, *136*(Pt 5), 1609–1625.
- Gordon, K. A., Cushing, S. L., Easwar, V., Polonenko, M. J., Papsin, B. C. (2017). Binaural integration: A challenge to overcome for children with hearing loss. *Curr Opin Otolaryngol Head Neck Surg*, *25*, 514–519.
- Grantham, D. W., Ricketts, T. A., Ashmead, D. H., Labadie, R. F., Haynes, D. S. (2008). Localization by postlingually deafened adults fitted with a single cochlear implant. *Laryngoscope*, *118*, 145–151.
- Grieco-Calub, T. M., & Litovsky, R. Y. (2010). Sound localization skills in children who use bilateral cochlear implants and in children with normal acoustic hearing. *Ear Hear*, *31*, 645–656.
- Kerber, S., & Seeber, B. U. (2012). Sound localization in noise by normal-hearing listeners and cochlear implant users. *Ear Hear*, *33*, 445–457.
- Khing, P. P., Swanson, B. A., Ambikairajah, E. (2013). The effect of automatic gain control structure and release time on cochlear implant speech intelligibility. *PLoS One*, *8*, e82263.
- Killan, C. F., Harman, S., Killan, E. C. (2018). Changes in sound-source localization for children with bilateral severe to profound hearing loss following simultaneous bilateral cochlear implantation. *Cochlear Implants Int*, *19*, 284–291.
- Killan, C., Scally, A., Killan, E., Totten, C., Raine, C. (2019). Factors affecting sound-source localization in children with simultaneous or sequential bilateral cochlear implants. *Ear Hear*, *40*, 870–877.
- Kim, C., Mason, R., Brookes, T. (2013). Head movements made by listeners in experimental and real-life listening activities. *J Audio Eng Soc*, *61*, 425–438.
- King, A. J. (2009). Visual influences on auditory spatial learning. *Philos Trans R Soc Lond B Biol Sci*, *364*, 331–339.
- Kolarik, A. J., Moore, B. C., Zahorik, P., Cirstea, S., Pardhan, S. (2016). Auditory distance perception in humans: A review of cues, development, neuronal bases, and effects of sensory loss. *Atten Percept Psychophys*, *78*, 373–395.
- Kopčo, N., & Shinn-Cunningham, B. G. (2011). Effect of stimulus spectrum on distance perception for nearby sources. *J Acoust Soc Am*, *130*, 1530–1541.
- Kral, A., & Tillein, J. (2006). Brain plasticity under cochlear implant stimulation. *Adv Otorhinolaryngol*, *64*, 89–108.
- Kral, A., Dorman, M. F., Wilson, B. S. (2019). Neuronal development of hearing and language: Cochlear implants and critical periods. *Annu Rev Neurosci*, *42*, 47–65.
- Kühnle, S., Ludwig, A. A., Meuret, S., Küttner, C., Witte, C., Scholbach, J., Fuchs, M., Rübsamen, R. (2013). Development of auditory localization accuracy and auditory spatial discrimination in children and adolescents. *Audiol Neurootol*, *18*, 48–62.
- Kumpik, D. P., Campbell, C., Schnupp, J. W. H., King, A. J. (2019). Re-weighting of sound localization cues by audiovisual training. *Front Neurosci*, *13*, 1164.
- Litovsky, R. Y., Parkinson, A., Arcaroli, J., Peters, R., Lake, J., Johnstone, P., Yu, G. (2004). Bilateral cochlear implants in adults and children. *Arch Otolaryngol Head Neck Surg*, *130*, 648–655.
- Litovsky, R. Y. (2011). Review of recent work on spatial hearing skills in children with bilateral cochlear implants. *Cochlear Implants Int*, *12*(Suppl 1), S30–S34.
- Litovsky, R. (2015). Development of the auditory system. *Handb Clin Neurol*, *129*, 55–72.
- Loomis, J. M., Klatzky, R. L., Philbeck, J. W., Golledge, R. G. (1998). Assessing auditory distance perception using perceptually directed action. *Percept Psychophys*, *60*, 966–980.
- Middlebrooks, J. C., & Green, D. M. (1991). Sound localization by human listeners. *Annu Rev Psychol*, *42*, 135–159.
- Middlebrooks, J. C. (2015). Sound localization. *Handb Clin Neurol*, *129*, 99–116.
- Misurelli, S. M., & Litovsky, R. Y. (2012). Spatial release from masking in children with normal hearing and with bilateral cochlear implants: Effect of interferer asymmetry. *J Acoust Soc Am*, *132*, 380–391.
- Mueller, M. F., Meisenbacher, K., Lai, W. K., Dillier, N. (2014). Sound localization with bilateral cochlear implants in noise: How much do head movements contribute to localization? *Cochlear Implants Int*, *15*, 36–42.
- Pastore, M. T., Natale, S. J., Yost, W. A., Dorman, M. F. (2018). Head movements allow listeners bilaterally implanted with cochlear implants to resolve front-back confusions. *Ear Hear*, *39*, 1224–1231.
- Pavani, F., 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. *Hear Res*, *344*, 24–37.
- Perrett, S., & Noble, W. (1997). The effect of head rotations on vertical plane sound localization. *J Acoust Soc Am*, *102*, 2325–2332.
- Platt, B. B., & Warren, D. H. (1972). Auditory localization: The importance of eye movements and a textured visual environment. *Percept Psychophys*, *12*, 245–248.

- Rakerd, B., & Hartmann, W. M. (1986). Localization of sound in rooms, III: Onset and duration effects. *J Acoust Soc Am*, *80*, 1695–1706.
- Seeber, B. U., Baumann, U., Fastl, H. (2004). Localization ability with bimodal hearing aids and bilateral cochlear implants. *J Acoust Soc Am*, *116*, 1698–1709.
- Shelton, B. R., & Searle, C. L. (1980). The influence of vision on the absolute identification of sound-source position. *Percept Psychophys*, *28*, 589–596.
- Shinn-Cunningham, B. G., Kopco, N., Martin, T. J. (2005). Localizing nearby sound sources in a classroom: Binaural room impulse responses. *J Acoust Soc Am*, *117*, 3100–3115.
- Shinn-Cunningham, B., Best, V., Lee, A. K. (2017). Auditory object formation and selection. In J. C. Middlebrooks, J. Z., Simon, A. N. Popper, R.R. Fay (Eds.), *The Auditory System at the Cocktail Party* (pp. 7–40). Springer.
- Simon, F., Roman, S., Truy, E., Barone, P., Belmin, J., Blanchet, C., Borel, S., Charpiot, A., Coez, A., Deguine, O., Farinetti, A., Godey, B., Lazard, D., Marx, M., Mosnier, I., Nguyen, Y., Teissier, N., Virole, B., Lescanne, E., Loundon, N. (2019). Guidelines (short version) of the French Society of Otorhinolaryngology (SFORL) on pediatric cochlear implant indications. *Eur Ann Otorhinolaryngol Head Neck Dis*, *136*, 385–391.
- Stöbich, B., Zierhofer, C. M., Hochmair, E. S. (1999). Influence of automatic gain control parameter settings on speech understanding of cochlear implant users employing the continuous interleaved sampling strategy. *Ear Hear*, *20*, 104–116.
- Tabry, V., Zatorre, R. J., Voss, P. (2013). The influence of vision on sound localization abilities in both the horizontal and vertical planes. *Front Psychol*, *4*, 932.
- Valzolgher, C., Verdelet, G., Salemme, R., Lombardi, L., Gaveau, V., Farné, A., Pavani, F. (2020a). Reaching to sounds in virtual reality: A multisensory-motor approach to promote adaptation to altered auditory cues. *Neuropsychologia*, *149*, 107665.
- Valzolgher, C., Alzhaler, M., Gessa, E., Todeschini, M., Nieto, P., Verdelet, G., Salemme, R., Gaveau, V., Marx, M., Truy, E., Barone, P., Farné, A., Pavani, F. (2020b). The impact of a visual spatial frame on real sound-source localization in virtual reality. *Curr Res Behav Sci*, *1*, 100003.
- Van Deun, L., van Wieringen, A., Scherf, F., Deggouj, N., Desloovere, C., Offeciers, F. E., Van de Heyning, P. H., Dhooge, I. J., Wouters, J. (2010). Earlier intervention leads to better sound localization in children with bilateral cochlear implants. *Audiol Neurootol*, *15*, 7–17.
- Van Hoesel, R. (2004). Exploring the benefits of bilateral cochlear implants. *Audiol Neurootol*, *9*, 234–246.
- van Wieringen, A., & Wouters, J. (2015). What can we expect of normally-developing children implanted at a young age with respect to their auditory, linguistic and cognitive skills? *Hear Res*, *322*, 171–179.
- Verdelet, G., Desoche, C., Volland, F., Farné, A., Coudert, A., Hermann, R., Truy, E., Gaveau, V., Pavani, F., Salemme, R. (2019). Assessing spatial and temporal reliability of the vive system as a tool for naturalistic behavioural research. *2019 International Conference on 3D Immersion (IC3D)*, 1–8. <https://doi.org/10.1109/IC3D48390.2019.8975994>
- Wallach, H. (1940). The role of head movements and vestibular and visual cues in sound localization. *J Experimental Psychol*, *27*: 339–68
- Wightman, F. L., & Kistler, D. J. (1999). Resolution of front-back ambiguity in spatial hearing by listener and source movement. *J Acoust Soc Am*, *105*, 2841–2853.
- Zahorik, P. (2001). Estimating sound source distance with and without vision. *Optom Vis Sci*, *78*, 270–275.
- Zheng, Y., Godar, S. P., Litovsky, R. Y. (2015). Development of sound localization strategies in children with bilateral cochlear implants. *PLoS One*, *10*, e0135790.

Journal Pre-proof

Reaching to sounds in virtual reality: A multisensory-motor approach to promote adaptation to altered auditory cues

Chiara Valzolgher, Grégoire Verdelet, Romeo Salemme, Luigi Lombardi, Valerie Gaveau, Alessandro Farné, Francesco Pavani

PII: S0028-3932(20)30337-7

DOI: <https://doi.org/10.1016/j.neuropsychologia.2020.107665>

Reference: NSY 107665

To appear in: *Neuropsychologia*

Received Date: 30 January 2020

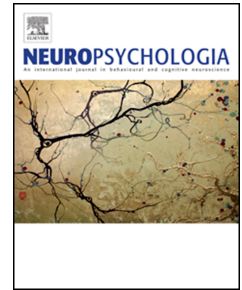
Revised Date: 25 July 2020

Accepted Date: 24 October 2020

Please cite this article as: Valzolgher, C., Verdelet, G., Salemme, R., Lombardi, L., Gaveau, V., Farné, A., Pavani, F., Reaching to sounds in virtual reality: A multisensory-motor approach to promote adaptation to altered auditory cues, *Neuropsychologia*, <https://doi.org/10.1016/j.neuropsychologia.2020.107665>.

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2020 Elsevier Ltd. All rights reserved.



CRedit author statement

Chiara Valzolgher and Francesco Pavani: *Conceptualization, Methodology, Writing - Review & Editing, Writing - Original Draft*

Chiara Valzolgher: *Investigation, Data Curation, Visualisation*

Chiara Valzolgher, Luigi Lombardi and Francesco Pavani: *Formal analysis, Data Curation*

Alessandro Farnè and Francesco Pavani: *Funding acquisition, Supervision*

Grègoire Verdelet and Romeo Salemme: *Software*

Valerie Gaveau: *Validation, Resources*

Journal Pre-proof

Reaching to sounds in virtual reality: A multisensory-motor approach to promote adaptation to altered auditory cues

Chiara Valzolgher^{1,2}, Grégoire Verdelet¹, Romeo Salemmè^{1,4}, Luigi Lombardi³, Valerie Gaveau¹,
Alessandro Farné^{1,2,4} & Francesco Pavani^{1,2,3}

¹ IMPACT, Centre de Recherche en Neurosciences Lyon (CRNL), France

² Centre for Mind/Brain Sciences (CIMeC), University of Trento, Italy

³ Department of Psychology and Cognitive Sciences (DiPSCo), University of Trento, Italy

⁴ Neuro-immersion, Centre de Recherche en Neurosciences Lyon (CRNL), France

Corresponding Author:

Chiara Valzolgher

INSERM U1028 CNRS UMR5292 — Claude Bernard University, Lyon 1 — Lyon Hospitals
16, ave Doyen Lépine
69676 Bron Cedex, France

Phone: +33 (0)4 72 91 34 07

chiara.valzolgher@inserm.fr

ABSTRACT

When localising sounds in space the brain relies on internal models that specify the correspondence between the auditory input reaching the ears, initial head-position and coordinates in external space. These models can be updated throughout life, setting the basis for re-learning spatial hearing abilities in adulthood. In addition, strategic behavioural adjustments allow people to quickly adapt to atypical listening situations. Until recently, the potential role of dynamic listening, involving head-movements or reaching to sounds, have remained largely overlooked. Here, we exploited visual virtual reality (VR) and real-time kinematic tracking, to study the role of active multisensory-motor interactions when hearing individuals adapt to altered binaural cues (one ear plugged and muffed). Participants were immersed in a VR scenario showing 17 virtual speakers at ear-level. In each trial, they heard a sound delivered from a real speaker aligned with one of the virtual ones and were instructed to either reach-to-touch the perceived sound source (Reaching group), or read the label associated with the speaker (Naming group). Participants were free to move their heads during the task and received audio-visual feedback on their performance. Most importantly, they performed the task under binaural or monaural listening. Results show that both groups adapt rapidly to monaural listening, improving sound localisation performance across trials and changing their head-movement behaviour. Reaching the sounds induced faster and larger sound localisation improvements, compared to just naming its position. This benefit was linked to progressively wider head-movements to explore auditory space, selectively in the Reaching group. In conclusion, reaching to sounds in an immersive visual VR context proved most effective for adapting to altered binaural listening. Head-movements played an important role in adaptation, pointing to the importance of dynamic listening when implementing training protocols for improving spatial hearing.

KEYWORDS

Dynamic listening, sound localisation, binaural cues, reaching to sounds, head movements.

HIGHLIGHTS

- We studied adaptation to monaural listening when listeners interact with sounds
- Reaching to sounds reduced errors more than just naming the sound sources
- Patterns of spontaneous head-movements changed during adaptation
- Exploration of a wider portion of space with the head was functional for adaptation
- Reaching to sounds promoted a more effective head-related behaviour

Journal Pre-proof

1. INTRODUCTION

Spatial hearing is a remarkable ability of the brain. To determine the spatial coordinates of sounds in the environment, the cognitive system exploits auditory cues resulting from the interactions between sound waves, the head and outer ears (Middlebrooks, 2015; Middlebrooks & Green, 1991; Wallach, 1940). Auditory cues are distinguished in binaural cues, which involve a combined computation of the signals reaching the two ears and are used for sound localisation in the horizontal dimension (Rayleigh, 1907), and monaural cues which involve the information available at each single ear and are crucial to localise sound in elevation, depth and anterior-posterior axis (Angell & Fite, 1901). In normal hearing listening condition, both monaural and binaural cues are weighted and combined to achieve optimal localisation of auditory events in the 3D space (Carlile, Martin & McAnally, 2005; see also Shinn-Cunningham, Santarelli, & Kopco, 2000). Auditory cues change during the life span, due anatomical modifications of one's own body during development (Clifton, Gwiazda, Bauer, Clarkson & Held, 1998) or changes in hearing threshold with ageing (Cranford, Andres, Piatz & Reissig, 1993; Dobрева, O'Neill & Paige, 2010). In addition, they change on a daily basis, due to the diversity of listening contexts to which we are all exposed (Majdak, Goupell & Laback, 2010). It is now clear that the brain remains capable to cope with these continuous changes in auditory cues throughout life (e.g., Carlile, Balachandar & Kelly 2014; Keating & King, 2015; Rabini, Altobelli & Pavani, 2019; Strelnikov, Rosito & Barone, 2011; Valzolgher, Campus, Rabini, Gori & Pavani, 2020; Van Wanrooij, John & Opstal, 2005). In the present work, we examined reaching to sounds in virtual reality as a multisensory-motor strategy for adapting sound localisation abilities in adulthood.

Updating of sound-space correspondences typically occurs in a multisensory environment. Whenever the other sensory systems (e.g., vision) provide reliable spatial information about auditory events, the brain exploits these additional sensory sources to calibrate and optimize internal models for spatial hearing (Keating & King, 2015). This notion emerged from studies that examined re-learning of sound-space correspondences when binaural auditory cues were temporarily altered using monaural ear-plugs (Rabini et al., 2019; Strelnikov et al., 2011; Trapeau & Schönwiesner, 2015; Valzolgher et al., 2020;), or monaural auditory cues were modified through ear molds (Van Wanrooij & Van Opstal, 2005) or non-individualised HRTFs (Head-Related Transfer Functions; Hendrickx, Stitt, Messonier, Lyzwa, Katz & De Boishéraud, 2017; Honda, Shibata, Gyoba, Saitou, Iwaya & Suzuki, 2007; Parseihian, Jouffrais & Katz, 2014; Parseihian & Katz, 2012; Steadman, Kim, Lestang, Goodman & Picinali, 2019; Stitt, Picinali & Katz, 2019). In these

simulated altered-listening conditions, multisensory training procedures proved effective for adapting to novel auditory cues (for reviews see: Carlile, 2014; Keating & King, 2015; Knudsen & Knudsen, 1985; Mendonça, 2014; Irving & Moore, 2011). For instance, Strelnikov and colleagues (2011) studied the effects of audio-visual vs. auditory-only training on sound localisation in monaurally-plugged adults. They found that performance improvements were larger after a training that exploited spatially and temporally congruent audio-visual inputs, compared to a training based on auditory information alone. These findings complement other research in which participants received feedback after their sound localisation response, through audio-visual (Shinn-Cunningham, Durlach, & Held, 1998; Zahorik, Bangayan, Sundareswaran, Wang & Tam, 2006; see also Majdak, Walder & Laback, 2013) or visual cues (Bauer, Matuzsa, Blackmer & Glucksberg, 1966; Kumpik, Kacelnik & King, 2010; Mendonça, Campos, Dias & Santos, 2013). Finally, converging evidence in favour of multisensory-based training emerged from research in ferrets with unilateral and bilateral cochlear implants. Performance in early-deafened ferrets with bilateral cochlear implants improved consistently after a multisensory training that used interleaved auditory and visual stimuli (Isaiah, Vongpaisal, King, & Hartley, 2014).

In natural, off-laboratory conditions, spatial hearing is a dynamic and sensorimotor task. Listeners spontaneously orient their head, eyes and trunk to the auditory source (e.g., when turning to listen to a speaker behind us), or they move their body in the environment to approach the auditory source (e.g., when reaching to grasp the mobile phone ringing on the table). These dynamic and sensorimotor interactions with sound sources are almost ubiquitous in our everyday listening conditions. Head-orienting to sounds is a spontaneous behaviour: it is detected in infants 2-4 days after birth (Muir & Field, 1979) and it follows the shortest path to the sound already between 6 and 9 months of age (van der Meer, Ramstad, & van der Weel, 2008). Although adults have different propensities for head-movement (Fuller, 1992), they can easily orient their head to sounds when required by the task (Brimijoin, McShefferty, & Akeroyd, 2010). Experimental evidence shows that head-movements during listening improve sound localisation (Noble, 1981; Thurlow and Runge, 1967), and head-movement strategies change in altered hearing conditions in the attempt of coping with the hearing difficulty. For instance, Brimijoin and colleagues (2010) showed that people with hearing-impairment have more complex head-movements. Also, people with asymmetrical hearing impairment use head-orienting to maximize the level of a target sentence in their better ear when solving a speech-in-noise task (Brimijoin & Akeroyd, 2012). Yet, studies that examined adaptation to altered auditory cues have typically limited the possibilities for head-movements using a chinrest (Kumpik et al., 2010;

Strelnikov et al., 2011) or by instructing participants to refrain from head-movements during listening (Rabini et al., 2019; Van Wanrooij & Van Opstal, 2005; Zahorik et al., 2006).

However, notable exceptions to the static listening approach have emerged in the last decade. Using gamified scenarios, researchers have asked participants to use a hand-held tool to hit a moving sound presented in virtual auditory space (Honda et al., 2007; Honda et al., 2013; Ohunchi, Iwaya, Suzuki & Minekata, 2005), or to shoot audio-visual moving targets presented in virtual reality (Poirier-Quinot & Katz, 2020). In all these cases participants were free to move their heads and bodies. In more controlled conditions, Parseihian and Katz (2012) studied how people could adapt to non-individualized head-related transfer functions (HRTFs) when localizing virtual sound sources. They explored the effects of a multi-modal training platform in which participants were involved in an active game-like scenario. They actively searched for animal sounds, scanning the space around them (front and back), using a hand-held track-ball. Again, participants were free to move their heads and hands. This training improved vertical localisation performance and reduced front/back confusions, showing that people could rapidly adapt to novel monaural cues. Using an identical training approach, but with longer sessions and with perceptually worst-rated non-individualised HRTFs, Stitt and colleagues (2019) also found evidence for some degree of adaptation. Another recent study (Steadman et al., 2019) combined gamification, virtual reality and free head-movements in a training protocol aimed at ameliorating sound localisation with virtual sounds. These authors also trained participants to adapt to non-individualized HRTFs, hence to new monaural cues. They asked participants to listen to stationary virtual sounds, either with the head static or free to move. Compared to the static listening condition, when the head was free to move sound localisation improved even over a very short timescale, corroborating the importance of head movements when adapting to novel auditory cues.

Despite these studies provide initial evidence in support of a key role of dynamic and sensorimotor interactions with sounds when adapting to altered auditory cues, they did not investigate the qualitative and quantitative nature of the spontaneous head-movements during listening (e.g., head-movement number or amplitude, head-movement strategies). As such, they only provide limited evidence of the role of head-movements in adapting to novel listening conditions. Note that in the head-constrained context, researchers can only examine the cognitive mechanisms of adaptation, such as cue-remapping (i.e., changes in the internal models that specify the correspondence between auditory cues and spatial locations) or cue-reweighting (i.e., changes in the contributions each auditory cue in these internal models). By contrast, the conditions in which the head is free to move are informative about the strategies (implicit or explicit) that

participants use to cope with the altered listening conditions. In addition, although some of the previous literature point to a joint contribution of head-movements and reaching movements in improving sound localization with altered auditory cues (Honda et al., 2007; Honda et al., 2013; Ohunchi, Iwaya, Suzuki & Minekata, 2005; Parseihian & Katz, 2012), several questions remain open. Specifically, are reaching movements relevant in promoting the useful strategies for adapting to altered spatial hearing conditions? Could they trigger different head-movements during sounds compared to alternative and less-interacting responses (e.g., naming labels that identify the speaker position)? Finally, the vast majority of studies that examined the effects of dynamic and sensorimotor interactions with sounds when adapting to novel auditory cues focused on monaural cues. Thus, it remains an open question to what extent head-movements can also allow fast adaptation to altered listening when binaural rather than monaural cues are changed.

In the present study we addressed all these questions by asking two groups of participants to perform a sound localisation task in an interactive visual virtual reality scenario, with real sounds presented within reaching distance. Participants wore a Head-Mounted Display (HMD) and perceived themselves in the centre of an empty virtual room with visible virtual speakers. In each trial, a 3 seconds sound was emitted from a real speaker aligned with one of the visible virtual speakers. All participants were asked to identify from which virtual speaker the sound originated and were informed that they were free to move their head during sound emission to solve the task. Crucially, half of participants were instructed to stop sound emission by reaching to the correct source using a hand-held controller (Reaching group). If the participant reached to the correct speaker the sound stopped, promoting a sense of agency over the auditory change. If instead the participant reached towards the wrong speaker the sound continued playing and the correct speaker started flashing, providing multisensory cues to actual sound location until the participants gave the correct response (see Figure 1B). The other half of participants received identical stimulation and identical feedback but were instructed to stop sound emission by naming the number that identified the active speaker (Naming group).

To examine the role of these two response instructions when adapting to altered spatial hearing conditions we focused specifically on binaural cues alteration, using plugging and muffing of the right ear. Unlike previous related works (Honda et al., 2007; Honda et al., 2013; Ohunchi, Iwaya, Suzuki & Minekata, 2005; Parseihian & Katz, 2012; Stitt, Picinali & Katz, 2019), in our study monaural cues at the unplugged left ear remained unaltered. Instead, we aimed to simulate conductive hearing loss. In terms of auditory cues, our manipulation clearly affected inter-aural level differences (ILD). In addition, because we adopted a slowly

changing broad-band sound we made inter-aural time differences (ITD) largely ineffective (see Methods section).

If reaching to sounds play a role in adaptation to altered binaural cues, performance in the monaural listening condition should improve faster for the Reaching compared to the Naming group. The rationale for this hypothesis is linked to the fact that reaching to sounds requires a coordination of different effectors (eyes, head, hand) into a common reference frame (Cohen & Andersen, 2002). In turn, this may result in a more stable (or salient) spatial coding of sound source location and favour adaptation to altered spatial hearing conditions. In addition, reaching movement may help to direct attention towards the position occupied by the sound source, because of additional visual, proprioceptive and kinaesthetic cues resulting from the action.

2. MATERIAL AND METHODS

2.1. Participants

Twenty-eight participants (age: $M = 20.93$, $SD = 2.48$, range [18-30], 5 males, 14 right-handed in the Reaching Group, 13 right-handed in the Naming Group) were recruited to participate in the experiment, mostly among undergraduate students at the University of Trento. All participants signed an informed consent before starting the experiment, which was conducted according to the criteria of the Declaration of Helsinki (1964, amended in 2013) and approved by the ethical committee at the University of Trento (protocol: 2018-018). All had normal or corrected-to-normal vision, and reported no movement deficit. Hearing threshold was measured using an audiometer (Grason Stadler GSI 17 Audiometer) for all participants, testing different frequencies (250, 500, 1000, 2000, 4000 Hz), on the right and left ear separately. All participants had an average threshold below 11.7 dB HL.

2.2. Apparatus and stimuli

Visual virtual reality (VR) and kinematic tracking was implemented using an HTC Vive System, comprising one head-mounted display (HMD, resolution: 1080 x 1200 px, Field Of View (FOV): 110°, Refresh rate : 90 Hz), 1 controller (which was used by participants to interact with sounds), 1 tracker (which was placed above the speaker to track its position in real time) and 2 lighthouse base stations (which served for regular scanning of the position of the controller and the tracker). Tracking precision and accuracy of the HTC Vive System is adequate for behavioural research purposes, as recently measured by our research group (Verdelet et al., 2019; see also Gaveau et al., 2020). Specifically, the HTC Vive has submillimetre precision (0.237 mm) and near centimetre accuracy (9.0 mm when trackers are static; 9.4 mm when trackers are dynamic). The HMD was equipped with an SMI eye-tracking system (250 Hz). All stimuli were controlled and delivered using a LDLC ZALMAN PC (OS: Windows 10 (10.0.0) 64bit; Graphic card: NVIDIA GeForce GTX 1060 6GB; Processor: Intel Core i7-7700K, Quad-Core 4.2 GHz/4.5 GHz Turbo - Cache 8 Mo - TDP 95W) using Steam VR software and the development platform Unity.

The experiment was entirely run in a 4x3 meters ordinary room almost devoid of furniture, not treated for being anechoic and quiet. The subject sat in the centre of the room, with no constraints for her head (no chin-rest was used). The virtual environment in which participants were immersed was an ordinary squared room (grey wall), similar to the real experimental room. The room was empty and there were not objects on

the walls, except a door which was placed in the back wall. In front of participants 17 visible speakers were positioned at ear level. In each session, subjects wearing the HDM saw themselves in the virtual scenario, sitting 50 cm from the centre of a semicircle of 17 speakers, spanning about $\pm 80^\circ$ of visual angle. Each speaker was marked with a numeric label (from 1 to 17) that changed in each trial. Auditory stimuli were delivered by a real, unseen speaker (JBL GO Black), which was moved by the experimenter and whose position was tracked in space. The auditory stimulus was an amplitude-modulated (4 Hz) white noise bursts (about 60 dB SPL as measured by a Decibel Meter TES1350A placed at ears level).

This solution allowed us the possibility of tracking the position of the tracker (i.e., the real speaker), the controller (handled by the participant) and the Head Mounted Display, via sensors on the headset and VR controllers using IR led positioned on the base stations (frequency sample 250 Hz). We design the software to allow the experimenter to track and align the real loudspeaker (the sound source) with the desired position in the virtual environment (i.e., in correspondence of one of the visible loudspeakers). Specifically, in each experimental trial, the experimenter moved the loudspeaker to the desired position in 3D space, following visual instructions concerning the pre-determined speaker position. These instructions were visible only to the experimenter, and they were delivered using the stimulus visualization monitor placed in the testing room. The monitor provided a bird-eye view of the experimental room and conveyed the pre-determined position of the sound source for that trial and the actual real-time position of the speaker. Using these visual instructions, the experimenter reached for computer-determined position rapidly, keeping the speaker approximately at ear-level. The spanning of about $\pm 80^\circ$ permitted to the experimenter to stand still in front of the participant and move the speaker silently stretching her right or left arms. Although the noise produced by the experimenter were minimal, participants were explicitly informed to pay attention to the target sounds, as any other sound in the room could in fact be deceiving about target sound position. Before starting, the experimenter showed how she can stay to the right of participants while providing sound from the left of participants by stretching her arm of the hand with the speaker. This was a demonstration that background noise could provide erroneous cues and thus discourage their use by participants, which were invited to focus to the specific target sound. Most importantly, pilot work in our laboratory showed that participants cannot reliably point to the speaker if the same placement procedure is used but no target sound is delivered (Gaveau et al., 2020).

To simulate monaural hearing experience, the right ear of each participant was occluded using an ear-plug (3M PP 01 002; attenuation value for high frequencies = 30 dB SPL; attenuation value for medium

frequencies = 24 dB SPL; attenuation value for low frequencies = 22 dB SPL; Single Number Rating = 32 dB SPL; as reported by the manufacturer). In addition, a unilateral ear muff covered the right ear (3M 1445 modified to cover only the right ear; attenuation value for high frequencies = 32 dB SPL; attenuation value for medium frequencies = 29 dB SPL; attenuation value for low frequencies = 23 dB SPL; Single Number Rating = 28 dB SPL; as reported by the manufacturer). The combined use of an ear-plug and muff at the right ear clearly affected ILDs. In addition, by attenuating the intensity of our amplitude-modulated broadband signal to one ear we most likely reduced the available ITD cues. This because we reduced the possibility of comparing the fine structure at low frequency, as well as the amplitude modulation signals that reach the two ears. Thus, we altered both binaural cues while leaving intact the possibility of exploiting monaural cues at the unplugged left ear.

2.3. Procedure

We measured auditory performance during five different blocks, under two different listening conditions: binaural listening (B) and simulated monaural listening (M) in which the left ear was plugged (Figure 1A). In each block, participants performed 51 trials (3 trials for each source) of a sound localization task. First, participants were invited to sit down on a chair and wear the Head Mounted Display (HMD) and to perform an eye-tracker calibration (they had to follow a white dot moving on the screen with eyes). Then, they were immersed in the virtual scenario. Around them, the cylindrical speakers were located at the height of their ears. Participants were instructed that the starting position consisted in gazing at the central speaker (marked with a cross) and refrain from moving their head until the beginning of the sound. The system delivered the sound only when head (HMD) and eyes (tracked) were directed toward the centre. In such a way, we obtained replicable head and eyes posture across trials and participants. As soon as the sound started, participants were free to move their eyes and head as they wanted.

After each block, participants were given breaks and were asked to judge their performance (“How do you judge your performance?” from 1 = really bad to 9 = really good) and their perceived effort during the block (“How much effort (in terms of energy and cognitive resources) did the task require?” from 1 = none to 9 = many). At the end of 5 blocks, participants were invited to complete a questionnaire presented using Google Forms. The questionnaire comprised general questions about personal data (i.e., age, gender), questions to investigate their relationship with sound during life, specific questions to deepen the quality of our VR technology and its ergonomics, a question concerning their strategy to perform the task, two

questions concerning the feeling of being able to switch the sound off (Agency) and 15 items of 2 scales (Focus of Attention and Satisfaction), used typically to measure the engagement in video game-based environments (Figure 2). We adapted these items from the work of Wiebe and colleagues (2013) to make them more suitable for our VR experience. Participants indicated their agreement with each item using a 5-points Likert scale (from 1 = totally disagree to 5 = totally agree).

We divided participants in 2 groups: the Reaching group and the Naming group (Figure 1B). The Reaching group performed the task by moving a controller with their dominant hand to touch the speaker emitting the sound. The Naming group performed the same task by reading aloud the number located above the presumed source speaker (see Fig 1B). If the response was correct, the sound immediately stopped. If the response was wrong, the correct speaker produced a pulsing red light while the sound continued. This provided participants with audio-visual feedback about the correct position of the sound source. Participants either moved the controller in contact with the correct source or named the number of the correct speaker. When the correct speaker was reached (or named) the sound stopped. Importantly, participants were informed that they were free to exploit head-movements while listening to target sounds, without time pressure on their response. Both reaching and naming responses were followed by a brief vibration of the hand-held controller (which was kept in hand also during the naming task). When participants in the Naming group were tested, a second experimenter in the room promptly registered the responses on the computer. Specifically, the second experimenter pressed a key on an additional VIVE controller to specify whether the response was correct or wrong, and at the end of the trial entered the spoken labels in the computer.

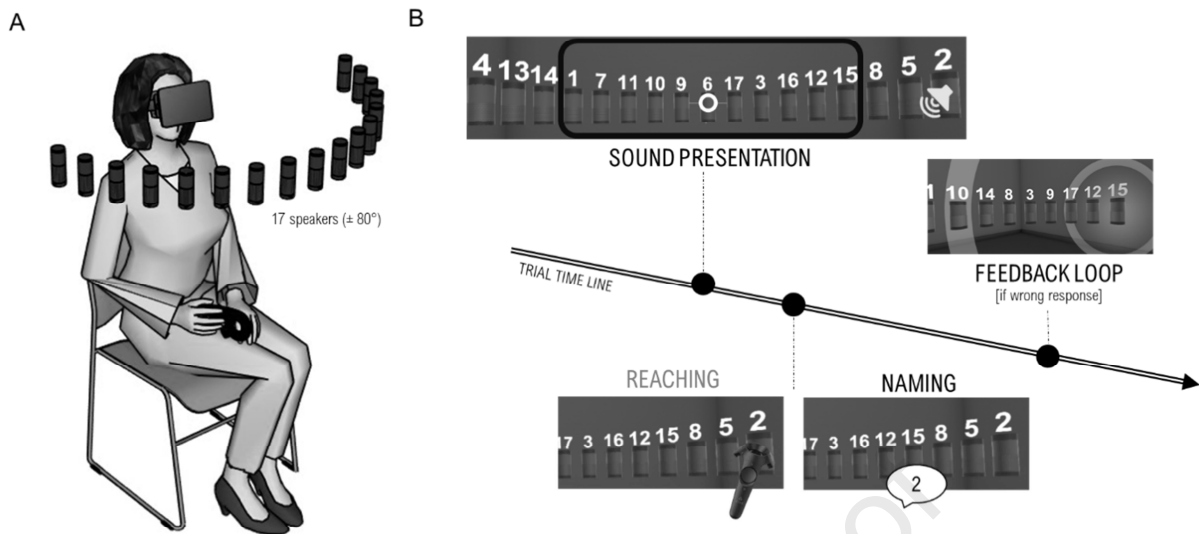


Figure 1. **EXPERIMENTAL PROCEDURE AND SETTING.** (A) Schematic representation of participant and speaker positions. (B) Schematic representation of 1 trial: participants saw the virtual room and part of the speakers' array (the black square represents participants starting view). They were allowed to turn their heads and move their eyes to see all speakers. In the Reaching group, participants stopped the sound by reaching and touching with the controller the virtual speaker emitting the sound. In the Naming group, the sound was stopped by the experimenter when participants read aloud the label above the speaker emitting the sound. If the answer was wrong, the correct source started to flash providing a visual feedback.

Head-movement pre-processing

To study head-movements, we calculated the tangential velocity on the x, y, z axis (expressed in degrees of rotation) using two-points central difference derivate algorithm (Bahill & McDonald, 1983) with 5 points for the half-window. The onset and the end of the movements were computed automatically using a velocity-based threshold ($10^\circ/\text{s}$). Head-movements were checked manually by visualizing the spatial rotation changes of the head and its speed using a home-made tool box in MATLAB R2018b.

3. RESULTS

3.1. Sound localisation performance

3.1.1. Immediate ear-plug effect

To assess the immediate effects of monaural plugging, we compared absolute and signed errors before and after ear-plugging. Absolute and signed errors (calculated in each trial as the absolute or signed difference in degrees between actual and reported azimuth position of the sound, respectively) were entered into separate Analysis of Variance (ANOVA) with TARGET POSITION (17 positions, from -80° to $+80^\circ$) and LISTENING CONDITION (binaural B, monaural M1) as within-participants variables and GROUP (reaching, naming) as between-participants variable. All statistical analyses using the software JASP 0.9.1.0 and R-studio (version 1.0.143).

The analysis on absolute errors revealed a main effect of LISTENING CONDITION, $F(1,26) = 64.79$, $p < .001$, $\eta^2 = .71$, a main effect of TARGET POSITION, $F(3.95, 102.60) = 5.96$, $p < .001$, $\eta^2 = .18$, and the 2-way interaction between LISTENING CONDITION and TARGET POSITION, $F(3.9, 101.51) = 7.36$, $p < .001$, $\eta^2 = .22$. Sound localisation error increased significantly in M1 compared to the B listening condition (see black and light grey lines in Figures 2A and 2B), particularly for target delivered more towards the plugged ear (left). This effect emerged irrespective of GROUP (main effect and all interactions, p -values $> .27$). Likewise, the analysis on signed errors revealed a main effect of TARGET POSITION, $F(4.69, 121.90) = 9.42$, $p < .001$, $\eta^2 = .26$, and the 2-way interaction between LISTENING CONDITION and TARGET POSITION, $F(4.33, 112.64) = 8.08$, $p < .001$, $\eta^2 = .23$. Signed errors increased particularly for target delivered more towards the plugged ear (left). For signed errors we also found a main effect GROUP, $F(1,26) = 6.45$, $p = .02$, $\eta^2 = .20$ and the interaction between GROUP and LISTENING CONDITION ($F(1,26) = 4.36$, $p = .05$, $\eta^2 = .26$). Simple main effects revealed that the signed error of the Naming group was significantly impacted by listening condition ($F = 4.9$, $p = 0.05$), whereas this did not occur in the Reaching group ($F = 0.26$, $p = .62$). As shown in Figure 2B, during monaural listening the Naming group localised sound more to the unplugged side (Binaural: $M = -0.39$; $SD = 3.14$; Monaural: $M = 4.36$; $SD = 16.53$), whereas the Reaching group did not present this bias (Binaural: $M = -0.88$; $SD = 3.55$; Monaural: $M = -1.69$; $SD = 14.34$).

Participants were aware of their performance change after monaural plugging. Performance assessment scores ("How do you judge your performance?" from 1 = really bad to 9 = really good) were

lower in M1 (Reaching: $M = 3.29$; $DS = 1.90$; Naming: $M = 3.50$, $DS = 1.40$) compared to the B listening condition (Reaching: $M = 6.86$; $DS = 1.35$; Naming: $M = 6.57$, $DS = 0.85$; Chi-square = 24.14, $p < .001$, in non-parametric Friedman test), irrespective of GROUP. Participants also reported that monaural plugging required additional cognitive effort. Effort assessment scores (*"How much effort (in terms of energy and cognitive resourced used) did the task require?"* from 1 = none to 9 = many) were larger in M1 (Reaching: $M = 6.71$; $DS = 1.33$; Naming: $M = 6.21$, $DS = 1.72$) compared to the B listening condition (Reaching: $M = 3.29$; $DS = 1.49$; Naming: $M = 3.57$, $DS = 1.95$; Chi-square = 28, $p < .001$, in non-parametric Friedman test), again irrespective of GROUP.

3.1.2. Adaptation to monaural listening

To study adaptation to monaural listening across successive monaural blocks, we entered absolute and signed errors in separate ANOVAs with TARGET POSITION (17 positions, from -80° to $+80^\circ$) and MONAURAL BLOCKS NUMBER (M1, M2, M3 and M4) as within-participants variables, and GROUP (reaching, naming) as between-participants variable. The analysis on absolute errors revealed a main effect of MONAURAL BLOCK NUMBER, $F(3,78) = 8.51$, $p < .001$, $\eta^2 = .24$, a main effect of TARGET POSITION, $F(16, 416) = 8.89$, $p < .001$, $\eta^2 = .24$, and the 2-way interaction between MONAURAL BLOCK NUMBER and TARGET POSITION, $F(48, 1248) = 3.02$, $p < .001$, $\eta^2 = .10$. Participants improved across M blocks, especially for target positions more towards the plugged ear (left) (Figures 2A and 2B). No main effect or interaction involving the GROUP variable emerged (all p-values $> .18$), when considering overall performance in a block. A similar analysis on signed errors revealed a main effect of MONAURAL BLOCK NUMBER, $F(3,78) = 5.252$, $p = .002$, $\eta^2 = .16$, a main effect of TARGET POSITION, $F(16, 416) = 10.68$, $p < .001$, $\eta^2 = .28$, and the 2-way interaction between monaural block number and target position, $F(48, 1248) = 2.64$, $p < .001$, $\eta^2 = .09$. Participants bias changed across M blocks, especially for target positions on the plugged side (left) (Figures 2A and 2B). While participants perceived targets more toward the unplugged side at the beginning, repeating the task their started perceived the targets more toward the plugged side. Despite this pattern seems to differ in the 2 groups (the Naming group minimize the bias across blocks, while the Reaching group increased it toward the plugged side), no main effect or interaction involving the group variable emerged (all p-values $> .08$), when considering overall performance in a block.

To further investigate changes in localisation error as a function of practice in the two groups, we analysed performance changes across the 204 trials (51 trials in each of the 4 blocks) using a linear mixed

model (LMM). Specifically, we entered the absolute error in a model with GROUP (naming, reaching) and PRACTICE (as a continuous variable of 204 items). Participant and target position were considered as random effects in the model. This analysis revealed that absolute errors were influenced by PRACTICE ($t = -4.45$, $p < .001$), revealing that participants learned across monaural trials. Crucially, our model showed that error reduction was across trials was influenced by the type of task performed by participants. The Reaching group reduced localisation errors more rapidly and to a greater extent compared to the Naming group ($t = -3.09$, $p = .002$) as shown in Figure 3 A. A similar analysis on signed error revealed that signed errors was generally influenced by GROUP $t = -2.58$, $p = .01$). As shown in Figure 3B, the Reaching group in general perceived the sound as shifted toward the plugged side, while the Naming group perceived it more toward the unplugged. Interestingly, both groups modified their bias by repeating the task ($t = -6.73$, $p < .001$), but differently ($t = 2.71$, $p = .007$). While the Naming group changed the direction of the bias across trial repetition, the Reaching group increased the bias toward the plugged side.

Participants were aware that their performance improved across monaural blocks. Performance self-assessment scores increased across monaural blocks (M4: Reaching: $M = 4.29$; $DS = 1.38$; Naming: $M = 3.93$, $DS = 1.69$; Chi-square = 10.02, $p < .001$, in non-parametric Friedman test), irrespective of GROUP. Participants did not report any change in terms of energy and cognitive resources used to do the task across monaural blocks: effort assessment scores did not change across monaural blocks in either GROUP (all $ps > .23$).

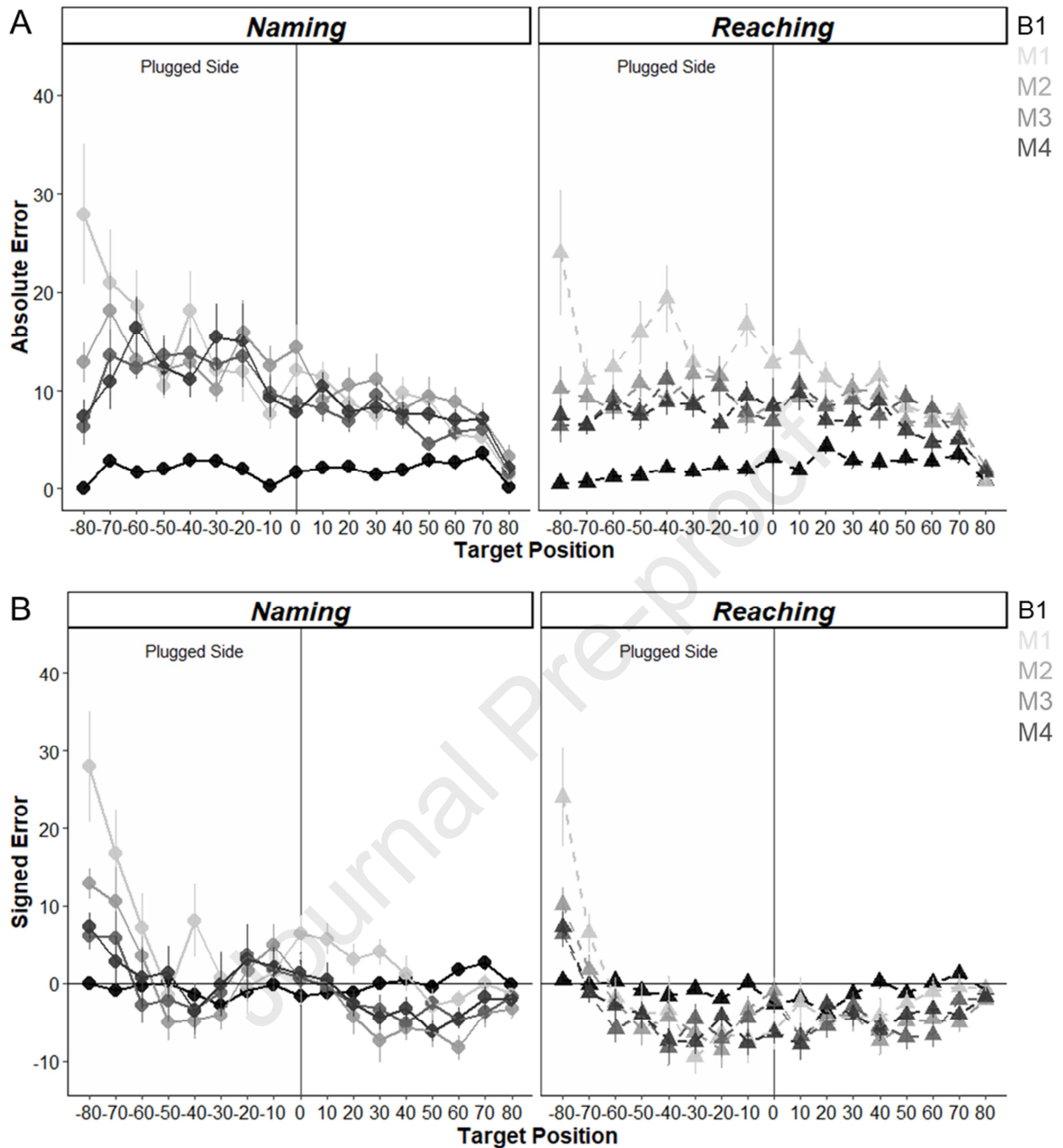


Figure 2. **SOUND LOCALISATION PERFORMANCE.** Absolute (A) and signed (B) errors (in degrees) as a function of Blocks (B = binaural block, M1 = first monaural block, M2 = second monaural block, M3 = third monaural block, M4 = fourth monaural block) and group (Naming: left panel, circles; Reaching: right panel, triangles). Error bars represent the standard error of the mean.

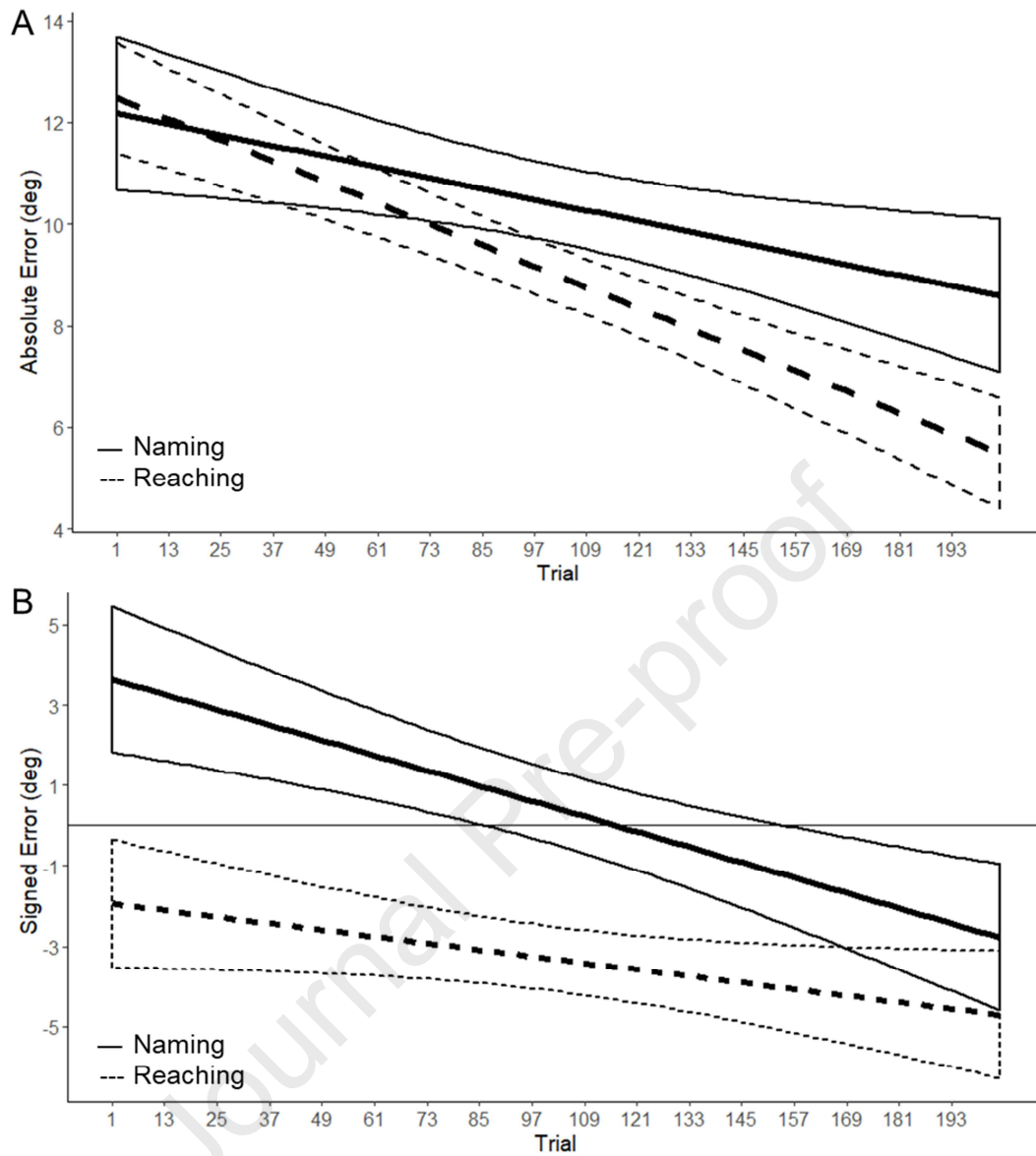


Figure 3. **SOUND LOCALISATION PERFORMANCE AS FUNCTION OF TRIAL.** Absolute (A) and signed (B) error as a function of trial for Naming (continuous line) and Reaching (dashed line) groups.

3.2. Head movements

To study head-movement behaviour, we measured the amplitude of the head rotation around the vertical axis for each detected movement. Next, we transformed the movement amplitude into a categorical variable with 10 bins (amplitude in each bin = 15,8 degrees; bins: 1 = 1.0° - 16.8°; 2 = 16.8° - 32.7°; 3 = 32.7° - 48.5°; 4 = 48.5° - 64.3°; 5 = 64.3° - 80.1°; 6 = 80.2° - 96.0°; 7 = 96.3° - 111.9°; 8 = 113.9° - 126.8°; 9 = 126.8° - 143.2°; 10 = 143.2° - 159.3°) and counted the number of head movements in each amplitude bin, separately for each group and experimental block.

Furthermore, we considered two additional variables computed across all trials: head-rotation around vertical axis as a function of target position and overall head-direction bias during an entire block. The latter two indices were extracted by linear fitting of the changes in head-rotation extent as a function target position (see example in Figure 3). Head-rotation extent was calculated as the average between the leftmost and rightmost head-rotation, for any given target position. The slope of the fitting line captures head-rotation around the vertical axis: the larger the slope the more participants rotated their head as a function of target eccentricity. In the representative participant (ID=21) shown in Figure 4, the slope remained stable from B to M1 (0.24 vs. 0.22) but increased from M1 to M4 (0.22 vs. 0.40), indicating a larger propensions to rotate the head towards peripheral targets in M4 compared to M1. The intercept of the fitting line captures the overall head-direction bias during an entire block, with negative values indicating a bias towards the left (plugged) side. In the same representative participant, the intercept became progressively more negative from B to M1 to M4 (-0.22, -1.73, -17.14, respectively) revealing the bias of this participant to rotate the head further to the plugged side, thus bringing the unplugged ear towards the target.

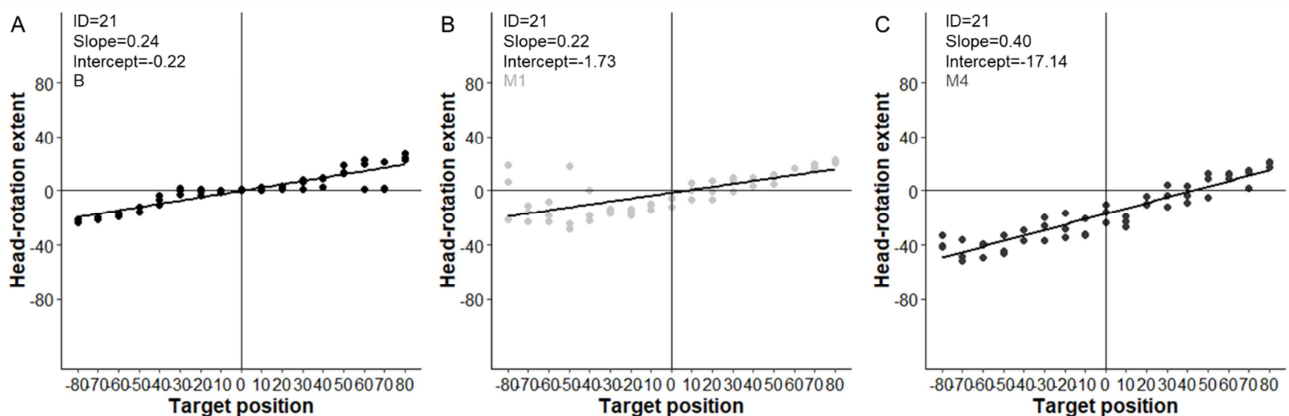


Figure 4. **SLOPE AND INTERCEPT INDECES CALCULATION** Head-rotation extent as a function of target position of participants 21 during first binaural block B (A), the first monaural block M1 (B) and the last monaural block M4 (C).

3.2.1. Immediate ear-plug effect

We started by studying the immediate effect of monaural plugging on the number of head movements in the two groups, as a function of head-movement amplitude. To this aim, a Poisson regression model was fitted on the observed number of movements using AMPLITUDE BIN (covariate), GROUP (Naming, Reaching), LISTENING CONDITION (Binaural, Monaural), and all interaction terms. The participant was included as random effect. The deviance table¹ for this model revealed a significant effect of AMPLITUDE BIN ($\chi^2(1) = 157.1710$, $p < 0.001$). Number of movements was inversely related with amplitude, as evident from the corresponding negative Poisson parameter estimate ($\beta_{\text{Amplitude}} = -0.58333$, $z = -12.537$, $p < 0.001$). Number of movements also increased significantly from the binaural (B1) to the first monaural (M1) listening block ($\beta_{m1} = 0.67464$, $z = 6.647$, $p < 0.001$) for both groups (main effect of Block, $\chi^2(1) = 44.1786$, $p < 0.001$). This increase concerned primarily head-movements of smaller amplitude (Amplitude by Block, $\chi^2(1) = 15.861$, $p < 0.001$). All the other effects were not statistically significant ($ps > 0.05$) (Figure 5A).

Next, we examined how participants explored auditory space with their head during sound emission, in the binaural and in the first monaural block. To this aim we focused on (1) head-rotation around vertical axis as a function of target position (slope of the fitting line, as specified above) and (2) overall head-direction bias during an entire block (intercept of the fitting line). Both measures were entered separately into an ANOVAs with LISTENING CONDITION (B, M1) as within-participants variable and GROUP (reaching, naming) as between-participants variable. The analyses on head-rotation around vertical axis revealed that participants in the Reaching group turned their head as a function of target position to a greater extent compared to those in the Naming group (main effect of group, $F(1,26) = 5.29$, $p = .03$, $\eta^2 = .17$). In other words, they explored auditory space with their head to a greater extent (compare B and M1 in Figures 6A). No main effect or interaction involving LISTENING CONDITION emerged, indicating that this group difference was already present in the binaural listening condition.

The analysis on the overall head-direction bias revealed instead a main effect of LISTENING CONDITION, $F(1,26) = 19.09$, $p < .001$, $\eta^2 = .42$. Intercepts became negative in M1 (Naming: $M = -3.41$, $SD = 6.20$; Reaching: $M = -3.39$, $SD = 4.31$) compared to the B listening condition (Naming: $M = 0.56$, $SD = 1.25$; Reaching: $M = 0.90$, $SD = 1.50$). This means that, when listening with the left ear plugged, all participants

¹ Deviance decomposition table for the Poisson model derived using the Anova function in the Car package by John Fox. We adopted the glmer function in the lme4 package for parameter estimates and the lmerTest package for the computation of the corresponding p-values.

turned their head leftward, to approach the sound with their unplugged (right) ear (compare B and M1 in Figure 6B).

3.2.2. Adaptation to monaural listening

To study the effect of the adaptation to monaural listening on head-movements, we started by examining the number of head movements in the two groups, as a function of head-movement amplitude, in the first and in the last monaural block. As before, a Poisson regression model was fitted on observed number of movements as a function of AMPLITUDE BIN (covariate), GROUP (Naming, Reaching), BLOCK (M1, M4) with all interaction terms. Participant number was included as random effect. The deviance table revealed a significant effect of AMPLITUDE BIN ($\chi^2(1) = 584.4407$, $p < 0.001$). Number of movements was inversely related with amplitude as evident from the corresponding negative Poisson parameter estimate ($\beta_{\text{Amplitude}} = -0.50891$, $z = -24.175$, $p < 0.001$). Number of movements also increased significantly from the first (M1) to the fourth (M4) monaural listening block, for larger amplitude levels ($\beta_{\text{Amplitude}, M4} = 0.11675$, $z = 4.321$, $p < 0.001$). This was also reflected from the significant interaction effect BLOCK by AMPLITUDE BIN ($\chi^2(1) = 18.6691$, $p < 0.001$). We also observed a significant effect for GROUP ($\chi^2(1) = 5.5511$, $p < 0.05$). The Reaching group was associated with a lower number of movements ($\beta_{\text{Reaching}} = -0.33190$, $z = -2.356$, $p < 0.05$). All the other effects were not statistically significant ($ps > 0.05$) (Figure 5B).

Finally, we examined how participants explored auditory space with their head across monaural blocks. Slopes and intercepts across participants were entered into separate ANOVAs with MONAURAL BLOCK NUMBER (M1, M2, M3 and M4), as within-participants variable and GROUP (reaching, naming) as between-participants variable. Both analyses revealed only a main effect of MONAURAL BLOCK NUMBER (slope: $F(3,78) = 12.67$, $p < .001$, $\eta^2 = .32$; intercept: $F(1.57,40.86) = 9.47$, $p < .001$, $\eta^2 = .27$). As monaural blocks progressed, participants in both groups increased their exploration of auditory space and turned their head more to approach sounds with their unplugged (right) ear (Figure 6A and 6B).

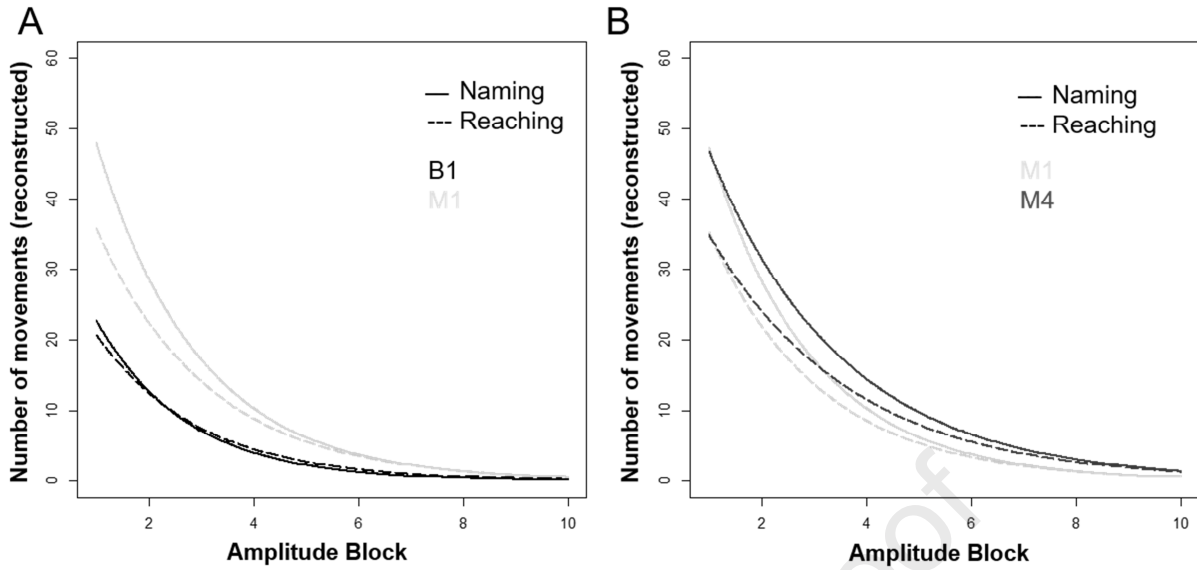


Figure 5. **NUMBER OF MOVEMENT AND AMPLITUDE** A) Poisson model representation. N° of movements as a function of Amplitude block (data binning: 10 bins: 1 = 1.0° - 16.8°; 2 = 16.8° - 32.7°; 3 = 32.7° - 48.5°; 4 = 48.5° - 64.3°; 5 = 64.3° - 80.1°; 6 = 80.2° - 96.0°; 7 = 96.3° - 111.9°; 8 = 113.9° - 126.8°; 9 = 126.8° - 143.2°; 10 = 143.2° - 159.3°); Black patterns correspond to binaural listening condition (B1) and light grey pattern correspond to monaural listening condition (M1). The continuous patterns represent the Naming group while dashed one the Reaching group. B) Poisson model representation. N° of movements as a function of Amplitude block. Light grey patterns correspond to the first monaural block (M1) and dark grey pattern correspond to the fourth monaural listening condition (M4). The continuous patterns represent the Naming group while dashed one the Reaching group.

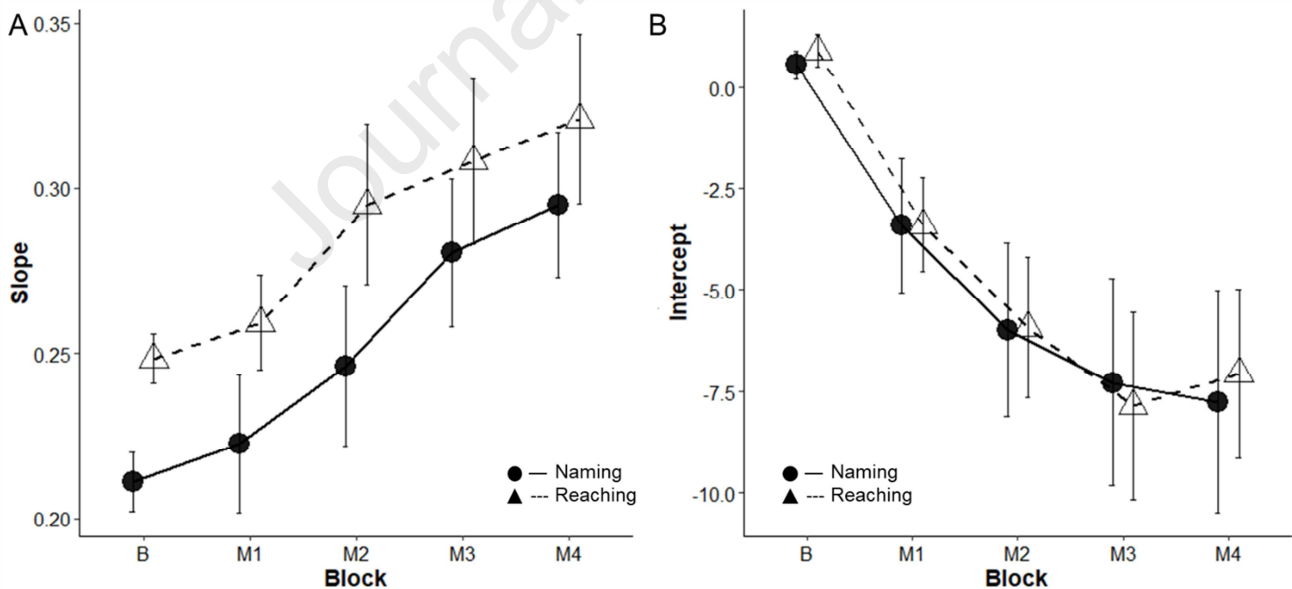


Figure 6. **SPACE EXPLORED.** Slope (A) and intercept (B) values as a function of block B1 M1 M2 M3 M4. Naming group (black circles) and Reaching group (white triangles).

3.3. Relation between changes in sound localisation performance and head-movements

3.3.1. Immediate ear-plug effect and head-movements

We first examined if the immediate ear-plug effect (i.e., changes in absolute localisation error between M1 and B; with positive values indicating larger plug-effect) correlated with changes in the number of head movements (i.e., changes between M1 and B). When studied separately for each group, using Pearson correlation with Bonferroni correction, a significant relation emerged only for the Naming group (naming: $r = 0.57$, $p = 0.04$; reaching: $r = 0.12$, $p = 0.68$). However, this correlation was driven by a single participant ($r = 0.22$, $p = 0.46$ with the participant 18 was removed).

Next, we examined if the immediate ear-plug effect (i.e., changes in absolute localisation error between M1 and B; positive values indicate larger plug-effect) correlated with changes in the space explored with the head (i.e., changes in slope between M1 and B). Unlike the case for number of head-movements described above, this correlation was reliable in each group (Pearson correlation with Bonferroni correction: reaching, $r = -0.66$, $p = 0.01$; naming, $r = -0.79$, $p < 0.001$; see Figure 7A). Irrespective of group, the more participants increased explored space between B and M1 blocks, the less the cost of ear plugging on their sound localisation performance.

3.3.2. *Adaptation to monaural listening and head-movements*

The number of head movements also increased between M1 and M4 block (see section 3.2.2). We examined if adaptation (i.e., changes in absolute localisation error between M4 and M1; positive values indicate larger improvement after trials repetition) correlated with the changes in the number of head movements (i.e., changes between M4 and M1). We did not find any relations for both groups (all $ps > 0.31$).

The portion of space explored with the head also changed across monaural blocks (comparison between M1-M4 blocks, see section 3.2.2). We examined if adaptation (i.e., changes in absolute localisation error between M4 and M1; positive values indicate larger improvement after trials repetition) correlated with changes in the space explored with the head (i.e., changes in slope between M4 and M1). This correlation is shown in Figure 7B, separately for each group (Pearson correlation with Bonferroni correction: reaching, $r = 0.60$, $p = 0.02$; naming, $r = 0.01$, $p = 0.97$). A positive relation emerged for the Reaching group, but not for the Naming group. Comparing correlations across group (as in Eld, Gollwitzer and Schmidt, 2011) produced a marginally significant between-group difference ($z = -1.602$, $p = 0.055$). This suggests that compensatory head-movement behaviour may relate with adaptation more for participants in the Reaching group.

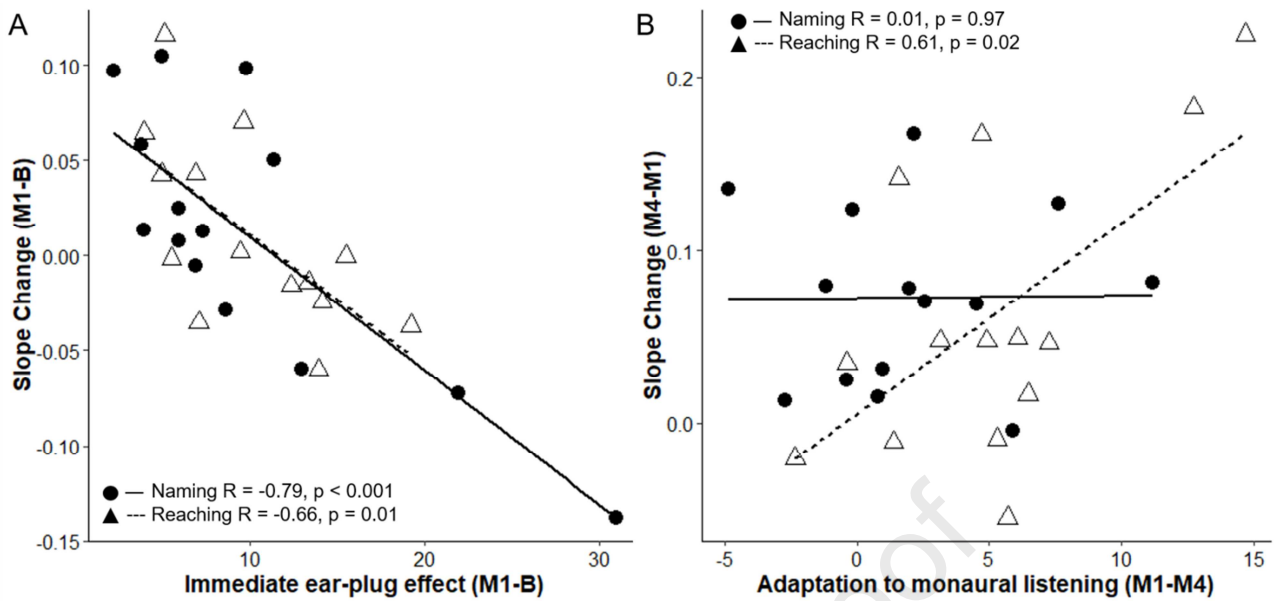


Figure 7. **CORRELATION BETWEEN SPACE EXPLORED ROTATING THE HEAD AND PERFORMANCE** A) Correlation between Ear-plug effect and Slope Change, B) Correlation between Slope Change and Adaptation to monaural listening. In both graphs participants in the Reaching group are indicated with white triangles (regression line shown as dashed) and those in the Naming group are indicated with black circles (regression line shown as continuous).

3.4. VR experience

At the end of the experiment, participant replied to questions about their experience using VR system, to assess also the feasibility of our approach. They judged that the overall experience with virtual reality was positive ($M = 7.86$, $SD = 1.15$, using a scale from 1 = negative to 9 = positive), that the scene appeared realistic ($M = 6.61$, $SD = 1.47$, using a scale from 1 = unrealistic to 9 = very realistic) and reported no feeling of losing balance ($M = 1.53$, $SD = 1.07$, using a scale from 1 = not at all to 9 = completely) or being annoyed ($M = 2.61$, $SD = 1.81$, using a scale from 1 = not at all to 9 = completely).

Participants were also queried about the sense of agency and the perceived gaming experience when performing the two tasks (i.e., reaching or naming).

Sense of agency

Participants indicated their agreement with 2 items (“I felt that I can turn off the sound”, “I felt that I can act on the sound”) using a 5-points Likert scale (from 1 = totally disagree to 5 = totally agree). Participants in both groups felt equally that they could act on sounds (reaching: $M = 2.29$, $SD = 0.99$; naming: $M = 1.71$, $SD = 0.99$; $W = 64$, $p = .14$ on Mann Whitney). However, the Reaching group reported the experience of turn off

the sound strongly ($M = 4.43$, $SD = 0.51$) compared to the Naming group ($M = 1.21$, $SD = 0.58$), $W = 0.00$, $p < 0.001$ (Mann Whitney).

Gaming experience

To investigate their gaming task-related experience, we proposed 15 items of 2 scale of the User Engagement Scale (UES) questionnaire, typically used to measure engagement during video-game playing (Wiebe, Lamb, Hardy & Sharek, 2014): focus of attention scale and satisfaction scale (Figure 8). We compared responses of the two groups for each item of the 2 scales, as well as the cumulative indices (average of the respective items, Cronbach's α : 0.8; 0.89) in t-test analyses. Overall, our virtual reality scenario was an effective method to create a game-like situation, capable of actively involving participants. The satisfaction scale items revealed a difference between tasks: participants of the Reaching group ($M = 3.71$, $SD = 0.83$) perceived the experience more rewarding compared to the Naming group ($M = 2.93$, $SD = 0.83$; $W = 47.50$, $p = .015$). Moreover, they expressed more interest in doing the task (Reaching: $M = 4.71$, $SD = 0.47$; Naming: $M = 4.07$, $SD = 0.92$, $W = 60$, $p = .05$).

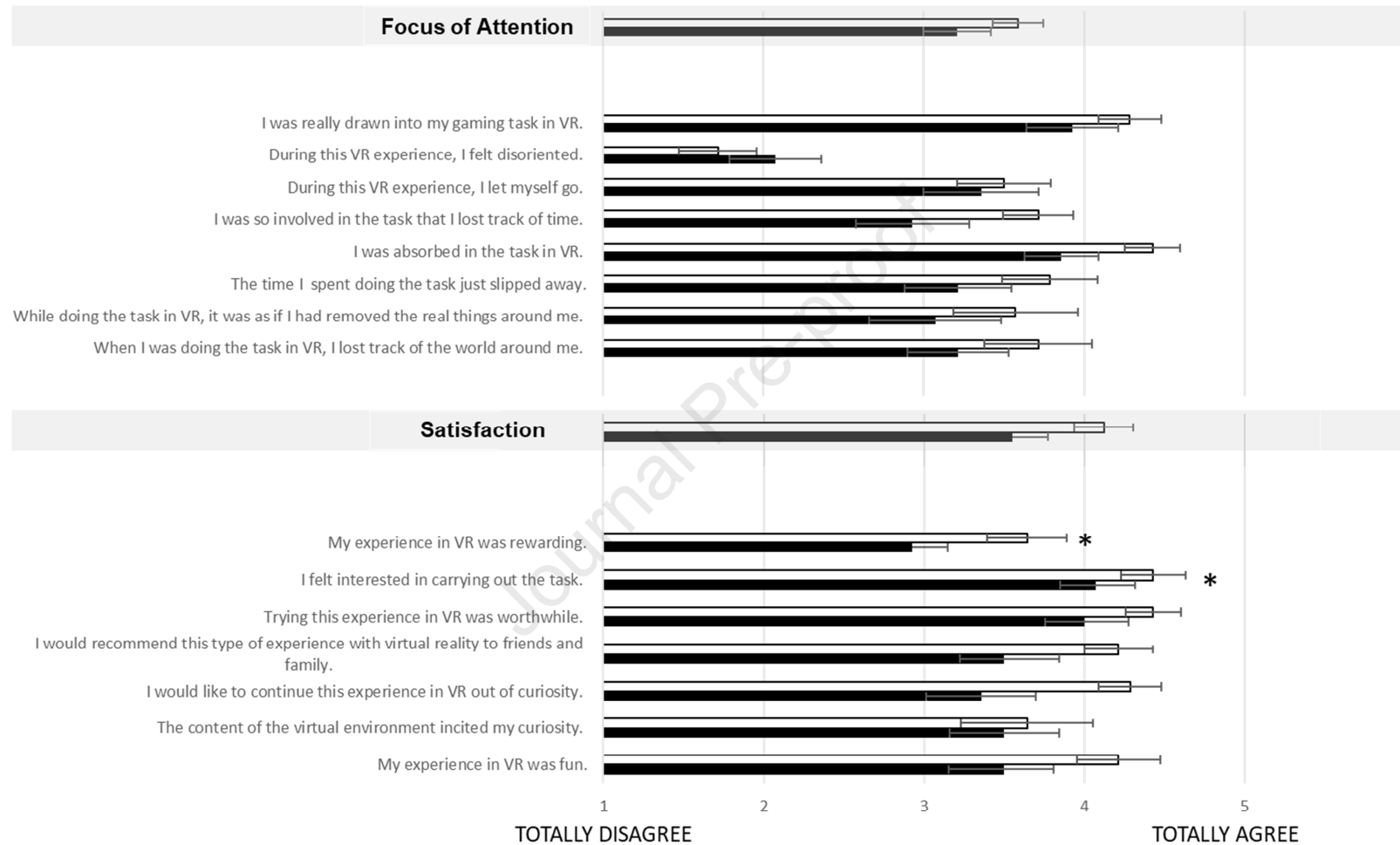


Figure 8. **VR EXPERIENCE.** Mean ratings for the Reaching (white bars) or Naming (black bars) group. Cumulative scores for focus of attention and satisfaction are highlighted in grey. The error bars represent the standard error of the mean.

4. DISCUSSION

The cognitive ability to localise sounds relies on internal models that specify the correspondences between the auditory input reaching the ears, initial head-position and coordinates in external space. It is now well established that these correspondences can be updated throughout life (Carlile, 2014; Keating & King, 2015; Rabini et al., 2019; Strelnikov et al., 2011; Van Wanrooij et al., 2005), setting the basis for re-learning spatial hearing abilities in adulthood. In addition, whenever auditory cues are altered, humans can exploit active perception strategies to adapt to the novel listening situations. For instance, people with hearing-impairment use head-movements to cope with their hearing difficulties (Brimijoin et al., 2010; see also Brimijoin & Akeroyd, 2012). Understanding the cognitive adjustments of internal models for spatial hearing, as well as the behavioural strategies adopted when localising sounds in novel listening conditions, is particularly important for individuals who experience long-term auditory alterations (e.g., hearing loss, hearing aids, cochlear implants). In addition, it may provide useful indications for training normal hearing individuals who must adapt to novel auditory cues (e.g., when listening with non-individualised HRTFs in virtual auditory environments).

Until now, however, several methodological constraints have limited our understanding of the mechanisms involved in adapting spatial hearing skills when exposed to novel listening situations. Specifically, the potential role of dynamic and sensorimotor interactions with the sounds have remained largely overlooked. These interactions include head-movements performed during sound, as well as reaching actions towards the sounds themselves. Using a novel methodology, based on virtual reality and real-time kinematic tracking of the participant's head and hand, we examined the contribution of reaching movements and head movements when adapting to monaural listening. Our experimental approach is based on a new apparatus (SPHERE; Pavani et al., 2017; Gaveau et al., 2020) that allows to present free-field sounds from a real loudspeaker, placed at pre-determined locations and aligned with a virtual reality scenario. This allows full control over visual cues about the visual environment (here, an empty room), cues to sound position (here, a visible array of virtual loudspeakers) and feedback (here, the flashing speaker that appeared in case of mistaken responses). Most importantly, our approach allows full monitoring of head position. This ensured identical straight-ahead initial posture for all trials and participants, but also permitted tracking of spontaneous head-movements during the response phase. Finally, precise kinematic of the hand response was measured in real-time. This allowed measuring of end-pointing positions in the Reaching group and ensured compliance with the instructions in the Naming group.

Four novel findings emerged. First, we documented that adaptation to monaural listening can occur over a relatively short period of time (about 200 trials, completed in approximately 50 minutes). Second, we showed that monaural plugging affected both sound localisation performance and spontaneous head-movements. Third, in line with our main prediction, we found that reaching to sounds proved more effective than naming sound position when adapting to altered binaural cues. Fourth, we document that head-movements played an important role in this fast adaptation: reduction in localisation errors was accompanied by wider portions of space explored with the head. Notably, performance improvements correlated with changes in head-movement extension selectively in the Reaching group, providing initial evidence for an interaction between reaching to sounds and head-movements when adapting spatial hearing abilities to monaural listening conditions.

4.1. Sound localisation improves rapidly during monaural listening

The present study demonstrates that normal hearing adults, listening with one ear plugged and muffled, can improve their sound localisation abilities over a short period of time. Both group of participants (Reaching and Naming) progressively reduced their error over trials, within a single testing session. We promoted adaptation to altered binaural cues through a combination of audio-visual feedback (Strelnikov et al., 2011; Rabini et al., 2019; Valzolgher et al., 2020) and dynamic sensorimotor interactions with the sounds (spontaneous head-movements, agency over the sound). Despite most of the participants were naïve to virtual reality, they easily understood the instructions, were engaged by the task, and gave a positive assessment to the VR experience overall. This confirms the feasibility of our novel approach, which represents an important pre-requisite for the use of similar VR applications in auditory perception research and clinical rehabilitation (see 3.4).

While adaptation to altered auditory cues has been documented in several previous reports (Carlile, 2014; Kacelnik, Nodal, Parsons & King, 2006; Keating & King, 2015; Nawaz, McNeill & Greenberg, 2014; Steadman et al., 2019), to the best of our knowledge, improvements on a trial-by-trial basis have not been studied and described before. Traditionally, spatial hearing adaptation has been examined in paradigms which comprised pre- and post-testing phases separated by training sessions which occurred over several days (e.g., 5 days in Strelnikov et al., 2011 and in Irving & Moore, 2011; 10 days in Ohuchi et al., 2005). Thus, even when adaptation of spatial hearing abilities occurred after a small number of short training session (3 sessions, lasting 15 minutes as in Steadman et al., 2019), they were performed across multiple

days. Although in our study we did not test the generalisation of the observed adaptation of spatial hearing abilities, we investigated the improvements occurring over the course of 204 trials administered in a single session.

Two mechanisms may subtend the fast adaptation we have documented. On the one hand, it is possible that participants learned to adjust the contribution of the different auditory cues to sound localisation. Re-weighting of the auditory cues, with increased reliance on monaural cues at the unplugged ear for localization, is one of the mechanisms by which sound localization in azimuth can improve during monaural listening (see Kumpik & King, 2019). We designed the experiment to selectively impact on binaural cues (both interaural-level differences and inter-aural time differences) while leaving monaural cues at the unplugged ear intact. Hence, participants may have learned to exploit monaural cues to a greater extent during the session. On the other hand, because participants were free to move their head during sound emission, it is possible that strategic (rather than cognitive) adaptation mechanisms could account for the observed fast improvements in sound localisation. Cognitive and strategic contributions to sound localisation improvements are clearly not mutually exclusive, and our study cannot determine their relative contributions to the observed performance changes. Yet, as described in the next section, we suggest a key role of compensatory behavioural strategies related to head-movements.

4.2. Monaural listening triggered compensatory head-movements

The role of head-movements in sound localisation has been emphasised since the 1940's (Wallach, 1940), but it has been studied sporadically until recent years (Iwaya, Suzuki & Kimura, 2003; Morikawa & Hirahara, 2013; Perrett & Noble, 1997; Toyoda, Morikawa & Hirahara, 2011). Here, we characterised head-movements with different variables (number and amplitude of head-movements, head-rotation around vertical axis, and overall head-direction bias during an entire block). Our results show that head-movements changed immediately when passing from binaural to monaural listening conditions. In the first monaural listening block, the number of head movements increased, specifically for movements of smaller amplitude. In addition, participants turned their heads to approach sounds with their unplugged ear (as evidenced by immediate changes in the intercept). Our results also show that participants adapt their head-movements across the monaural blocks to adapt to the altered listening condition. Specifically, the number of head-movement continued to increase, but this time for movements of larger amplitude (compare Figure 5B and 5A). Moreover, participants increased their exploration of auditory space by turning their head to more

eccentric positions (as evidenced by changes in slope; see Figure 6A) and turned their head even more to approach the sounds with their unplugged ear (see Figure 6b).

Many previous works showed that head-orientation and head-movements improve sound localisation performance in azimuth and elevation (Morikawa & Hirahara, 2013; Perret & Noble, 1997; Thurlow & Runge, 1967) and facilitate disambiguation of front/back confusions (Brimijoin et al., 2012; Kim, Barnett-Cowan, & Macpherson, 2013; McAnally & Martin, 2014; Perret & Noble, 1997; Wightman & Kistler, 1999; Pöntynen & Salminen, 2019). This document the well-known effect of head-movements on spatial hearing. However, much less work has described the actual changes in head-movements that occur in difficult or altered listening conditions – as here. In a task that required head orienting to spoken sentences, Brimijoin and colleagues (2010) observed that normal hearing listeners performed smoother and more sigmoidal head-movement trajectories compared to hearing-impaired listeners. In the latter group, abrupt velocity changes, reversal of directions and frequent corrections tended to prevail. Note however, that this study described the kinematic of head-movements when participants were explicitly instructed to point to the sound, whereas we reported the head-movement pattern observed in the spontaneous movements produced while the sound was playing and before the response. Although other works have examined spontaneous head-movements in novel and challenging listening conditions (as here), they did so when participants were not explicitly instructed to localise sounds. For instance, Brimijoin and colleagues (2012) has shown that participants use head-movements to increase the level of the signal in their better ear, or to increase the difference in level between the signal and the noise, when listening in complex auditory scenes. Remarkably, this type of head movement differs compared to those performed in more typical hearing situations (i.e., binaural hearing), whereby participants rather tend to face the speaker frontally when performing a sounds localisation task (Iwaya et al., 2003; Thurlow, Mangels & Runge, 1967). A recent study by Hadley and colleagues (Hadley, Brimijoin & Whitmer, 2019) has shown that, as noise increases, people move closer to the target rather than make use of the potentially beneficial effects of head-orientations.

In our study, participants appear to have moved their head during monaural listening to achieve two goals: (1) orienting of their unplugged ear to the sound; (2) exploring auditory space to a greater extent. The need to cope with monaural listening was fulfilled by increasing the frequency and extent of head-turns to favour perception with the unplugged ear. This peculiar type of movement could have allowed greater exploitation of monaural cues. Specifically, listeners could have rapidly made use of changes in spectral monaural cues across the blocks, or – perhaps more likely – they could have exploited the attenuating

effects of the head to identify the azimuthal position of sounds (for discussion of the different use of monaural cues in actual or simulated unilateral hearing loss see Kumpik & King, 2019). Along a similar line of reasoning, greater exploration of auditory space may have permitted participants to 'search' for the region of maximal sound intensity. A third possibility is that participants learned to use their asymmetric binaural cues to localise sounds in azimuth (as proposed by Agterberg, Snik, Hol, Van Wanrooij & Van Opstal, 2012, albeit for listeners with acquired unilateral conductive hearing loss who had much longer experience with the asymmetrical binaural cues). In particular, it is possible that they learned to determine the head-orientation that brought the target sound on their medial plane (i.e., 0 ITD orientation). In future works, it would be important to assess the relative contribution of spectral and intensity monaural information, by affecting the spectral information at the unplugged ear using an ear-mould.

4.3. Reaching to sounds promote adaptation to monaural listening more than naming

One clear result that emerged from our study is that the instructions on how to perform the task (Reaching vs. Naming) impacted on sound localization performance, head-movements during sounds and adaptation to monaural listening in general. In the first monaural listening condition (M1), the Naming group perceived sounds more toward the unplugged side, whereas the Reaching group did not present this bias (Figure 2B). In addition, across trial repetitions, participants of the Reaching group reduced localisation errors faster compared to those who interacted with sounds simply by reading labels (the Naming group). This result revealed that Reaching to sound was effective in promoting faster adaptation to monaural listening. Across monaural trials, the error changed differently. As indicated by the signed error (which refers to the lateralized bias of the response), the Naming group reduced the bias toward the unplugged side across trials, while the Reaching group developed a slight bias towards the plugged side (Figure 3B).

Response type also affected head-movements during sound emission. Participants in the Reaching group turned their head as a function of target position to a greater extent compared to the Naming group, irrespective of listening condition (binaural vs. first monaural block; see 3.2.1). Interestingly, when considering the progressive adaptation to monaural listening blocks, only the Reaching group showed a correlation between head-movement extension (slope) and performance (Figure 7). Thus, although participants in both groups were free to move the head during the task, only participants in the Reaching group appeared to have developed effective dynamic listening strategies involving head-movements, which in turn impacted on adaptation to monaural listening.

Why the response type (Reaching vs. Naming) influenced adaptation to monaural listening? As anticipated in the introduction, reaching to sounds requires a coordination of different effectors (eyes, head, hand) into a common reference frame (Cohen & Andersen, 2002). In turn, this may result in a more stable (or salient) spatial coding of sound source location and favour re-learning of sound-space correspondences. In addition, reaching movement may have helped directing attention towards the position occupied by the sound source, because of additional visual, proprioceptive and kinaesthetic cues resulting from the action. The analysis of spontaneous head-movements, however, revealed a more complex scenario than expected. Reaching participants explored a wider portion of space with the head compared to Naming participants, and this was functional for adaptation selectively for the Reaching group. Movements of smaller amplitudes, appeared to have played a less important role for adaptation: they were more numerous for Naming than Reaching participants, but only in the first monaural block. At present we can only speculate about the origin of this different head-movement behaviour (and consequent change in performance) in the Reaching group. One possibility, is that head-movements were more functional to the key task (sound localisation) for the Reaching group than the Naming group. Specifically, it could be argued that participants in the Naming group used head-movements also to explore the labels that identified the speakers (recall that labels changed randomly between trials and thus they had to be read anew every time). In turn, this could have resulted in head-movements that could have interfered with the ones needed to optimally adapt to the monaural listening condition. Another possibility is that Reaching was a more engaging task for participants than Naming. In the questionnaires, the Reaching participants reported a stronger feeling of agency on the sound sources. In turn, this could have led to a stronger motivation when performing the task (Aarhus, Grönvall, Larsen, & Wollsen, 2011). The Reaching group also perceived the experience as more rewarding overall, and expressed more interest in the task compared the Naming group. Recall that the participants in the Naming group did not directly interrupted the sound, as it was the experimenter that entered the response. Future works should make this aspect more comparable between groups, and assess to what extent motivation and sense of agency per se could be at the origin of the different head-movement behaviour.

4.4. Conclusions

In this article we show that reaching to sounds can promote fast adaptation to altered binaural cues, more than naming sounds. In addition, we document that listening with one ear plugged triggers compensatory head-movement behaviour, which in turn can modulate sound localization performance. The head-

movement behaviour we observed can be interpreted as a strategic (rather than cognitive) adaptation mechanism to altered spatial hearing. However, cognitive and strategic contributions to sound localisation improvements are clearly not mutually exclusive and, in fact, it is plausible that head-related behavioural strategies may trigger cognitive changes. The implications of these results are two-fold. First, they show the importance of considering dynamic listening behaviour (reaching to sounds and head-movements) when studying how humans adapt to altered auditory cues. This is now increasingly possible thanks to virtual reality and kinematic tracking approaches, like the one we adopted in the present study and other recent works (Parseihian, Jouffrais & Katz, 2014; Parseihian & Katz, 2012; Steadman, Kim, Lestang, Goodman & Picinali, 2019; Stitt, Picinali & Katz, 2019). More generally, researchers should pursue the novel opportunity to study spatial hearing in more realistic and ecological environments (see Hadley et al., 2019). Second, our findings could pave the way for new approaches to the rehabilitation of spatial hearing difficulties in people suffering from hearing loss or using hearing-enabling devices such as hearing aids or cochlear implants.

ACKNOWLEDGEMENTS

C.V. was supported by a grant of the Università Italo-Francese (UIF)/Université Franco-Italienne (UFI) and the Ermenegildo Zegna Founder's Scholarship. F.P. was supported by a grant of the Agence Nationale de la Recherche (ANR-16-CE17-0016, VIRTUALHEARING3D, France) and by a prize of the Fondation Medisite (France) by the Neurodis Foundation (France) and by a grant from the Italian Ministry for Research and University (MUR, PRIN 20177894ZH). We thank Francesca Guglielmi for helping in collecting data and Giordana Torresani for graphic support (Figure 1A).

All figures will be black and white.

REFERENCES

- Aarhus, R., Grönvall, E., Larsen, S. B., & Wollsen, S. (2011). Turning training into play: Embodied gaming, seniors, physical training and motivation. *Gerontechnology*, 10(2), 110-120. <https://doi.org/10.4017/gt.2011.10.2.005.00>;
- Agterberg, M.J., Snik, A.F., Hol, M.K., Van Wanrooij, M.M., Van Opstal, A.J., 2012. Contribution of monaural and binaural cues to sound localization in listeners with acquired unilateral conductive hearing loss: improved directional hearing with a bone-conduction device. *Hear. Res.* 286, 9e18. <https://doi.org/10.1016/j.heares.2012.02.012>;
- Angell, J. R., & Fite, W. (1901). From the Psychological Laboratory of the University of Chicago: The monaural localization of sound. *Psychological Review*, 8(3), 225–246. <https://doi.org/10.1037/h0073690>
- Bahill, A. T., & McDonald, J. D. (1983). Frequency limitations and optimal step size for the two-point central difference derivative algorithm with applications to human eye movement data. *IEEE Transactions on Biomedical Engineering*, (3), 191-194. <https://doi.org/10.1109/TBME.1983.325108>;
- Bauer, R. W., Matuzsa, J. L., Blackmer, R. F., & Glucksberg, S. (1966). Noise Localisation after Unilateral Attenuation. *The Journal of the Acoustical Society of America*, 40(2), 441–444. <https://doi.org/10.1121/1.1910093>;
- Biguer, B., Prablanc, C., & Jeannerod, M. (1984). The contribution of coordinated eye and head movements in hand pointing accuracy. *Experimental brain research*, 55(3), 462-469. <https://doi.org/10.1007/BF00235277>;
- Brimijoin, W. O., & Akeroyd, M. A. (2012). The role of head movements and signal spectrum in an auditory front/back illusion. *i-Perception*, 3(3), 179-182. <https://doi.org/10.1068/i7173sas>;
- Brimijoin, W. O., McShefferty, D., & Akeroyd, M. A. (2010). Auditory and visual orienting responses in listeners with and without hearing-impairment. *The Journal of the Acoustical Society of America*, 127(6), 3678-3688. <https://doi.org/10.1121/1.3409488>;
- Carlile, S. (2014). The plastic ear and perceptual relearning in auditory spatial perception. *Frontiers in neuroscience*, 8, 237. <https://doi.org/10.3389/fnins.2014.00237>;
- Carlile, S., Balachandar, K., & Kelly, H. (2014). Accommodating to new ears: The effects of sensory and sensory-motor feedback. *The Journal of the Acoustical Society of America*, 135(4), 2002–2011. <https://doi.org/10.1121/1.4868369>;
- Carlile, S., Martin, R., & McAnally, K. (2005). Spectral Information in Sound Localization. *International Review of Neurobiology*, 70, 399–434. [https://doi.org/10.1016/S0074-7742\(05\)70012-X](https://doi.org/10.1016/S0074-7742(05)70012-X)
- Clifton, R. K., Gwiazda, J., Bauer, J. A., Clarkson, M. G., & Held, R. M. (1988). Growth in head size during infancy: Implications for sound localisation. *Developmental Psychology*, 24(4), 477. <http://dx.doi.org/10.1037/0012-1649.24.4.477>;
- Cohen, Y. E., & Andersen, R. A. (2002). A common reference frame for movement plans in the posterior parietal cortex. *Nature Reviews Neuroscience*, 3(7), 553. <https://doi.org/10.1038/nrn873>;
- Cranford, J. L., Andres, M. A., Piatz, K. K., & Reissig, K. L. (1993). Influences of age and hearing loss on the precedence effect in sound localisation. *Journal of Speech, Language, and Hearing Research*, 36(2), 437-441. <https://doi.org/10.1044/jshr.3602.437>;

- Dobрева, M. S., O'Neill, W. E., & Paige, G. D. (2011). Influence of aging on human sound localisation. *Journal of neurophysiology*, *105*(5), 2471-2486. <https://doi.org/10.1152/jn.00951.2010>;
- Eid, M., Gollwitzer, M., & Schmitt, M. (2011). Statistik und Forschungsmethoden. Lehrbuch (2., korr. Aufl.). *Grundlagen Psychologie*;
- Fogt, N., Uhlig, R., Thach, D. P., & Liu, A. (2002). The influence of head movement on the accuracy of a rapid pointing task. *Optometry (St. Louis, Mo.)*, *73*(11), 665-673;
- Francesco Pavani, Valerie Gaveau, Anael Belle, Eric Koun, Romeo Salemme, Clement Desoche, Eric Truy, Alessandro Farnè (2017). Active sound localisation in 3D space in unilateral and bilateral cochlear implant users: a new approach to spatial hearing assessment through immersive virtual reality. *International Federation of ORL Societies (IFOS)*, Paris, June 24-27th;
- Fuller, J. H. (1992). Head movement propensity. *Experimental Brain Research*, *92*(1), 152-164. <https://doi.org/10.1007/BF00230391>;
- Gaveau, V., Coudert, A., Salemme, R., Koun, E., Desoche, C., Truy, E., ... & Pavani, F. (2020). SPHERE: A novel approach to 3D and active sound localization. *bioRxiv*.
- Hadley, L. V., Brimijoin, W. O., & Whitmer, W. M. (2019). Speech, movement, and gaze behaviours during dyadic conversation in noise. *Scientific reports*, *9*(1), 1-8. <https://doi.org/10.1038/s41598-019-46416-0>;
- Hendrickx, E., Stitt, P., Messonnier, J. C., Lyzwa, J. M., Katz, B. F., & De Boishéraud, C. (2017). Influence of head tracking on the externalization of speech stimuli for non-individualized binaural synthesis. *The Journal of the Acoustical Society of America*, *141*(3), 2011-2023. <https://doi.org/10.1121/1.4978612>
- Honda, A., Shibata, H., Gyoba, J., Saitou, K., Iwaya, Y., & Suzuki, Y. (2007). Transfer effects on sound localisation performances from playing a virtual three-dimensional auditory game. *Applied Acoustics*, *68*(8), 885–896. <https://doi.org/10.1016/J.APACOUST.2006.08.007>;
- Honda, A., Shibata, H., Hidaka, S., Gyoba, J., Iwaya, Y., & Suzuki, Y. (2013). Effects of head movement and proprioceptive feedback in training of sound localization. *i-Perception*, *4*(4), 253-264. <https://doi.org/10.1068/i0522>;
- 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>;
- Isaiah, A., Vongpaisal, T., King, A. J., & Hartley, D. E. (2014). Multisensory training improves auditory spatial processing following bilateral cochlear implantation. *Journal of Neuroscience*, *34*(33), 11119-11130. <https://doi.org/10.1523/JNEUROSCI.4767-13.2014>;
- Iwaya, Y., Suzuki, Y., & Kimura, D. (2003). Effects of head movement on front-back error in sound localisation. *Acoustical science and technology*, *24*(5), 322-324. <https://doi.org/10.1250/ast.24.322>;
- Kacelnik, O., Nodal, F. R., Parsons, C. H., & King, A. J. (2006). Training-induced plasticity of auditory localisation in adult mammals. *PLoS biology*, *4*(4), e71. <https://doi.org/10.1371/journal.pbio.0040071>;
- Keating, P., & King, A. J. (2015). Sound localisation in a changing world. *Current opinion in neurobiology*, *35*, 35-43. <https://doi.org/10.1016/j.conb.2015.06.005>;
- Kim, J., Barnett-Cowan, M., & Macpherson, E. A. (2013). Integration of auditory input with vestibular and neck proprioceptive information in the interpretation of dynamic sound localisation cues.

- In *Proceedings of meetings on acoustics ICA2013*, 19, (1), 050142). ASA. <https://doi.org/10.1121/1.4799748>;
- Knudsen, E. I., & Knudsen, P. F. (1985). Vision guides the adjustment of auditory localisation in young barn owls. *Science*, 230(4725), 545-548. <https://doi.org/10.1126/science.4048948>;
- Kumpik, D. P., Kacelnik, O., & King, A. J. (2010). Adaptive reweighting of auditory localisation cues in response to chronic unilateral earplugging in humans. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 30(14), 4883–4894. <https://doi.org/10.1523/JNEUROSCI.5488-09.2010>;
- 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>
- Majdak, P., Goupell, M. J., & Laback, B. (2010). 3-D localisation of virtual sound sources: effects of visual environment, pointing method, and training. *Attention, perception, & psychophysics*, 72(2), 454-469.; <https://doi.org/10.3758/APP.72.2.454>;
- Majdak, P., Walder, T., & Laback, B. (2013). Effect of long-term training on sound localisation performance with spectrally warped and band-limited head-related transfer functions. *The Journal of the Acoustical Society of America*, 134(3), 2148–2159. <http://asa.scitation.org/doi/10.1121/1.4816543>;
- McAnally, K. I., & Martin, R. L. (2014). Sound localisation with head movement: implications for 3-d audio displays. *Frontiers in neuroscience*, 8, 210. <https://doi.org/10.3389/fnins.2014.00210>;
- Mendonça, C. (2014). A review on auditory space adaptations to altered head-related cues. *Frontiers in neuroscience*, 8, 219. <https://doi.org/10.3389/fnins.2014.00219>;
- Mendonça, C., Campos, G., Dias, P., & Santos, J. A. (2013). Learning auditory space: Generalization and long-term effects. *PloS one*, 8(10), e77900. <https://doi.org/10.1371/journal.pone.0077900>;
- Middlebrooks, J. C. (2015b). Sound localisation. *Handbook of Clinical Neurology* (1st ed., Vol. 129). Elsevier B.V. <https://doi.org/10.1016/B978-0-444-62630-1.00006-8>;
- Middlebrooks, J. C., & Green, D. M. (1991). Sound localisation by human listeners. *Annual review of psychology*, 42(1), 135-159. <https://doi.org/10.1146/annurev.ps.42.020191.001031>;
- Morikawa, D., & Hirahara, T. (2013). Effect of head rotation on horizontal and median sound localisation of band-limited noise. *Acoustical Science and Technology*, 34(1), 56-58. <https://doi.org/10.1250/ast.34.56>;
- Morikawa, D., Toyoda, Y., & Hirahara, T. (2011, July). Impact of head movement on sound localization with band-limited noise. In *INTER-NOISE and NOISE-CON Congress and Conference Proceedings* (Vol. 2011, No. 6, pp. 1963-1968). Institute of Noise Control Engineering;
- Muir, D., & Field, J. (1979). Newborn infants orient to sounds. *Child development*, 431-436;
- Nawaz, S., McNeill, C., & Greenberg, S. L. (2014). Improving sound localisation after cochlear implantation and auditory training for the management of single-sided deafness. *Otology & Neurotology*, 35(2), 271-276. <https://doi.org/10.1097/MAO.0000000000000257>;
- Noble, W. G. (1981). Earmuffs, exploratory head movements, and horizontal and vertical sound localisation. *Journal of Auditory Research*;

- Ohuchi, M., Iwaya, Y., Suzuki, Y., & Munekata, T. (2005). Training effect of a virtual auditory game on sound localisation ability of the visually impaired. Georgia Institute of Technology; <http://hdl.handle.net/1853/50191>;
- Parseihian, G., Jouffrais, C., & Katz, B. F. (2014). Reaching nearby sources: comparison between real and virtual sound and visual targets. *Frontiers in neuroscience*, 8, 269. doi: 10.3389/fnins.2014.00269;
- Parseihian, G., & Katz, B. F. G. (2012). Rapid head-related transfer function adaptation using a virtual auditory environment. *The Journal of the Acoustical Society of America*, 131(4), 2948–2957. <https://doi.org/10.1121/1.3687448>;
- Perrett, S., & Noble, W. (1997). The contribution of head motion cues to localisation of low-pass noise. *Perception & Psychophysics*, 59(7), 1018–1026. <https://doi.org/10.3758/BF03205517>;
- Poirier-Quinot, D., & Katz, B. F. (2018, August). Impact of HRTF individualization on player performance in a VR shooter game II. In *Audio Engineering Society Conference: 2018 AES International Conference on Audio for Virtual and Augmented Reality*. Audio Engineering Society.
- Rabini, G., Altobelli, E., & Pavani, F. (2019). Interactions between egocentric and allocentric spatial coding of sounds revealed by a multisensory learning paradigm. *Scientific reports*, 9(1), 7892. <https://doi.org/10.1038/s41598-019-44267-3>;
- Rayleigh, Lord. (1907). On our perception of sound direction. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 13(74), 214–232. <https://doi.org/10.1080/14786440709463595>
- Shinn-Cunningham, B. G., Durlach, N. I., & Held, R. M. (1998). Adapting to supernormal auditory localisation cues. I. Bias and resolution. *The Journal of the Acoustical Society of America*, 103(6), 3656. <https://doi.org/10.1121/1.423088>;
- Shinn-Cunningham, B. G., Santarelli, S., & Kopco, N. (2000). Tori of confusion: Binaural localization cues for sources within reach of a listener. *The Journal of the Acoustical Society of America*, 107(3), 1627. <https://doi.org/10.1121/1.428447>
- Steadman, M. A., Kim, C., Lestang, J. H., Goodman, D. F., & Picinali, L. (2019). Short-term effects of sound localisation training in virtual reality. *Scientific Reports*, 9(1), 1-17. <https://doi.org/10.1038/s41598-019-54811-w>;
- Stitt, P., Picinali, L., & Katz, B. F. (2019). Auditory accommodation to poorly matched non-individual spectral localization cues through active learning. *Scientific Reports*, 9(1), 1-14, <https://doi.org/10.1038/s41598-018-37873-0>.
- Strelnikov, K., Rosito, M., & Barone, P. (2011). Effect of audiovisual training on monaural spatial hearing in horizontal plane. *PLoS one*, 6(3), e18344. <https://doi.org/10.1371/journal.pone.0018344>;
- Thurlow, W. R., & Runge, P. S. (1967). Effect of Induced Head Movements on Localisation of Direction of Sounds. *The Journal of the Acoustical Society of America*, 42(2), 480–488. <https://doi.org/10.1121/1.1910604>;
- Thurlow, W. R., Mangels, J. W., & Runge, P. S. (1967). Head movements during sound localization. *The Journal of the Acoustical society of America*, 42(2), 489-493. <https://doi.org/10.1121/1.1910605>;

- 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>;
- Valzolgher, C., Campus, C., Rabini, G., Gori, M. & Pavani, F. (accepted for publication). Updating spatial hearing abilities through multisensory and motor cues. *Cognition*;
- van der Meer, A. L., Ramstad, M., & van der Weel, F. R. (2008). Choosing the shortest way to mum: Auditory guided rotation in 6-to 9-month-old infants. *Infant Behavior and Development*, 31(2), 207-216. <https://doi.org/10.1016/j.infbeh.2007.10.007>;
- van Wanrooij, M. M., John, A., & Opstal, V. (2005). Relearning Sound Localisation with a New Ear. *The Journal of Neuroscience*, 25(22), 5413–5424. <https://doi.org/10.1523/JNEUROSCI.0850-05.2005>;
- Vercher, J. L., Magenes, G., Prablanc, C., & Gauthier, G. M. (1994). Eye-head-hand coordination in pointing at visual targets: spatial and temporal analysis. *Experimental brain research*, 99(3), 507-523. <https://doi.org/10.1007/BF00228987>;
- Verdelet, G., Desoche, C., Volland, F., Farnè, A., Coudert, A., Hermann, R., ... & Salemme, R. (2019, December). Assessing spatial and temporal reliability of the vive System as a tool for naturalistic behavioural research. In 2019 International Conference on 3D Immersion (IC3D) (pp. 1-8). IEEE.
- Wallach, H. (1940). The role of head movements and vestibular and visual cues in sound localisation. *Journal of Experimental Psychology*, 27(4), 339. <https://doi.org/10.1037/h0054629>;
- Wiebe, E. N., Lamb, A., Hardy, M., & Sharek, D. (2014). Measuring engagement in video game-based environments: Investigation of the User Engagement Scale. *Computers in Human Behavior*, 32, 123-132. <https://doi.org/10.1016/j.chb.2013.12.001>;
- Wightman, F. L., & Kistler, D. J. (1999). Resolution of front–back ambiguity in spatial hearing by listener and source movement. *The Journal of the Acoustical Society of America*, 105(5), 2841-2853. <https://doi.org/10.1121/1.426899>;
- World Medical Association. (2013). *WMA declaration of Helsinki: Ethical principles for medical research involving human subjects*. Retrieved from <https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/>;
- Zahorik, P., Bangayan, P., Sundareswaran, V., Wang, K., & Tam, C. (2006). Perceptual recalibration in human sound localisation: learning to remediate front-back reversals. *The Journal of the Acoustical Society of America*, 120(1), 343–359. <https://doi.org/10.1121/1.2208429>

HIGHLIGHTS

- We studied adaptation to monaural listening when listeners interact with sounds
- Reaching to sounds reduced errors more than just naming the sound sources
- Patterns of spontaneous head-movements changed during adaptation
- Exploration of a wider portion of space with the head was functional for adaptation
- Reaching to sounds promoted a more effective head-related behaviour

Journal Pre-proof



The impact of a visual spatial frame on real sound-source localization in virtual reality

Chiara Valzolgher^{a,b,*}, Mariam Alzhaler^d, Elena Gessa^c, Michela Todeschini^c, Pauline Nieto^d, Gregoire Verdelet^{a,g}, Romeo Salemme^{a,g}, Valerie Gaveau^{a,e}, Mathieu Marx^h, Eric Truy^f, Pascal Barone^d, Alessandro Farnè^{a,b,g}, Francesco Pavani^{a,b,c}

^a Integrative, Multisensory, Perception, Action and Cognition Team (IMPACT), Lyon Neuroscience Research Center, France

^b Center for Mind/Brain Sciences (CIMEC), University of Trento, Italy

^c Department of Psychology and Cognitive Sciences (DiPSCo), University of Trento, Italy

^d Centre de Recherche Cerveau & Cognition, Toulouse, France

^e University of Lyon 1, France

^f Hospices Civils de Lyon, Lyon, France

^g Neuroimmersion, Lyon, France

^h Hospices Civils, Toulouse, France

ARTICLE INFO

Keywords:

Spatial hearing
Virtual reality
Visual information
Active listening
Head movements tracking
Gaze tracking

ABSTRACT

Studies on audio-visual interactions in sound localization have primarily focused on the relations between the spatial position of sounds and their perceived visual source, as in the famous ventriloquist effect. Much less work has examined the effects on sound localization of seeing aspects of the visual environment. In this study, we took advantage of an innovative method for the study of spatial hearing – based on real sounds, virtual reality and real-time kinematic tracking – to examine the impact of a minimal visual spatial frame on sound localization. We tested sound localization in normal hearing participants (N=36) in two visual conditions: a uniform gray scene and a simple visual environment comprising only a grid. In both cases, no visual cues about the sound sources were provided. During and after sound emission, participants were free to move their head and eyes without restriction. We found that the presence of a visual spatial frame improved hand-pointing in elevation. In addition, it determined faster first-gaze movements to sounds. Our findings show that sound localization benefits from the presence of a minimal visual spatial frame and confirm the importance of combining kinematic tracking and virtual reality when aiming to reveal the multisensory and motor contributions to spatial-hearing abilities.

1. Introduction

In humans, as well as in other animals that can hear, the ability to localize sounds in space has evolved over the years within a multisensory environment. Under this ecological pressure, spatial hearing co-evolved with other sensory systems such as vision, which provides distal information about the environment (Heffner and Heffner, 1992, Heffner and Heffner, 2014). In addition, studies in animal and human models clearly showed that vision plays a critical role in the development of acoustic space perception (Knudsen and Knudsen, 1985, Hofman et al., 1998). When a listener is engaged in a sound localization task, there are at least two ways in which vision can contribute useful spatial information. First, vision can provide *direct* information about the auditory target, by revealing the position of the sound source in the environment

(e.g., the listener hears and sees the bird tweeting on the tree). Second, vision can provide *indirect* information about the auditory targets, by revealing from which sector of space they may originate or by providing general information about the environmental spatial frame for encoding sound position (e.g., the listener cannot see the bird tweeting, but perceives the tree branches from which the stimulus originates).

The vast majority of studies that investigated audio-visual interactions in spatial hearing have been carried out in a context in which vision provides direct visual cues about sound position. In the typical experiment of this sort, the onset of the target sound is accompanied by a visual cue. When the visual information is veridical, listeners are more precise in sound localization compared to when no information is provided (Shelton and Searle, 1980, Tonelli et al., 2015). Instead, when the visual information is not veridical, visual-capture of sound position typically

* Corresponding author. Present address: Integrative, Multisensory, Perception, Action and Cognition Team (IMPACT), Centre de Recherche en Neurosciences de Lyon Inserm U1028 - CNRS UMR5292 Bâtiment Inserm 16 avenue Doyen Lépine 69676 BRON Cedex.

E-mail addresses: chiara.valzolgher@unitn.it, chiara.valzolgher@inserm.fr (C. Valzolgher).

<https://doi.org/10.1016/j.crbeha.2020.100003>

Received 12 August 2020; Received in revised form 19 October 2020; Accepted 10 November 2020

2666-5182/© 2020 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>)

emerges (the well-known ‘ventriloquist effect’) (Alais and Burr, 2004). For instance, Bolognini et al. (2007) demonstrated that veridical visual cues can enhance sound localization. Participants sat in front of a plastic semicircular apparatus which comprised eight loudspeakers hidden behind a curtain. They were required to verbally judge sound location, reading aloud labels marking the position of each speaker. Crucially, the auditory stimulus was either presented alone or together with the visual stimulus. Results showed that spatially and temporally coincident visual stimuli improved sound localization accuracy. Consistent results have been observed also in infants (Morrongiello and Rocca, 1987). Moreover, these direct visual contributions to sound localization have been proved useful when training acoustic space perception (Rabini et al., 2019, Strelnikov et al., 2011, Valzolgher et al., 2020).

Considerably less research has instead investigated indirect visual contributions to sound localization. Yet, the idea that vision of the surrounding environment can provide useful spatial information for spatial hearing dates back to the 1970s, when it was termed ‘visual facilitation’ (Warren, 1970). Warren was among the first to report that sound localization accuracy can improve when listeners localize unseen auditory targets with their eyes open than closed. In eye-open conditions, speakers were hidden from the subject using a fabric screen, but the overall environment was clearly visible. When interpreting the advantage observed in eye-open conditions, Warren proposed that participants use visual cues from the environment to place the auditory stimulus into a visual spatial representation, instead of using only an auditory frame of reference. In line with this early observation, a decade later Shelton and Searle (1980) showed that when a speaker array is placed directly in front of the subjects, vision of the sources can facilitate accuracy compared to a blindfolded condition – even when the exact position of the target sound remains unknown. More recent works have replicated these pioneering studies, suggesting that seeing the environment can also enhance the precision of motor response and may thus affect participants’ performance by facilitating their motor behavior. For instance, Redon and Hay (2005) showed that the presence of a visually structured background reduces pointing bias to visual targets. Interestingly, even the brief observation of the overall environment can improve spatial hearing (Tonelli et al., 2015).

Yet, in these classic studies as well as in more recent ones, it was difficult to disambiguate indirect contributions to sound localization that resulted from seeing the overall structure of the environment (e.g., a visible room), from the contributions resulting from seeing the possible space occupied by the sound sources (e.g., a panel hiding sound sources placed in front of the participant). In the first case, participants could code sound position with respect to existing references; in the second case, participants can develop some sort of visual prior about the position of the sounds in the environment (Parise et al., 2014). It is worth noting that in most studies exploring spatial hearing, the position of the sound sources is either directly visible or can be easily inferred. For instance, when all sources are hidden behind a curtain participants can nonetheless infer that sounds can originate from a restricted portion of the space (Pavani et al., 2017). In this case, although participants have no detailed prior about the spatial layout of the speakers, they have continuous visual priors about the hemisphere (front or back), elevation and distance of the speaker array.

One way to disentangle between these indirect visual contributions to sound localization is to exploit virtual reality. In a recent study, Ahrens et al. (2019) asked participants to perform a sound localization task in different visual scenarios created using virtual reality technology. In some conditions, visual information about the room was entirely prevented. In other conditions, participants were allowed to see a virtual version of the real room, comprising or not the speakers around them. Using VR, these authors were able to control the effect of both having the *structure of the overall auditory environment* and knowing the *spatial likelihood* of the auditory targets. They found that the reference frame provided by the visual information of the room without loudspeaker was enough to decrease error both in the horizontal and verti-

cal dimensions compared to a blindfolded condition. Moreover, vision of the speaker array provided a further benefit compared to receiving only visual information of the room. Along a similar line, Majdak et al. (2010) have manipulated the whole visual background during a sound localization task to study the impact of seeing a simple visual spatial frame (a grid) compared to a condition of total darkness. The results showed that even a simple grid improved precision in the horizontal plane. Furthermore, the visible grid reduced quadrant errors in the vertical plane, particularly in the front hemisphere. In this work, sounds were always delivered using HRTF making more difficult to study the distance dimension. Likewise, the visible grid lacked binocular cues and did not provide information about distance. Finally, participants were forced into a static head listening posture. While this posture is most common in sound localization studies, it is a constraint that may limit the benefit from the visible visual environment on sounds localization. As proposed by some authors (Shelton and Searle, 1980), the spatial hearing facilitation that can result from seeing the overall environment, may reflect the active sensory-motor exploration of the auditory scene.

The present study aimed to test the contribution of a visual spatial frame to sound localization when participants are free to move their head without restriction. To this aim, we tested participants in two visual conditions: a uniform gray scene, in which no cues about the sound sources or the auditory environment were provided, and a simple visual spatial frame, in which a visible grid was the only visual information available about the environment. To present the different visual scenarios and to allow recording of spontaneous head and gaze movements, we took advantage of a new approach for the study of spatial hearing developed in our laboratories (Verdelet et al., 2020, Gaveau et al., 2020). Our approach is based on real sounds, virtual reality and real-time kinematic tracking and it allows: (1) accurate positioning of real sounds at pre-determined locations with respect to the head; (2) measuring the participant’s hand responses in the three dimensions of space (i.e., in 3D); (3) control over the visual scenarios and (4) free and measurable head-movements and gaze during sounds playback. The latter aspect is particularly advantageous for the present study, as we hypothesized that head and eyes orienting to the sounds could provide implicit measures of performance to further investigate the effect of visual manipulation. A secondary purpose of the present work was also to validate the feasibility of our VR approach for the study of sound localization abilities in humans.

2. Methods

2.1. Participants

Thirty-six participants (age: $M = 25.08$, $SD = 2.96$, range [20-32], 13 males, 34 right-handed) were recruited to participate in the experiment at the University of Trento (Italy), at the Integrative, Multisensory, Perception, Action and Cognition Team (IMPACT) lab in Lyon and at the center de Recherche Cerveau et Cognition (CerCo) of Toulouse (France). All participants signed an informed consent before starting the experiment, which was conducted according to the criteria of the Declaration of Helsinki (1964, amended in 2013) and approved by the respective ethical committees. All had normal or corrected-to-normal vision and reported no movement deficit. Hearing thresholds were measured using an audiometer for all participants, testing different frequencies (250, 500, 1000, 2000, 4000 Hz), on the right and left ear separately. All participants had an average threshold below 20 dB HL.

2.2. Apparatus and stimuli

Virtual reality (VR) and kinematic tracking was implemented using 3 identical HTC Vive Systems, one for each testing site. This method

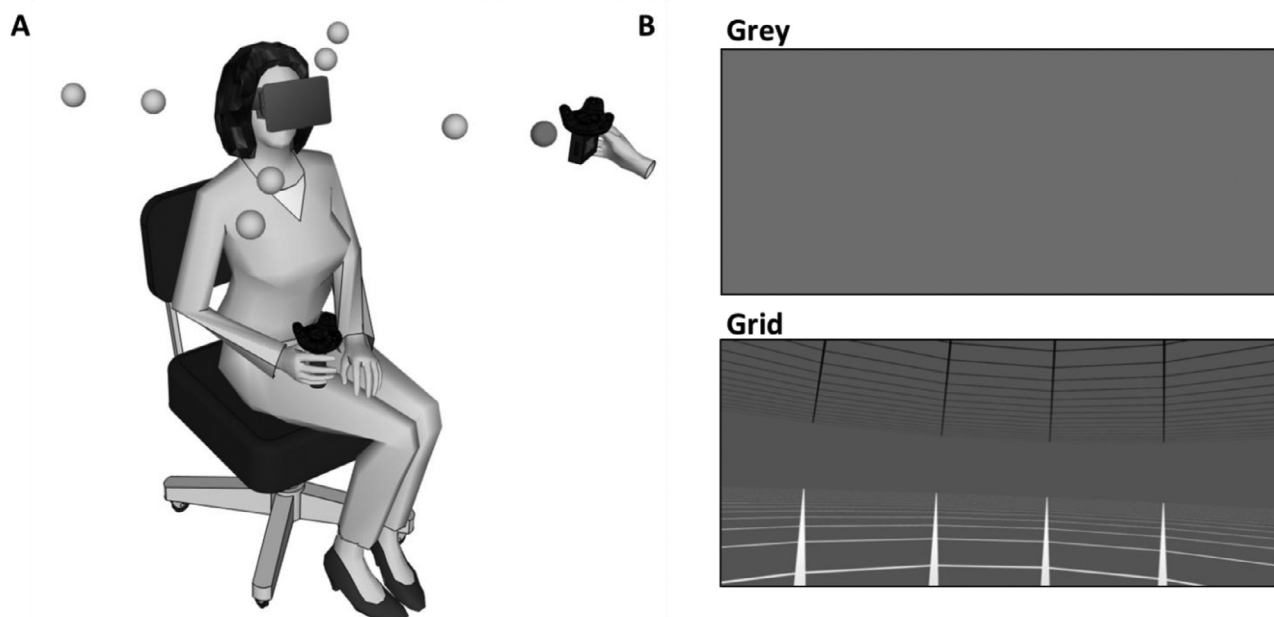


Fig. 1. Experimental procedure and setting. (A) Schematic representation of participant wearing the HMD and holding the tracker used for pointing responses; the eight spheres around the participant's head indicate the pre-determined speaker positions; the experimenter brought the tracked speaker (also shown in figure) at the pre-determined location identified in each trial. (B) Representation of the two 3D visual environments used in the study: gray and grid. A video showing dynamically the grid condition as seen from the participant's perspective is available at: <https://youtu.be/89xPLzr3fyQ>.

(European patent n°17723294.6–1115) has been developed in our laboratory (see also Gaveau et al., 2020 and Valzolgher et al. 2020). Each system (Fig. 1A) comprised one head-mounted display (HMD, resolution: 1080×1200 px, Field Of View (FOV): 110° , Refresh rate: 90 Hz), 1 controller (used by experimenter to calibrate head-position and to interrupt trial data registration), 2 trackers (one mounted on a short rod and held by participants to indicate the position of the sound and the other mounted above the speaker to track its position in real time) and 2 lighthouse base stations (scanning the position of the controller and trackers). Tracking precision and accuracy of the HTC Vive System is adequate for behavioral research purposes (Ahrens et al., 2019). The HMD was equipped with an SMI eye-tracking system (250 Hz). At all testing sites, stimuli were controlled and delivered using a LDLC ZALMAN PC (OS: Windows 10 (10.0.0) 64bit; Graphic card: NVIDIA GeForce GTX 1060 6GB; Processor: Intel Core i7–7700 K, Quad-Core 4.2 GHz/4.5 GHz Turbo - Cache 8 Mo - TDP 95 W) using Steam VR software and the development platform Unity3D (Unity Technologies, San Francisco, CA).

Participants were seated on a rotating armless chair with no chin rest, in the center of the room. The rooms had the following dimensions: Trento: 4×3 m; Lyon: 3×6 m; Toulouse: 3×5 m. All rooms were quiet, but none was specifically treated for being anechoic and sound-proof.

Real free field auditory stimuli were delivered by an unseen loudspeaker (JBL GO Portable, $68.3 \times 82.7 \times 30.8$ mm, Output Power 3.0 W, Frequency response 180 Hz – 20 kHz, Signal-to-noise ration > 80 dB), whose position was continuously tracked in space. They were 3 s white noise bursts, amplitude-modulated at 4 Hz, and delivered at about 60 dB SPL, as measured from the participant's head (using Decibel Meter TES1350A placed at ears level).

This solution allowed us to track the position of the speaker, the hand of participant and the Head Mounted Display, via a dead reckoning process using gyroscope and accelerometer, plus a correction signal from the lighthouse system every 8.333 ms. Both tracking method allowed us to track this position with a frequency sample of 250 Hz. The software is designed to guide the experimenter to align the real loudspeaker (the sound source) with each of the 8 pre-determined position in the virtual environment in each trial.

2.3. Procedure

Before starting the experiment, participants were instructed about the task and informed about the use of the VR equipment. Specifically, participants were asked to listen carefully to each sound until it finished, and then to indicate its location in space using the tracker held in their right hand. They were informed that sounds could be delivered anywhere in the 3D around them, always within a reaching space. They were also informed that, during sound emission, they were not allowed to move their hand (which rested on their laps) nor their trunk (which was oriented straight ahead). After sound emission, they could freely rotate the chair to indicate the sound source. Note that head-movements remain unconstrained both during and after sound emission, allowing the possibility of spontaneous active listening behavior (e.g., orienting the head in the direction of the sound).

Participants performed sound localization under two visual conditions: a uniform gray scene (gray) and a more structured scene with spatial references (grid) (Fig. 1B and Video). The grid comprised two horizontally laid figures, drawn like spiderwebs of 50-meters radius, with 19 straight sides (angle around 20°) and 20 sub-figures plan separated by 2.5 m. The first horizontal web was placed at floor level ($Y = 0$ m) and the second was placed at 10 m height ($Y = 10$ m). In creating the grid, we aimed to obtain a structured environment, with distance clue that conveyed the idea of a vast and open space (hence, the 50-meter radius of the spiderwebs). In addition, we aimed to avoid any vertical line that could be used by participants as visual anchor during sound perception and/or during the pointing response. The overall grids were centered on the participant's position. A video showing dynamically the grid condition as seen from the participant's perspective is available at: <https://youtu.be/89xPLzr3fyQ> (note that the video does not represent an experimental trial, but only depicts the environment visible to the participants).

After the participant sat on the chair in the experimental room and wore the HMD, the experiment began the head-center calibration which was performed by collecting the 3D position of the two ears using the controller. These head-center coordinates served as origin of the polar frame of reference that defined speaker, head and gaze positions

throughout the experiment. Then, the eye-tracker calibration was performed: participants were asked to follow a moving dot with their eyes. Both head and eye calibrations were repeated whenever the HMD was temporarily removed during the experiment (i.e., during pauses).

The loudspeaker position in 3D space was calculated for each trial, with reference to the center of the head. In this way, despite participants sat without any chin-rest, we could carefully control the position of each sound source with respect to the ears. Eight pre-determined positions were used throughout the experiment, resulting from the combination of 4 different azimuths (-45° , 45° , -135° or 135°), 2 different distances (35 cm or 55 cm) and a single elevation (0° , i.e., ear-level) (Fig. 1).

In each experimental trial, the experimenter moved the loudspeaker to the desired position in 3D space, following visual instruction generated in real-time by the computer. Instructions conveyed the pre-determined azimuth and distance position for the speaker. These instructions were visible only to the experimenter, and they were delivered using the stimulus visualization monitor (IIYAMA PROLITE E2280HS 22", resolution: 1980×1080 , format 16:9) placed in the testing room. The monitor provided a bird-eye view of the experimental room and conveyed the pre-determined position of the sound source for that trial (as a red circle) and the actual real-time position of the speaker (as tracker picture). Using this visual instruction, the experimenter reached for computer-determined position rapidly, keeping the speaker approximately at ear-level. The computer considered the loudspeaker correctly positioned when it entered a sphere of 6 cm diameter centered on the pre-determined sound position.

The noises produced by the experimenter while placing the speaker were minimal. Nonetheless, participants were explicitly informed to pay attention to the target sounds, as any other sound in the room could be deceiving. To prove this point at the beginning of the testing session, the experimenter showed how she could stay to the right of participants while delivering the sound from their left by stretching the arm and hand holding the speaker. This was a demonstration that occasional noise cues about the movement of the experimenter could provide misleading information and thus discourage participants to rely upon this information. Most importantly, pilot work in our laboratory showed that participants ($N=6$) cannot reliably point to the speaker when the same placement procedure is used but no target sound is delivered. In this scenario the three-dimensional vector distance between the speaker and the response is on average 73.3 cm ($SD = 31.7$ cm).

The computer delivered the target sound only when three concurrent criteria were met: (1) the loudspeaker was in the 3D position pre-determined for the trial; (2) the participant's head was facing straight ahead; (3) the participant's eyes were directed straight ahead. Participants complied with criterion 2 (head pointing straight ahead) and criterion 3 (eyes gazing straight ahead) by taking advantage of visual stimuli displayed in the HMD. At the beginning of each trial two crosses were presented to the participant: a white cross in the background indicated the desired position of the head and eyes; and a thin white cross, indicating the actual head-position of the participant. Participants were instructed to move their head to align the two crosses. When the alignment was achieved the thin cross turned blue. Likewise, participants were instructed to stare at the cross center, a feedback of their gaze location was given by a blue circle. Once the three criteria were achieved simultaneously (speaker position; head position and eye gaze), all visual feedback disappeared, and the sound was delivered. Participants were instructed to respond only after the end of the sound, bringing the tracker to the perceived location of the sound and holding it still a few seconds. The experimenter terminated the registration of the tracking by pressing a button on the controller. No feedback on performance was provided (similar procedure was used also in Gaveau's study (Gaveau et al., 2020), see also (Valzolgher et al., 2020)).

The experimental session was organized in 4 successive blocks, with a pause between each block. Visual conditions (gray or grid) alternated between blocks of trials. Half of the participants followed a grid-gray-grid sequence, whereas the other half followed a gray-grid-grid-

gray sequence. Each block comprised 40 trials (i.e., 5 trials for each of the 8 pre-determined positions), resulting in a total of 160 trials (i.e., 10 trials for each pre-determined position in each visual condition). The entire experimental protocol lasted approximately 45 min.

2.4. Analyses

The position of all tracked elements (loudspeaker, head center and direction, hand and eyes) was inspected manually for each trial. Loudspeaker position was calculated as the mean of x, y, z coordinates from the beginning of the sound to the end. Head and hand positions were analyzed using custom-made software for the kinematic analysis of movements, running on MATLAB R2019b.

To study head and gaze movements, we calculated the tangential velocity on the x, y, z axis (expressed in degrees of rotation) using two-points central difference derivative algorithm (Terry Bahill and McDonald, 1983) with 5 points for the half-window. To determine the sequence of head and hand movements, the beginning and the end of all movements were automatically detected using a velocity threshold procedure ($10^\circ/s$ for head and 400 mm/s for eyes). The results of this procedure were inspected off-line and corrected manually, if necessary. This procedure served to establish the spatio-temporal profile of head and hand behavior and extract relevant parameters for subsequent analyses (reaction times (RT) of the first head movement and first eyes movement). Head movements below 10° were discarded. Importantly, it also served to reject all trials in which participants did not comply with the instructions (i.e., they made anticipatory hand movement during sound playback) or because of artefacts or lack of data.

All statistical analyses and data visualizations were performed using R, R-studio environment and JASP 0.9.1.0.

3. Results

Our approach to sound localization allowed to decompose localization errors in the three dimensions of space: azimuth, elevation and distance. We started by studying participants' ability to discriminate sound source location in the uniform gray condition – which we considered as baseline – and then we focused on our key experimental question, concerning the effects of seeing a visual grid. This two-step approach was motivated also by our interest to validate our VR approach to spatial hearing which, for several aspects, is novel with respect to other previous methods. Unlike classic works using real sounds, our approach allows free but carefully controlled positioning of the speaker in head-centered coordinates. Moreover, unlike classic auditory virtual reality studies, it uses actual sounds delivered in space and aligned with the visual reality scenario.

3.1. Hand pointing responses: the uniform gray condition

Sound localization responses for each participant are shown in Fig. 2A and B in bird-eye and lateral views, respectively, as a function of sound position. Participants clearly discriminated stimulus side (overall left/right discrimination errors = 1.0%, $SD = 0.7$), and made very few front/back confusions (overall errors = 1.4%, $SD = 0.6$). In addition, as shown in Fig. 2C, responses in distance were clearly segregated for near and far targets (near: $M = 42.7$ cm, $SD = 10.8$ cm; far: $M = 58.6$ cm, $SD = 11.7$ cm; $t(35) = 16.15$, $p < 0.001$).

For each dimension, we studied absolute and signed errors as a function of target position, using an Analysis of Variance (ANOVA) with SIDE (left or right), ANTERO-POSTERIOR SECTOR (front or back) and DISTANCE (near, 35 cm or far, 55 cm) as within-participants factors.

3.1.1. Azimuth

The overall absolute error in azimuth was 11.5° ($SD = 6.5^\circ$). The analysis on absolute error in azimuth revealed a main effect of ANTERO-POSTERIOR SECTOR, $F(1,35) = 6.00$, $p = 0.02$, $\eta^2 = 0.07$, caused by larger

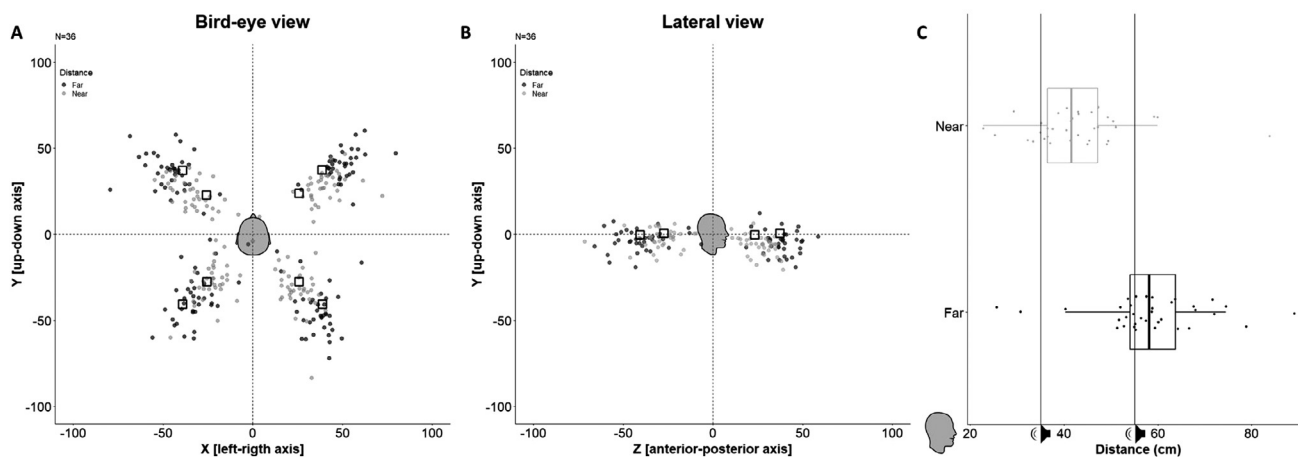


Fig. 2. Sound localization performance. (A) Bird-eye view of all target positions (squares with black border) and hand-pointing responses (smaller circles) in all trials and participants. (B) Lateral view of all target positions and responses. Responses for each participant are averaged across side (left or right) and distance (near, 35 cm and far, 55 cm). Responses are color-coded as a function of target distance (far is dark gray and near is light gray). (C) Distance of participants' hand-pointing as a function of target position (Near or Far). Vertical lines represent the real position of the targets (Near: 35 cm; Far = 55 cm).

errors in back space ($M=13.5^\circ$, $SD=9.1^\circ$) compared to front space ($M=9.4^\circ$, $SD=7.2^\circ$). The main effect of DISTANCE also reached significance, $F(1,35)=13.10$, $p < 0.001$, $\eta^2=0.02$. Errors were larger for near (35 cm) ($M=12.6^\circ$, $SD=7.8^\circ$) compared to far (55 cm) targets ($M=10.4^\circ$, $SD=5.5^\circ$). No other main effect or interaction reached significance (all $F_s < 2.51$, all $p_s > 0.12$).

A similar analysis on signed error (speaker position minus participant's tracker position, positive values indicate a rightward bias) revealed only a main effect of SIDE, $F(1,35)=26.39$, $p < 0.001$, $\eta^2=0.22$. This main effect reflects the fact that participants' pointing responses were more eccentric than actual sound azimuth (left: $M=-5.2^\circ$, $DS=8.2^\circ$; right: $M=5.3^\circ$, $DS=6.4^\circ$) as shown in Fig. 2A. Thus, a positive (rightward) bias emerged for right sounds, where a negative (leftward) bias emerged for left sounds.

3.1.2. Elevation

The overall absolute error in elevation was 13.5° ($SD=5.4^\circ$). The analysis on absolute error in elevation revealed a main effect of ANTERO-POSTERIOR SECTOR, $F(1,35)=4.17$, $p=0.05$, $\eta^2=0.04$, caused by larger errors in front space ($M=14.9^\circ$, $SD=6.8^\circ$) compared to back space ($M=12.1^\circ$, $SD=6.7^\circ$). In addition, there was a main effect of DISTANCE, $F(1,35)=22.49$, $p < 0.001$, $\eta^2=0.05$. Elevation errors were larger for near ($M=15.0^\circ$, $SD=5.6^\circ$) compared to far sounds ($M=12.0^\circ$, $SD=5.8^\circ$). No other main effect or interaction reached significance (all $F_s < 3.27$, all $p_s > 0.08$).

A similar analysis on signed errors (positive values indicate an upward bias) revealed a main effect of ANTERO-POSTERIOR SECTOR, $F(1,35)=25.15$, $p < 0.01$, $\eta^2=0.31$. Participants positioned the hand tracker below actual speaker location in front space (i.e., a downward bias; $M=-9.2$, $SD=10.5$), but pointed above actual speaker location in back space (i.e., upward bias: $M=4.2$, $SD=8.7$). These biases were more pronounced for near than far targets, resulting in an interaction between ANTERO-POSTERIOR SECTOR and DISTANCE, $F(1,35)=9.37$, $p=0.004$, $\eta^2=0.01$.

3.1.3. Distance

The overall absolute error in distance was 12.8 cm ($SD=6.7$ cm). The analysis on absolute error in distance revealed a main effect of SIDE, $F(1,35)=4.63$, $p < 0.04$, $\eta^2=0.02$, caused by larger errors for right ($M=13.6$ cm, $SD=7.5$ cm) than left targets ($M=12.1$ cm, $SD=6.6$ cm). In addition, all 2-way interactions reached significance: SIDE and DISTANCE, $F(1,35)=13.43$, $p < 0.001$, $\eta^2=0.01$; SIDE and ANTERO-POSTERIOR SECTOR, $F(1,35)=8.06$, $p < 0.007$, $\eta^2=0.01$;

ANTERO-POSTERIOR SECTOR and DISTANCE, $F(1,35)=15.69$, $p < 0.001$, $\eta^2=0.02$. Overall these interactions capture the larger error in distance estimation that occurred for near targets in front right space ($M=16.6$, $SD=10.2$).

A similar analysis on signed errors in distance (positive values indicate overestimation of target distance) revealed the main effects of SIDE (left or right), ANTERO-POSTERIOR SECTOR and DISTANCE all ($F_s > 9.87$, all $p_s < 0.003$). Overall, sound distance was overestimated ($M=5.8$ cm, $SD=10.7$ cm), but more for right ($M=9.1$ cm, $SD=10.5$ cm) compared to left sounds ($M=2.5$ cm, $SD=11.5$ cm). Overestimation was also larger for front ($M=11.0$ cm, $SD=10.7$ cm) compared to back sounds ($M=0.6$ cm, $SD=12.7$ cm), and for near ($M=7.2$ cm, $SD=10.8$ cm) compared to far sounds ($M=4.4$ cm, $SD=11.3$ cm). Several higher-order interactions also reached significance: SIDE and ANTERO-POSTERIOR SECTOR, $F(1,35)=5.65$, $p < 0.02$, $\eta^2=0.003$, SIDE and DISTANCE, $F(1,35)=5.46$, $p < 0.03$, $\eta^2=0.002$, and between SIDE, ANTERO-POSTERIOR SECTOR and DISTANCE, $F(1,35)=22.57$, $p < 0.001$, $\eta^2=0.005$. This 3-way interaction did not reveal any difference between near and far signed errors targets when placed in the back-right sector.¹

3.2. Hand pointing responses: the effects of adding a visual frame

Having established performance in the three dimensions in the gray visual condition, we turned to examine to what extent adding a visual frame affected sound localization performance. To this aim, we entered absolute and signed errors into ANOVAs with VISUAL CONDITION (gray, grid) and ANTERO-POSTERIOR SECTOR as within-participant variables, separately for each spatial dimension. We focused on the antero-posterior position of sounds because we predicted that the effect of the

¹ We interpreted these differences observed for near sounds in right space as the consequence of bio-mechanical constraints related to the fact that all participants were asked to point to the sounds using their right hand. To assess this hypothesis empirically, we analyzed the data of a pilot experiment in which 12 participants performed the same task reported in the present manuscript but using their preferred hand on a trial-by-trial basis. An ANOVA on absolute errors in distance on this dataset, using again SIDE, ANTERO-POSTERIOR SECTOR and DISTANCE as factors, yielded no significant main effect or interaction (all $F_s < 3.20$, all $p_s > 0.10$). On signed errors in distance, we found only a significant interaction between ANTERO-POSTERIOR SECTOR and DISTANCE, $F(1,11)=7.22$, $p=0.02$, $\eta^2=0.03$. This provides initial evidence that at least part of the specificities observed for right-sided sounds could depend upon the imposed use of the right hand for the response.

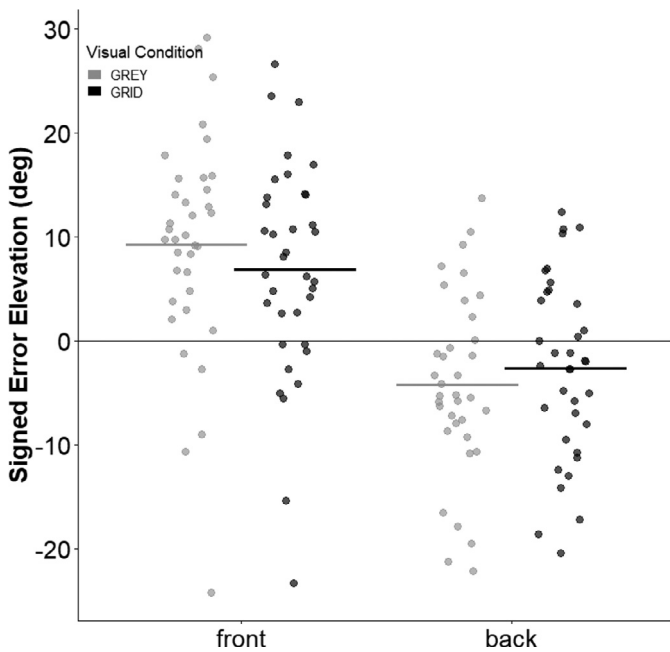


Fig. 3. Signer Error in Elevation (deg) as a function of Visual condition (gray, grid) and antero-posterior sector (front, back). In both graphs horizontal bars represent the mean of each condition, while points show participants value. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

visual manipulation could be maximal for sounds were delivered in the front space rather than back space.

The analyses on absolute and signed errors in azimuth, elevation and distance, did not reveal any main effect or interaction involving VISUAL CONDITION (all $F_s < 1.97$, all $p_s > 0.17$). One notable exception was signed error in elevation. For this measure, we found a significant interaction between ANTERO-POSTERIOR SECTOR and VISUAL CONDITION, $F(35) = 12.72$, $p = 0.001$, $\eta^2 = 0.27$. When sounds were delivered in the front space, participants judged their position lower than their actual location. This bias was smaller in the grid condition ($M = -6.8^\circ$, $DS = 10.3^\circ$) compared to the gray condition ($M = -9.2^\circ$; $DS = 10.5^\circ$; simple main effect: $F = 12.12$, $p = 0.001$). When sounds were delivered in the back space, participants judged their position as higher than their actual location. Again, this bias smaller in the grid condition ($M = 2.7^\circ$, $DS = 8.6^\circ$) compared to the gray condition ($M = 4.2^\circ$; $DS = 8.7^\circ$; simple main effect: $F = 5.84$, $p = 0.02$) (Fig. 3 and Table 1).

3.3. Head rotation

Head-movements occurred in 68.3% of trials on average ($SD = 38.3\%$; median = 89.9, range = 0–100%), for targets in front and back space (71.3% and 65.3% of trials, respectively). Fig. 4A shows the substantial variability of this spontaneous behavior in both visual conditions. In front space, the first head movement for sounds at 45° to the right was of 40.1° ($SD = 7.2^\circ$), whereas for sounds at 45° to the left it was 39.2° ($SD = 7.3^\circ$). In back space, the first head rotation for sounds at 135° to the right was of 67.9° degrees ($SD = 21.0^\circ$), whereas for sounds at 135° to the left it was 68.7° ($SD = 22.5^\circ$) (see Fig. 4B).

To analyze head movements rotation around the vertical axis, we considered two variables: 1) rotation amplitude of the first head movement (deg) and 2) time to first head-movement (the time between the beginning of the sound emission and the beginning of the first head movement in millisecond). Trial without head movements were excluded from these analyses. This resulted in the exclusion of 6 participants which did not move at all and 5 participants for whom less than 12 trials per condition remained. In addition, we focused only on those trials in which participants rotated towards the correct sound

side. This occurred on 97.5% ($SD = 2.0$) of trials on average, indicating again that participants easily disambiguate sound side.

To test the effect of the presence of the grid on these two variables (head-rotation latency and amplitude), we entered each measure separately in an ANOVA with VISUAL CONDITION and ANTERO-POSTERIOR SECTOR as within variables. ANOVA on head-rotation amplitude revealed no main effect or interaction (all $F_s < 1.26$, all $p_s > 0.27$). The analysis on time to first head-movement revealed that participants started to rotate their head 978 milliseconds ($SD = 155$) after sound emission on average. In addition, a significant main effect of ANTERO-POSTERIOR SECTOR emerged, $F(1,24) = 7.90$, $p = 0.02$, $\eta^2 = 0.25$. When sounds were emitted in front space, participants responded faster ($M = 957$ ms, $SD = 154$, ms) compared to when they were emitted from the back ($M = 999$ ms, $SD = 156$ ms) (Fig. 5A).

3.4. Gaze direction

Similar to head rotation analyses, we focused on trials in which participants gazed towards the same side of the sound (correct trials). This resulted in the exclusion of 10 participants for whom less than 12 trials per condition remained. On average participants gazed towards the same side of the target on 97.2% ($SD = 1.7$) of trials. Fig. 4C shows gaze direction as a function of target position. In front space, first gaze was directed to 42.1° ($SD = 18.9^\circ$) for sounds 45° right, and to 44.6° ($SD = 20.7^\circ$) for sounds 45° left. In back space, first gaze was directed 67.2° ($SD = 27.3^\circ$) for sounds 135° right, and to 70.5° ($SD = 25.2^\circ$) for sounds 135° left.

To analyze gaze movements, we considered two variables: (1) direction of the first gaze movement (deg) and (2) time to first gaze-movement (the time between the beginning of the sound emission and the beginning of the first gaze movement in millisecond).

To test the effect of our visual manipulation we entered these two dependent variables in separate ANOVAs, with VISUAL CONDITION and ANTERO-POSTERIOR SECTOR as within variables. The ANOVA on direction of the first gaze movement revealed a significant two-way interaction, $F(1,25) = 5.70$, $p = 0.03$, $\eta^2 = 0.19$, caused by a less symmetrical gaze directions in response to sounds in back space ($M = -3.4^\circ$, $DS = 9.0^\circ$), specifically in the grid condition (simple main effect: $F = 5.01$, $p = 0.03$).

A similar ANOVA on time to first gaze-movement revealed that participants started to gaze sounds 698 milliseconds ($SD = 164$) after sound emission. In addition, we found a significant main effect of ANTERO-POSTERIOR SECTOR, $F(1,25) = 19.54$, $p < 0.001$, $\eta^2 = 0.44$. Unsurprisingly, participants oriented their gaze faster to front ($M = 651$ ms, $SD = 153$ ms) compared back sounds ($M = 744$ ms, $SD = 174$ ms) (Fig. 5B). More interestingly, we also found a significant main effect of VISUAL CONDITION, $F(1,25) = 12.85$, $p = 0.001$, $\eta^2 = 0.34$. When the visual scenario was the grid, participants latency was shorter ($M = 666$ ms, $SD = 146$ ms) compared to when it was gray ($M = 728$ ms, $SD = 271$ ms). To test specifically whether VISUAL CONDITION had a greater impact in front than back space we run simple main effects. These revealed that the effect of VISUAL CONDITION was more substantial in the front ($F = 19.46$, $p < 0.001$) than back space ($F = 3.82$, $p = 0.06$; note that the 2-way interaction between VISUAL CONDITION and ANTERO-POSTERIOR SECTOR also approached significance: $F(1,25) = 6.46$, $p = 0.075$, $\eta^2 = 0.12$) (Fig. 5B).

4. Discussion

The present study examined the effect of seeing a simple visual spatial frame on sound localization, in a context in which participants were free to move their head and eyes while listening to sounds. To this aim, we leveraged a virtual reality (VR) approach which allows accurate control over the visual scenarios, accurate playback of free field sounds in three-dimensional space, and free and measurable movements of head and gaze during sounds playback.

Table 1
Absolute error and Signed error of 3 dimensions of the space as a function of target position for grid and gray visual conditions. Standard deviation (SD) between brackets.

		Azimuth (deg)		Elevation (deg)		Distance (cm)	
		Absolute Error	Signed Error	Absolute Error	Signed Error	Absolute Error	Signed Error
GRAY							
NEAR							
FRONT							
	Left	10.3 (7.8)	-5.8 (9.2)	16.0 (8.1)	-12.0 (10.9)	12.8 (9.0)	9.8 (12.1)
	Right	11.5 (13.1)	4.9 (13.5)	17.7 (9.5)	-10.0 (15.4)	16.6 (10.2)	15.8 (10.9)
BACK							
	Left	15.8 (13.2)	-5.7 (16.0)	14.0 (8.8)	5.1 (12.3)	10.4 (7.6)	-1.0 (12.1)
	Right	12.8 (9.6)	6.6 (11.1)	12.2 (7.6)	4.7 (10.5)	11.8 (9.3)	4.4 (13.3)
FAR							
FRONT							
	Left	8.5 (6.1)	-4.5 (7.5)	12.7 (6.5)	-8.0 (10.7)	11.9 (9.3)	6.7 (12.9)
	Right	7.5 (6.2)	4.2 (5.3)	13.2 (8.1)	-6.7 (12.4)	14.0 (8.3)	11.8 (10.2)
BACK							
	Left	13.7 (12.9)	-4.7 (11.4)	11.5 (10.1)	3.2 (7.4)	13.2 (8.8)	-5.2 (14.6)
	Right	11.7 (7.2)	5.6 (10.1)	10.6 (7.2)	3.8 (10.3)	12.1 (8.5)	4.2 (13.5)
GRID							
NEAR							
FRONT							
	Left	11.2 (8.8)	-4.9 (9.4)	15.4 (8.3)	-8.7 (12.6)	12.9 (8.1)	10.8 (10.6)
	Right	9.3 (7.0)	6.4 (6.5)	15.7 (8.9)	-8.8 (14.0)	17.5 (10.1)	16.3 (11.7)
BACK							
	Left	14.8 (11.2)	-9.0 (14.2)	13.8 (7.9)	1.7 (12.8)	10.3 (7.7)	-1.2 (12.4)
	Right	12.4 (7.6)	7.3 (10.2)	12.2 (7.5)	3.8 (10.7)	12.6 (8.1)	5.1 (12.2)
FAR							
FRONT							
	Left	8.9 (6.4)	-4.9 (8.7)	11.8 (6.4)	-5.3 (10.1)	11.4 (8.6)	5.7 (12.8)
	Right	7.4 (8.8)	3.7 (6.0)	12.6 (8.8)	-4.6 (11.2)	14.3 (9.1)	12.0 (11.7)
BACK							
	Left	13.6 (9.5)	-6.3 (12.1)	11.6 (10.7)	2.6 (12.0)	13.9 (9.1)	-6.6 (12.9)
	Right	11.2 (5.6)	7.7 (7.4)	9.8 (7.3)	2.5 (8.8)	12.5 (10.3)	3.7 (15.3)

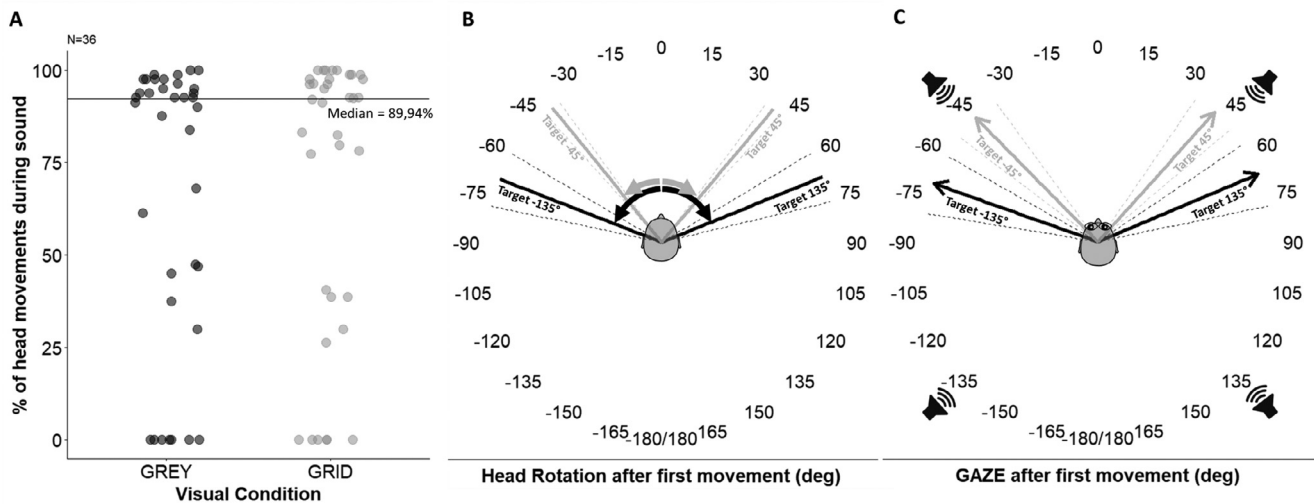


Fig. 4. (A) Boxplot of percentage head-movements: the percentage of trials in which head movements occurred for each participant. (B) Average of the rotation around vertical axis of the first head movement only when participants turned to the side of the target (same dataset of the analysis). (C) Average of the direction of the first gaze movement only when participants turned to the side of the target (same dataset of the analysis). Dashed lines represent confidence intervals.

When no visual spatial frame was available (gray condition), sound localization in azimuth was better in front than back space. When placing the tracker in back space participants had to turn around and modify their posture. This adjustment entails re-coding of sound position from egocentric to world-centered coordinates, leading to more uncertain estimates of sound position (Kopinska and Harris, 2003, Pavani et al., 2008). In addition, during the gray condition the participants' responses were characterized by a bias to point to positions

more eccentric than actual speaker locations. Errors in azimuth were higher for far rather than near targets. Although elevation was not varied, errors in the vertical dimension were larger for near than far sounds, and frontal sounds were perceived as lower than their actual position. Sound distance was overestimated, particularly for front and right sounds. Concerning distance, it is important to notice that our approach led us to measure sounds localization abilities also considering this dimension.

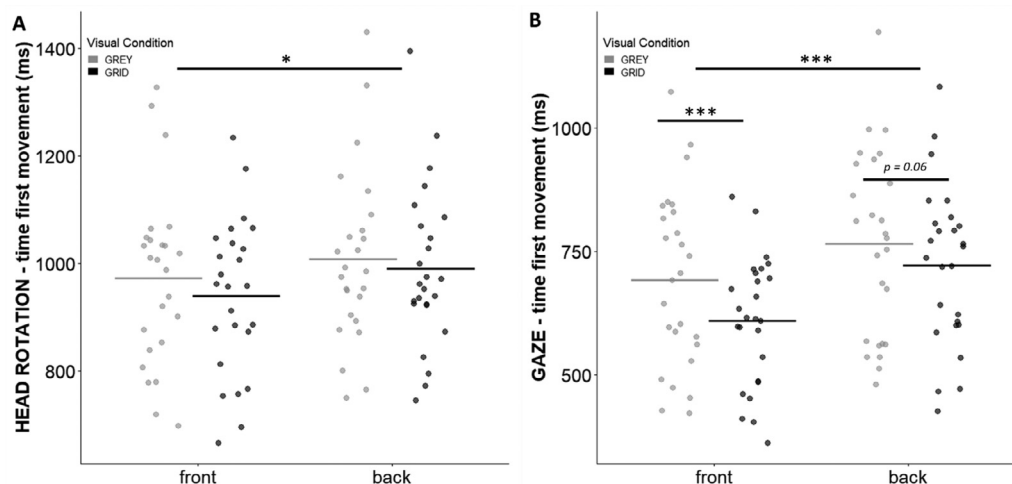


Fig. 5. T, the time between the beginning of the sound stimulation and the beginning of the first head or eyes movement. (A) Boxplot of Head Rotation time first movement as a function of target position (front, back) color coded as a function of Visual Condition (gray, grid). (B) Boxplot of GAZE time first movement as a function of target position (front, back) color coded as a function of Visual Condition (gray, grid).

Adding a visual spatial frame (grid condition) did not change the overall pattern of hand pointing responses in azimuth or distance. However, it resulted in improved accuracy in elevation and it affected participant's gaze behavior, producing faster orienting responses to the sounds.

4.1. Studying 'visual facilitation' with virtual reality

The question on the role played by visuo-spatial references on sound localization dates back to the 1970s (Warren, 1970), when it was termed 'visual facilitation'. Yet, until the advent of VR methods, visual facilitation could only be addressed with somewhat crude manipulations (e.g., blindfolding participants). Most importantly, when visual information about the environment was provided (i.e., eyes-open conditions), it was a methodological challenge to disentangle the contributions related to seeing the overall structure of the environment, from the contribution related to seeing the potential sources of sound. Using a VR approach, we investigated the role of visually structured information on sound localization, in the absence of visual priors about the speakers' position. While immersed in the VR scene, participants were only informed that target sounds would be delivered within reaching distance, but had no further information on their positions — i.e., they expected sounds to appear all around the body but inside an estimated range of distance (Gaveau et al., 2020).

To the best of our knowledge, only two previous works have addressed a similar question using VR techniques (Ahrens et al., 2019, Majdak et al., 2010). Majdak et al. (2010) tested the effect of seeing a simple visual environment on sound localization. Subjects were immersed in a virtual environment which comprised a sphere (diameter 5 m), marked with grid lines every 5° horizontally and every 11.25° vertically. Furthermore, participants required to judge elevation and azimuth of the sound taking advantage of grid lines. They reported that participants' hand pointing to sounds was enhanced by the visual grid in both horizontal and vertical dimensions, compared to a condition of total darkness. Our results are consistent with these findings. However, in the present study we observed an advantage only in terms of elevation, whereas sound localization in azimuth remained unchanged in the presence of the visual grid. This different result could reflect methodological discrepancies between the two studies. First, our visual grid was intentionally conceived to avoid any vertical line that could serve as anchor for sound localization. Second, our virtual environment conveyed also a sense of distance, which could have introduced additional uncertainty

in the interpretation of the auditory cues. Third, we used real sounds instead of virtual ones, possibly leading to more precise sound localization overall.

Nonetheless, both studies converge in stressing the importance of having a visual frame when localizing sound. In particular, the grid condition may have provided a detailed visual map for positioning sound sources.

Ahrens and colleagues also took advantage of VR (Ahrens et al., 2019), allowing participants to see the room dimensions as well as their hand-position, whereas the speakers' array was not visible. Using sounds beyond reaching distance (2.4 m from the head), they found that visual information decreased errors both in the horizontal and vertical dimensions, as compared to a blindfolded condition. In Ahrens et al.'s study the VR reconstructed room was a careful visual replica of the in which the experiment was conducted, with the foam wedges of the anechoic chamber providing extremely rich vertical and horizontal visual references all around the participant. As discussed above for the study by Majdak and colleagues, it may be that these substantial visual cues played a role in the improvements observed in the horizontal dimension, serving as references or place-holders for sound localization when visible. In our more minimalist visual scenario we intentionally avoided all vertical visual references.

One difference with respect to both these previous works is that we presented sounds within reaching distance. In Majdak et al., sounds were delivered through headphones and although they were likely perceived externalized, it is difficult to establish at which distance they were perceived. In Ahrens et al., sounds were instead delivered away from the body, at 2.4 m. Our rationale in presenting sounds within reaching distance was to allow participants to respond using the hand-held tracker and measure their accuracy in distance estimation. However, this choice limited our studied space to the near-field and potentially influenced the interactions with the visual environment we created. Future works could examine to what extent the impact of visual environment on sound localization could emerge differently for sounds at different distances from the body, or — in case of enclosed spaces like rooms — at different relative distances with respect to the visible surfaces.

Another important difference is that our participants were free to move their heads while listening to sounds. This spontaneous orienting behavior also involved the eyes, and likely made azimuth localization easier for our participants. In turn, this could have reduced the possibility of observing visual condition effects in azimuth. However, this gaze (i.e., head and eyes) orienting behavior revealed interesting findings.

We found that correct gaze orienting responses in the direction of the sound started earlier when exposed to a visual grid compared to the gray condition. A similar trend was observed also for head rotation. We believe that the interest of analyzing this measure is related to the fact that, unlike hand pointing, gaze orienting is a more implicit measure of sound localization performance — particularly when the first gaze to sound response is considered. Measuring participants' gaze allowed us to capture an early and implicit aspect of acoustic space perception, which differs from the explicit head pointing method that has been used in sound localization tasks as an alternative to hand pointing. Taking this perspective, our results suggest that even a minimal visual spatial frame can speed up the right-left implicit disambiguation of sound position.

These effects of the visual features of the environment on sound localization are complementary with the line of research that examined the effect of the *acoustic* feature of the environment on sound localization and visual scene perception. For instance, Gil-Carvajal et al. (2016) have examined the effect of a mismatch between playback and recording room on perceived distance and azimuth direction of sounds. They found that sound distance ratings decreased when measured in an environment that was more reverberant than the original recording room, whereas azimuth direction remained unaffected. Most interestingly, they also observed that changes in the visual attributes of the room were ineffective and concluded that visual congruency is less crucial than the correspondence between the acoustical features of environment and the stimuli.

Other works examined the interplay between auditory and visual features of the environment. Etchemendy et al. (2017) have found that auditory reverberation cues impact on the perception of visible room size, revealing that the auditory environmental context can modulate visual distance perception. At the same time, Schutte et al. (2019) have shown that visual room impression does not affect people's abilities to estimate rooms' reverberations. The latter evidence in particular is directly relevant to the present work, as it suggests that the effect of the visual grid proposed in our experiment may not have influenced the extraction of the acoustic features from the environment.

In the present study conducted in three different experimental sites, we did not measure the rooms' reverberation limiting the possibility to deepen the acoustic influences on both sound localization performance and visual manipulation effects. Further work should aim to combine the contribution of visual and auditory features of the environment with the active sound localization approach introduced by the present work, to examine to what extent the relative contribution and potential interactions between these multisensory contextual contributions to spatial hearing. It would also be important to assess sound localization beyond reaching (as here) to address more directly whether sounds that perceived further away from the head (e.g., at 2.4 m as in Ahrens et al., 2019 or beyond) could interact more with the wider spaces of rooms or open VR environments like the ones simulated in our grid condition.

5. Conclusion

In this study, we documented that providing visual frame that is totally uninformative about sound source position in space helps sound source localization in active listening conditions. These findings contribute to emphasize the indirect but positive contribution of minimally structured vision to sound localization and further promote the idea that sound localization ability should be conceived as a multisensory process. Our findings also underlie the importance of allowing and measuring spontaneous head and gaze movements. The here adopted VR technology allowed us to go beyond traditional approaches in the study of spatial hearing by allowing participants to move their eyes and head during and after sound presentation, while retaining full control over sound placement and recording of our dependent variables using eco-

logically valid contexts, which is crucial in studying hearing experience as suggested by recent studies in the field (Hadley et al., 2019).

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We thank all the participants and the colleagues who helped during research. We thank Giordana Torresani for graphic solutions (Fig. 1A). **Funding:** C.V. was supported by a grant of the Università Italo-Francesca (UIF)/Université Franco-Italienne (UFI) and the Ermenegildo Zegna Founder's Scholarship. F.P., C.V. and AF were supported by a grant of the Agence Nationale de la Recherche (ANR-16-CE17-0016, VIRTUAL-HEARING3D, France) and by a prize of the Fondation Medisite (France). The study was supported by the IHU CeSaMe ANR-10-IBHU-0003 and it was performed within the framework of the LABEX CORTEX (ANR-11-LABX-0042) of Université de Lyon. We thank JL Borach, S Alouche, S Terrones for administrative support, and Eric Koun for informatics support.

References

- Ahrens, A., Lund, K.D., Marschall, M., Dau, T., 2019. Sound source localization with varying amount of visual information in virtual reality. *PLoS ONE* 14 (3) Mar.
- Alais, D., Burr, D., 2004. The ventriloquist effect results from near-optimal bimodal integration. *Curr. Biol.* 14 (3), 257–262 Feb.
- Bolognini, N., Leor, F., Passamonti, C., Stein, B.E., Làdavas, E., 2007. Multisensory-mediated auditory localization. *Perception* 36 (10), 1477–1485 Oct.
- Etchemendy, P.E., et al., 2017. Auditory environmental context affects visual distance perception. *Sci. Rep.* 7 (1), 1–10 Dec.
- Gaveau, V., et al., 2020. SPHERE: a novel approach to 3D and active sound localization. *bioRxiv* p. 2020.03.19.998906, Mar.
- Gil-Carvajal, J.C., Cubick, J., Santurette, S., Dau, T., 2016. Spatial hearing with incongruent visual or auditory room cues. *Sci. Rep.* 6 (1), 1–10 Nov.
- Hadley, L.V., Brimjoin, W.O., Whitmer, W.M., 2019. Speech, movement, and gaze behaviours during dyadic conversation in noise. *Sci. Rep.* 9 (1), 1–8 Dec.
- Heffner, R.S., Heffner, H.E., 1992. Visual factors in sound localization in mammals. *J. Comp. Neurol.* 317 (3), 219–232 Mar.
- Heffner, H.E., Heffner, R.S., 2014. *The Behavioral Study of Mammalian Hearing*. Springer, New York, NY, pp. 269–285.
- Hofman, P.M., Van Riswick, J.G.A., Van Opstal, A.J., 1998. Relearning sound localization with new ears. *Nat. Neurosci.* 1 (5), 417–421.
- Knudsen, E.I., Knudsen, P.F., 1985. Vision guides the adjustment of auditory localization in young barn owls. *Science* (80-). 230 (4725), 545–548.
- Kopinska, A., Harris, L.R., 2003. Spatial representation in body coordinates: evidence from errors in remembering positions of visual and auditory targets after active eye, head, and body movements. *Can. J. Exp. Psychol.* 57 (1), 23–37.
- Majdak, P., Goupell, M.J., Laback, B., 2010. 3-D localization of virtual sound sources: Effects of visual environment, pointing method, and training. *Attent. Percept. Psychophys.* 72 (2), 454–469 Feb.
- Morronegello, B.A., Rocca, P.T., 1987. Infants' localization of sounds in the horizontal plane: effects of auditory and visual cues. *Child Dev.* 58 (4), 918 Aug.
- Parise, C.V., Knorre, K., Ernst, M.O., 2014. Natural auditory scene statistics shapes human spatial hearing. *Proc. Natl. Acad. Sci. U. S. A.* 111 (16), 6104–6108 Apr.
- Pavani, F., Husain, M., Driver, J., 2008. Eye-movements intervening between two successive sounds disrupt comparisons of auditory location. *Exp. Brain Res.* 189 (4), 435–449 Aug.
- Pavani, F., et al., 2017. Spatial and non-spatial multisensory cueing in unilateral cochlear implant users. *Hear. Res.* 344, 24–37.
- Rabini, G., Altobelli, E., Pavani, F., 2019. Interactions between egocentric and allocentric spatial coding of sounds revealed by a multisensory learning paradigm. *Sci. Rep.* 9 (1) Dec.
- Redon, C., Hay, L., 2005. Role of visual context and oculomotor conditions in pointing accuracy. *Neuroreport* 16 (18), 2065–2067.
- Schutte, M., Ewert, S.D., Wiegand, L., 2019. The percept of reverberation is not affected by visual room impression in virtual environments. *J. Acoust. Soc. Am.* 145 (3), EL229–EL235.
- Shelton, B.R., Searle, C.L., 1980. The influence of vision on the absolute identification of sound-source position. *Percept. Psychophys.* 28 (6), 589–596 Nov.
- Strelnikov, K., Rosito, M., Barone, P., 2011. Effect of audiovisual training on monaural spatial hearing in horizontal plane. *PLoS ONE* 6 (3), e18344 Mar.
- Terry Bahill, A., McDonald, J.D., 1983. Frequency limitations and optimal step size for the two-point central difference derivative algorithm with applications to human eye movement data. *IEEE Trans. Biomed. Eng. BME-30* (3), 191–194.

Tonelli, A., Brayda, L., Gori, M., 2015. Task-dependent calibration of auditory spatial perception through environmental visual observation. *Front. Syst. Neurosci.* 9, 84 no. June/June.

Valzolgher, C., Campus, C., Rabini, G., Gori, M., Pavani, F., 2020a. Updating spatial hearing abilities through multisensory and motor cues. *Cognition* 204, 104409 Nov.

Valzolgher, C., Verdelet, G., Salemme, R., Lombardi, L., Gaveau, V., Farné, A., Pavani, F., 2020b. Reaching to sounds in virtual reality: A multisensory-motor ap-

proach to promote adaptation to altered auditory cues. *Neuropsychologia*, 107665 doi:10.1016/j.neuropsychologia.2020.107665.

Verdelet G. et al., "Assessing spatial and temporal reliability of the vive system as a tool for naturalistic behavioural research," 2020, pp. 1–8.

Warren, D.H., 1970. Intermodality interactions in spatial localization. *Cogn. Psychol.* 1 (2), 114–133 May.

ASSESSING SPATIAL AND TEMPORAL RELIABILITY OF THE VIVE SYSTEM AS A TOOL FOR NATURALISTIC BEHAVIOURAL RESEARCH

G. Verdelet^{1,2,3}, C. Desoche^{1,2,3}, F. Volland^{1,2}, A. Farnè^{1,2,3,4}, A. Coudert^{1,2,5}, R. Hermann^{1,2,5}, E. Truy^{1,2,5}, V. Gaveau^{1,2}, F. Pavani^{1,2,3,4,6}, and R. Salemmé^{1,2,3}

1. Integrative Multisensory Perception Action & Cognition Team (ImpAct), INSERM U1028, CNRS UMR5292, Lyon Neuroscience Research Center (CRNL), Lyon, France
2. University of Lyon 1, Lyon, France
3. Neuro-immersion platform, Lyon, France
4. Center for Mind/Brain Sciences (CIMeC), University of Trento, Italy
5. ENT and cervicofacial surgery department, Edouard Herriot Hospital, HCL, Lyon, France
6. Department of Psychology and Cognitive Science, University of Trento, Italy

ABSTRACT

Nowadays behavioral and cognitive neuroscience studies have turned ‘naturalistic’, aiming at understanding brain functions by maintaining complexity close to everyday life. Many scholars started using commercially available VR devices, which, were not conceived as research tools. It is therefore important to assess their spatio-temporal reliability and inform scholars about the basic resolutions they can achieve. Here we provide such an assessment for the VIVE (HTC Vive) by comparing it with a VICON (BONITA 10) system. Results show a submillimeter Vive precision (0.237mm) and a nearest centimeter accuracy (8.7mm static, 8.5mm dynamic). we also report the Vive reaction to a tracking loss: the system takes 319.5 +/- 16.8 ms to detect the loss and can still be perturbed for about 3 seconds after tracking recovery. The Vive device allows for fairly accurate and reliable spatiotemporal measurements and may be well-suited for studies with typical human behavior, provided tracking loss is prevented.

Index Terms— Virtual reality, Head-mounted display, motion capture, HTC Vive, VICON, accuracy, precision, lost tracking, neuroscience, spatio-temporal.

1. INTRODUCTION

Our complex interactions with the physical environment entail precise sequence of movements that unfold in time and space. In the last decades, researchers interested in understanding human cognition through the study of behavior in naturalistic conditions have started to exploit Virtual Reality (VR) via Head-Mounted Displays (HMD). This allows to understand brain function going beyond the (way too) simplistic experimental settings typically used in

laboratory conditions and to delivered complex but rigorously controlled stimulations to human and non-human observers. Moreover, recent advances of VR technology for gaming have led to commercialized and affordable solutions for tracking whole-body movements in space. The combination of VR with tracking of motor responses in real-time has the potential to open new avenues for experimental and applied psychology, as well as for rehabilitation. Yet, the spatial and temporal reliability of these commercially available solutions for the purpose of research remains to be ascertained [1]. In the present study, we used a validated tool conceived for motion tracking in scientific research (i.e., the VICON system) to study the reliability of measurements obtained with the increasingly more popular HTC Vive system.

The HTC Vive system used 1 or 2 lighthouses (base stations), and a set of track objects (the HMD, plus one or more controllers or trackers). Each base station includes 2 lasers that emit infrared light (IR): one projects a horizontal array of light; the other projects a vertical array light. Each laser scans the tracked volume with a 60 Hz rate, so any tracked object inside the volume is hit at a rate of 120 Hz. Base stations are synchronized with one another, using either a sync cable or an LED array which pulse at a rate of 60 Hz, to switch from a base station to another [2]. Each tracked object contains a structure of photodiodes (TS3633) and embeds an internal gyroscope and accelerometer (IMU). In such a way, tracked objects location can be updated in real-time, with a form of dead reckoning tracking, and their position can be corrected via the lighthouses tracking system every 8.333 ms on one dimension, based on the current vertical or horizontal laser sweep. Using both lighthouse and IMU system, the tracking rate could up to 250 Hz for a Vive tracker [3].

With this tracking system, we can expect HTC Vive to have a submillimeter jitter and around the cm accuracy in a

tracking volume of 5x5x2 m³ for the HMD [4,5]. Although HMD and Vive tracker use the same technology, the different of tracker design may modify this performance. Unfortunately, the manufacturer does not provide any test of the reliability of the tracking across different static positions. Besides, the most important aspect to be certain would be the device reliability in space and time once operated in dynamic conditions – during a movement. Furthermore, in a research protocol, the stability over time and reliability of the measure is of utmost importance. We have already been warned that losing the HMD tracking may tilt the reference plane [5]. Yet, we want to put forward that the complementary, or substituting nature of the IMU tracking system when the lighthouses lose the tracked object, may be a cause of reduction in accuracy in the measure. It is thus important to identify the impact of lose tracking on the recorded measures.

The aim of the present study was to address each of these open questions, to extend the testing of the HTC Vive tracking system reliability to aspects that are relevant to behavioral research [5]. To this aim, our measurement methodology quantifies the accuracy and the precision of a Vive Tracker position for static and dynamic situations. Furthermore, our study examined the consequences of losing tracking on the performance measurement of Vive tracker position, in both static and dynamic conditions. As a benchmark for our measures, we used a motion capture system with high accuracy and precision, validated for research purposes (VICON). VICON is a motion capture system that uses cameras to track reflective markers illuminated by near infrared (NIR) light source with a submillimeter accuracy [6,7].

Table 1

Display	Dual AMOLED 3.6” diagonal
Resolution	1080 x 1200 per eye
Refresh rate	90 Hz
FOV	110°
Tracking system	Steam VR 1.0
Strobe	850 nm IR
Room Scale	
min	2m x 1.5m
max	3.5m x 3.5m
Required PC min	
Graphic Card	NVIDIA® GTX 970
CPU	Intel® Core™ i5-4590
Memory	4 GB + 16 GB RAM

Table 1. The characteristics of the head-mounted display used in the study, as reported by the manufacturer.

2. MATERIAL & METHODS

To test performance of the HTC Vive, we used a Vicon system composed by 7 cameras BONITA 10 as described below.

2.1. Instrumentation

The HTC Vive characteristics have been provided by its manufacturer “Table 1”. For our test, we use 2 lighthouses V1.0, the HMD and a tracker with its dongle. The Vive tracker comprises 23 photodiodes.

Vicon is a motion capture system used in life science and engineering industries. Our setting is composed of 7 BONITA 10 (small optical camera tracking reflective markers or constellations of markers illuminated by infrared light sources). Their characteristics have been provided by its manufacturer “Table 2”. The BONITA NIR LED and the base station IR LED do not appear to add noise to one another, but we have to be careful because some cases have been observed.

Table 2

System latency	2 ms
Strobe	68 high powered NIR LEDs @ 780 nm
Power	802.3af PoE
Resolution	1 megapixel (1024 x 1024)
Frame Rate	250 Hz
Precision	0.5 mm
Room Scale max	4m x 4m

Table 2. The characteristics of the Bonita 10 camera used in the study, as reported by the manufacturer.

To track the Vive tracker with HTC Vive and VICON systems at the same time, we have attached a reflective marker constellation to the tracker “Fig. 1”.

A VR-Ready computer with Windows 10 x64, equipped with an Intel® Core™ i5-6600K CPU @ 3.50 GHz, 16 GB of RAM, and a NVIDIA® GeForce® GTX 980 with 4 GB of GDDR5, was used for the HTC Vive. Another computer managed VICON data acquisition and was equipped with an Intel® Core™ i7-4771 CPU @ 3.5 GHz, 8 GB of RAM, and a NVIDIA® Quadro K2000.

The software to record HTC Vive devices position was Unity 5.6.0f3 x64 and used SteamVR. We used Vicon Tracker 2.0.1 for the tracking of the Vicon marker and we linked this software and Unity with the plugin unityVRPN.



Fig. 1. Vive Tracker with a reflective marker constellation.

2.2. Procedure

Three experiments were carried out to answer our questions. For each of them, we built a tracking volume of 2.5m x 2.5m x 2m where both Vive and VICON system were most efficient to track at the center, as shown in “Fig. 2”. We were careful that at least 2 VICON cameras could always localize the tracked object within the tracked volume. The base stations were positioned at 2 m height and 4 m distance from each other. Their inclination was of 45° downward and they were oriented towards the center of the preset volume, to allow ‘seeing’ both the object tracked in the volume and to the other base station.

The Vive system measures the object position in its own reference frame. To compare this value with that given by the VICON (defined in the VICON reference frame), we have applied an algorithm to achieve the transformation from the Vive’s to VICON’s reference frame. Once positions obtained by the Vive were transformed into the VICON reference frame, we could compare the tested system measured positions (Vive) with those obtained from our reference system (VICON). We calibrated the VICON and the Vive system to position Vive reference frame the closest possible than VICON reference frame to reduce the influence of passage algorithm (see Data Analysis section below).

For the first experiment, we were interested in the Vive tracking system dispersion and accuracy for static positions. We placed the tracked object at 3 calibration positions (see “Eq. 8”) and one of two sets of 16 positions at 1.6 m height (see “Fig. 3”) and acquired measures during 5 seconds at the Unity frame rate of 90 Hz with both Vive and VICON systems. We repeated the acquisition 10 times for the first dataset and 5 times for the second dataset, performing a new Vive and VICON calibration each time.

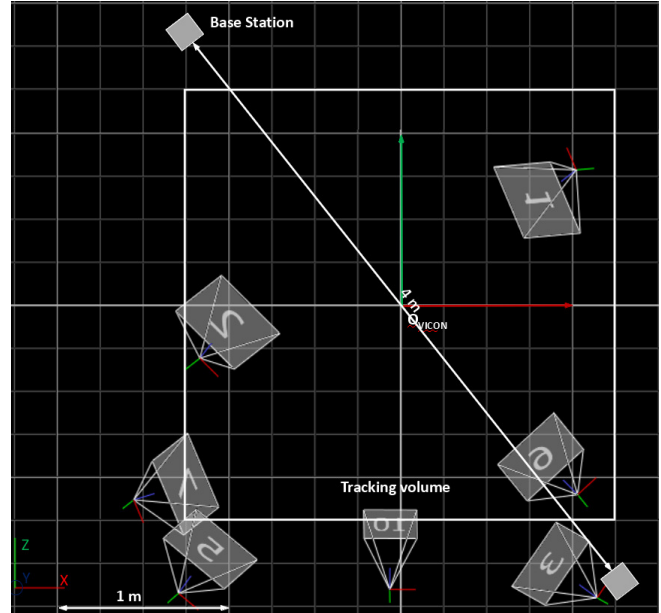


Fig. 2. Tracking volume and position of BONITA 10 and Base stations.

We calculated the mean dispersion for all static points for both VICON and Vive systems, then we estimated the Vive accuracy by calculating the distance between Vive and VICON measures of each target centroid. We distinguished accuracy across the tracking volume, so the mean of error at each point in the tracking volume, to check if the accuracy was variable in the space, and the global accuracy, so the mean of all point in the tracking volume, to estimate reliability of the measure anywhere.

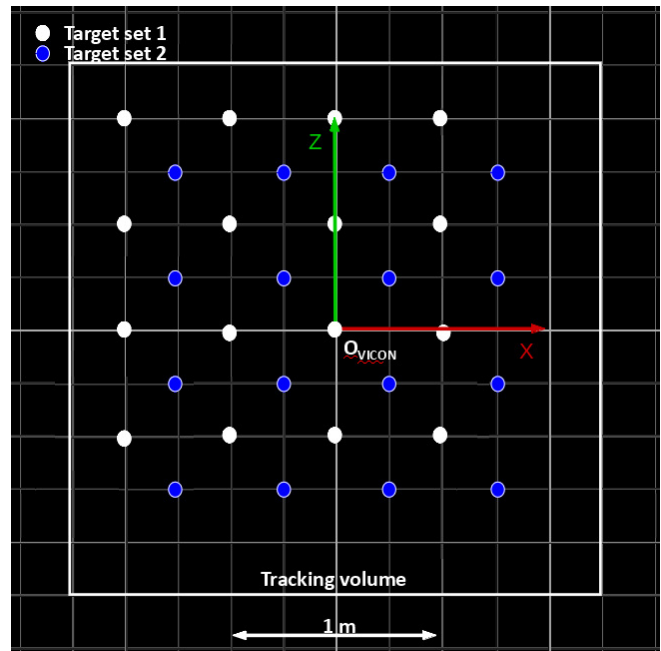


Fig. 3. Target positions in the tracking volume for the static analysis. We represented 2 different sets of static points.

For the second experiment, we wanted to compare the Vive performance in the static vs. the dynamic condition. We placed the tracking object at 3 calibration positions, then we recorded its position while it was manually displaced back and forth along a rail of 1.20 m at 0.8 m height 5 to 6 times, for a total recording time of 50-100 seconds. We repeated this procedure 6 times to have a total of 32 recorded movements. The movement speed was manually made to differ for each pass, alternating slow (20 seconds by pass) and fast movements (10 seconds by pass). Accuracy was estimated in a similar fashion to Experiment 1, except that here we compared positions frame by frame.

Experiment 3 aimed at identifying the impact of losing the tracked object by the base stations on its recorded position. We positioned the tracked object in the center of the tracking volume at 1.20 m height. We launched the recording, then we covered the object with a box to hide it from base stations and VICON cameras. While covered, we displaced the tracked object of approximately 50 cm and removed the box to record the new position. We repeated this procedure 9 times.

We considered the tracking was lost when VICON measures were not updated for 3 frames and we compared it with the Vive tracking loss to obtain the delay of lost tracking detection. Finally, we considered the first position as the centroid of the dataset (before tracking loss) and the final position as the centroid of the dataset after stabilization, see results section below, of the Vive measures.

2.3. Data Analysis

The dispersion of the position reported by the Vive was computed with the “Eq. 1” of a set of data $\{p_i\}$ acquired during 5 seconds at a rate of 90 Hz from a static position in the tracking volume:

Eq. 1:

$$Dispersion = \frac{\sum_{i=1}^N \|p_i - centroid(p)\|}{N}$$

To estimate the Vive accuracy, we considered two-point sets $\{p_i\}$ and $\{p'_i\}$; $i = 1, 2, \dots, N$ where p_i and p'_i were the gravity center of 5 seconds at a rate of 90 Hz for the position i in the tracking volume gathered by Vive and VICON system respectively. Both Vive and VICON system had their own reference frames which could be related by “Eq. 2” where R was a rotation matrix 3×3 , T a translation vector 3×1 , and N_i a noise vector:

Eq. 2:

$$p'_i = Rp_i + T + N_i$$

We tried two methods to compute a matrix and a vector to pass from vive reference frame to VICON reference frame so we were able to compare the real position with the expected position.

Method 1 was based on the article of K.S. Arun et al. [8], we found the R matrix vector by minimize the least-square “Eq. 3” and the vector T by $T = p' - Rp$ where p and p' were respectively the centroid of $\{p_i\}$ and $\{p'_i\}$:

Eq. 3:

$$\Sigma^2 = \sum_{i=1}^N \|p'_i - (Rp_i + T)\|^2$$

To minimize least-square, we computed the matrix 3×3 H as “Eq. 4”:

Eq. 4:

$$H = \sum_{i=1}^N q_i q_i^t$$

$$q_i = p_i - centroid(p)$$

$$q'_i = p'_i - centroid(p')$$

Then we calculated the singular value decomposition (SVD) “Eq. 5”:

Eq. 5:

$$H = U\Lambda V^t$$

Finally, the R matrix was determined according to the determinant of VU^t “Eq. 6”, we found the vector T as seen above:

Eq. 6:

$$\begin{cases} \text{If } \det(VU^t) = +1, \text{ then } R = VU^t \\ \text{If } \det(VU^t) = -1, \text{ then } R = V'U^t \\ \text{with } V' = [v_1, v_2, -v_3] \end{cases}$$

Although this method was used and confirmed in other studies, it might artificially decrease the error because it computes the best solution to find the minimum mismatch between the two sets of data. To counterbalance this bias, we also used method 2 to compute a transformation matrix to pass from vive reference frame to VICON reference frame.

First, we wanted to find the correspondent unit matrix from VICON reference frame to the Vive reference frame, “Eq. 7”:

Eq. 7:

$$[V'_x; V'_y; V'_z] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \rightarrow [V_x; V_y; V_z] = \begin{bmatrix} V_{x1} & V_{x2} & V_{x3} \\ V_{y1} & V_{y2} & V_{y3} \\ V_{z1} & V_{z2} & V_{z3} \end{bmatrix}$$

Here, we obtained V_x , V_y and, V_z by positioned the Vive/VICON tracked object in three specific positions P'_0 , P'_x and, P'_z in VICON reference frame and P_0 , P_x and P_z are corresponded Vive measures, as see in “Eq. 8”:

Eq. 8:

$$\begin{aligned} P'_x - P'_0 &= [1,0,0] \rightarrow P_x - P_0 = [V_{x1}, V_{x2}, V_{x3}] \\ P'_z - P'_0 &= [0,0,1] \rightarrow P_z - P_0 = [V_{z1}, V_{z2}, V_{z3}] \\ \text{cross}(P'_z, P'_x) &= [0,1,0] \rightarrow \text{cross}(P_z, P_x) = [V_{y1}, V_{y2}, V_{y3}] \end{aligned}$$

We built a passage matrix M and a translation vector T as define at “Eq. 9”:

Eq. 9:

$$M = \text{inv} \left(\begin{bmatrix} V_{x1} & V_{x2} & V_{x3} \\ V_{y1} & V_{y2} & V_{y3} \\ V_{z1} & V_{z2} & V_{z3} \end{bmatrix} \right)$$

$$T = P'_0 - P_0 M$$

Then, we had an equation to have a point p''_i from the reference frame Vive to p_i in the reference frame VICON, “Eq. 10”:

Eq. 10:

$$p''_i = p_i M + T$$

This method was dependent of the quality of calibration points. The error between vive and VICON position on the calibration points was reported on all other measure, so this method could increase the error found. Both methods were used and presented in the results.

3. RESULTS

3.1. Experiment 1: Static error

The experimental results showed a submillimeter dispersion for both VICON (0.027 +/- 0.011 mm) and Vive (0.237 +/-

0.076 mm) system “Fig. 4”. Nonetheless, precision of the Vive tracking system was one order of magnitude worse compared to the VICON precision for static condition.

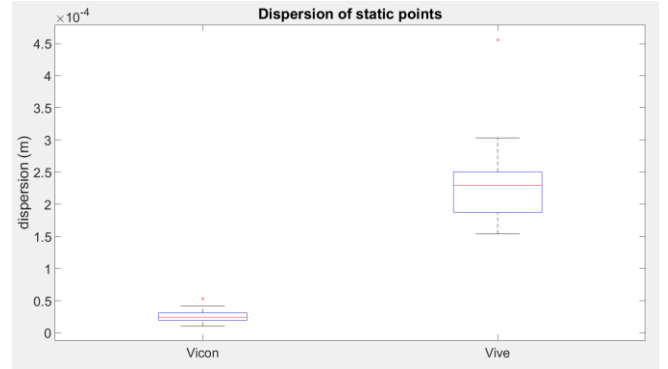


Fig. 4. Dispersion of static points. N = 14; dispersion VICON = 0.027 +/- 0.011 mm; dispersion Vive = 0.237 +/- 0.076 mm.

Regarding the accuracy across the tracking volume “Fig. 5”, both analysis methods showed less performance around the left top corner. The accuracy variability is more relevant with the method 2. Regarding the global accuracy, method 1 reveals better performance (5.7 +/- 1.2 mm; N = 14) than method 2 (9 +/- 1.5 mm; N = 12, 2 sets of data were excluded because the calibration points increase artificially the error mean).

3.2. Experiment 2: Dynamic error

The second experimental results concerned the comparison of accuracy in dynamic conditions across methods. With method 1 the accuracy was better (4.8 +/- 2.4 mm; N = 32) than with method 2 (9.4 +/- 4 mm; N = 26, 6 sets of data were excluded because bad calibration points).

Accuracy comparison between static and dynamic conditions “Fig. 6” showed that mean accuracy was similar for both method 1 (5.7 mm vs 4.8 mm) and method 2 (9 mm vs 9.4 mm). However, standard deviation in the dynamic condition was greater than the static condition, especially with method 2.

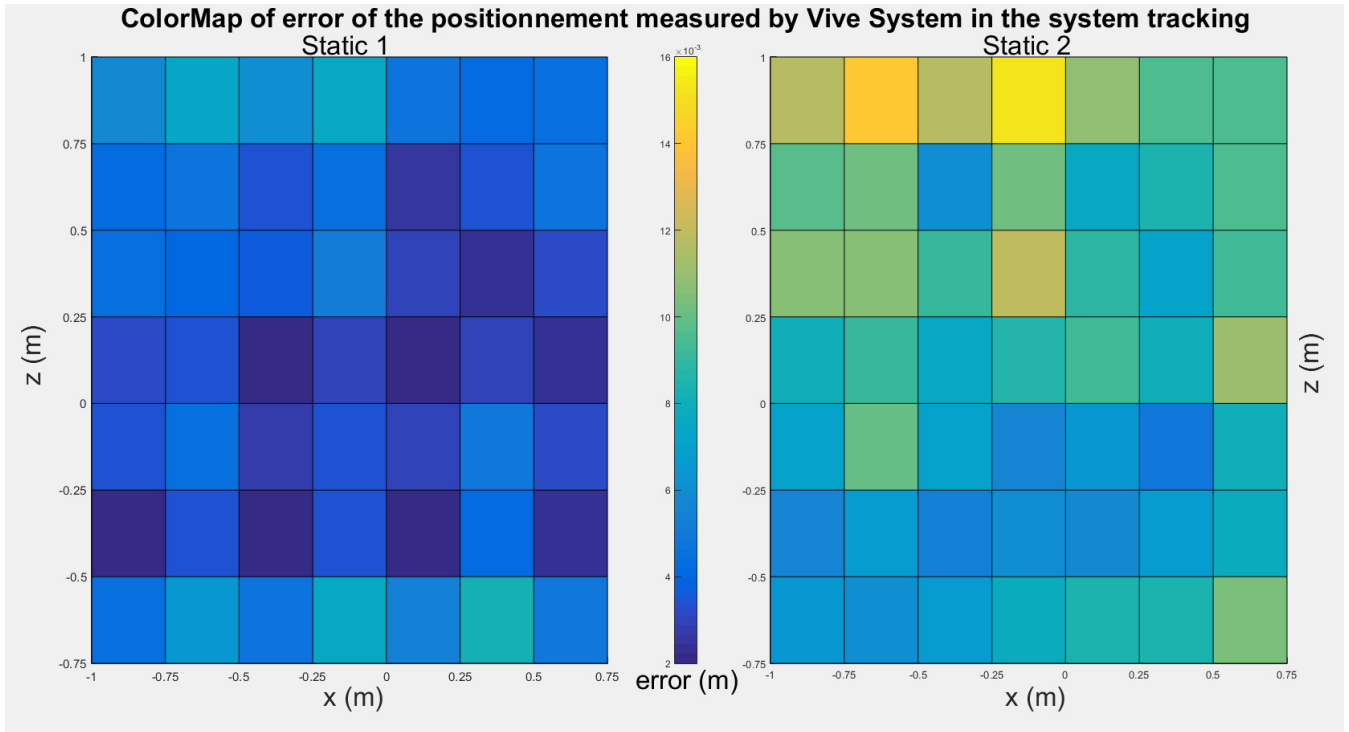


Fig. 5. Color map of error measured between Vive and VICON system to the position of object tracked across the tracking volume. Static 1 shows the fluctuation of the error as measured with method 1. The nearer the target to the center the greater the accuracy. Static 2 shows the fluctuation of the error measured with the method 2. Accuracy decreases around the left top corner.

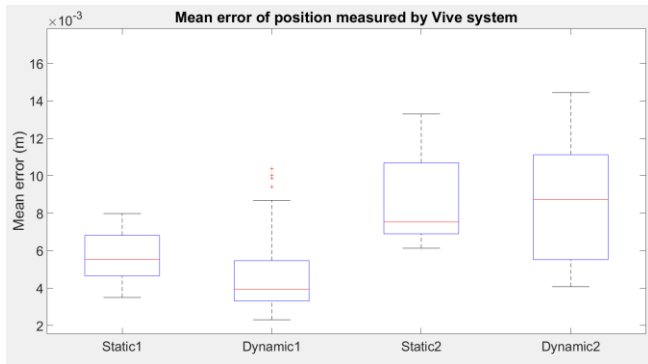


Fig. 6. Box plot of the mean error of position measured by the Vive system compared with the VICON position for both methods and for static and dynamic conditions. Static 1 represents the mean error of each trial in static experience using method 1 to change the reference frame, $N = 14$; error = 5.7 ± 1.2 mm. Dynamic 1 represents the mean error of each trial in dynamic experience using method 1 to change the reference frame, $N = 32$; error = 4.8 ± 2.4 mm. Static 2 represents the mean error of each trial in static experience using method 2 to change the reference frame, $N = 12$; error = 8.7 ± 2.4 mm. Dynamic 2 represents the mean error of each trial in dynamic experience using method 2 to change the reference frame, $N = 20$; error = 8.5 ± 3.2 mm.

3.3. Experiment 3: Tracking loss impact

During the third experiment we first estimated the Vive delay to detect the tracking lost to 319.5 ± 16.8 ms. Then, results showed a disruption of the Vive caused by a tracking lost “Fig. 7A”. Furthermore, the results showed that Vive system could take 3 seconds after recovering the tracking “Fig 7B”. During stable situation the dispersion was 0.0112 ± 0.097 mm but increase during the few seconds after recovering the tracking (3.2 ± 2 mm during the first second; 5.2 ± 2.3 mm during the second; 2.2 ± 1.6 mm during the third second) and return to normal after the third second (0.113 ± 0.09 mm).

4. DISCUSSION

Utilization of two methods to compute the passage matrix allowed us to estimate an accuracy around the centimeter for positioning the tracking object by Vive system for both static and dynamic conditions, regardless of algorithm influence. Even if method 1 used by Niehorster *et al.*'s study [5] shows better performance, method 2 provides greater confidence in result validation. We thus preconize using the latter to estimate whether the Vive system is relevant for a protocol. Furthermore, both methods reveal similar accuracy variability across the tracking volume. We cannot

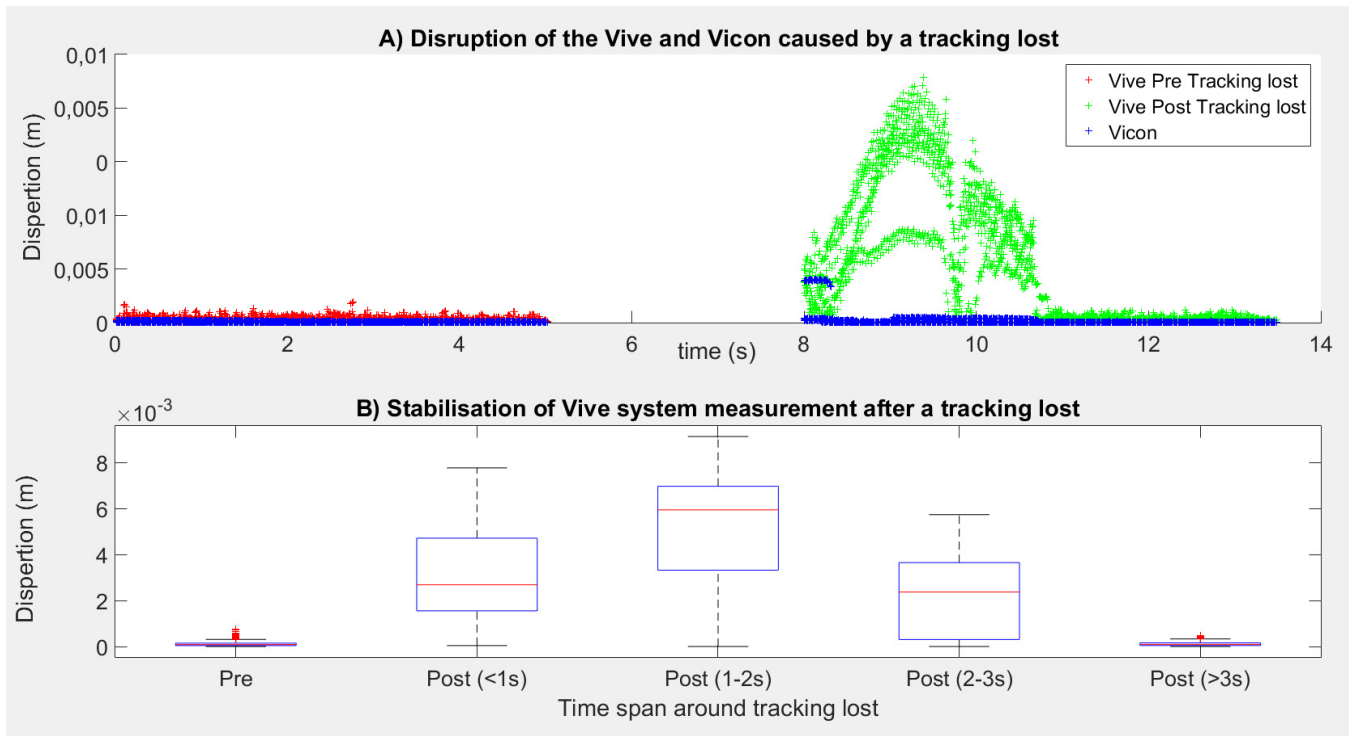


Fig. 7. Representation of the impact of a tracking lost on the Vive measure. **A.** Disruption of the Vive and the VICON system caused by the tracking lost. The blue crosses show that VICON is not disrupted by the tracking lost whereas the green crosses show the Vive tracking system after a tracking lost is not as stable as before a tracking lost (red crosses). **B.** Bow plot showing the Vive system dispersion through the time before and after the disruption caused by a tracking lost: Pre: $N = 2772$, dispersion = 0.112 ± 0.097 mm; Post (< 1 s): $N = 450$, dispersion = 3.2 ± 2 mm; Post (1-2 s): $N = 450$, dispersion = 5.2 ± 2.3 mm; Post (2-3 s): $N = 450$, dispersion = 2.2 ± 1.6 mm; Post (> 3 s): $N = 1032$, dispersion = 0.113 ± 0.09 mm. The Vive system takes about 3 s to stabilize after a disruption caused by a tracking lost.

fully exclude the possibility that NIR from BONITA 10 cameras might have perturbed the Vive system. We deem this occurrence unlikely, as other studies similarly reported poorer performance at the borders, as compared to the center of the tracking volume [4,5].

The ensemble of results from the three experiments converge in showing two main findings. First, quite obviously, the Vive system bears less accuracy and precision than the VICON. On purpose, our superior device was used as a benchmark testing tool. Second, and most important: when tested in either static, or in those dynamic conditions that are compatible with typical human behavior (such as reaching or pointing objects in the near space), the Vive device allows for fairly accurate and reliable spatiotemporal measurements. In summary, although not originally conceived for research purposes, the Vive allows for centimeter accuracy. This sets the lower threshold for use of this device for research. The dispersion values fit well what one would need to detect relatively small differences in performance in cases such as manually pointing to virtual objects, visually congruent or displaced by some degrees with respect to the environment. In future studies it would be

interesting to extend the performance analysis of the Vive system to a wider range of speeds. Recent work using HMD in similar situations has indeed brought valuable results both in basic [9-11] and clinical research [12]. Besides vision, innovative applications of HMD and manual tracking devices are under development [13] to map human's ability in localizing free-field sounds [14] in auditory perception. The present testing of the Vive device's reliability support the feasibility of its use for mapping 3D sound perception both in normal listeners and hearing impaired, or cochlear implant patients [14].

5. ACKNOWLEDGEMENTS

The study has been supported by ANR-10-IBHU-0003 Neuro-immersion, the ANR 16-CE17-0016-01 VirtualHearing3D, the ANR 16-CE28-0015-01 Dev-Tool and has been performed in the framework of the LABEX CORTEX (ANR-11-LABX-0042) A. Farnè was supported by the James S. McDonnell Scholar award.

6. REFERENCES

- [1] C. Chun-Ming, H. Cheng-Hsin, H. Chih-Fan, and C. Kuan-Ta. Performance measurements of Virtual Reality systems: Quantifying the timing and positioning accuracy. In *Proceedings of the 24th ACM international conference on Multimedia*, pages 655-659. 2016.
- [2] A. Malventano. SteamVR HTC Vive In-depth – Lighthouse Tracking System Dissected and Explored. *PC Perspective*. 2016.
- [3] Skarredghost. All you need to know about SteamVR Tracking 2.0... will it be the foundation of Vive 2? *The Ghost Howls*. 2017.
- [4] A. Borrego, J. Latorre, M. Alcañiz, and R. Llorens. Comparison of Oculus Rift and HTC Vive: Feasibility for virtual Reality-Based Exploration, Navigation, Exergaming, and Rehabilitation. In *Games Health J.*, 7:151-156. January 2018.
- [5] D.C. Niehorster, L. Li, and M. Lappe. The Accuracy and Precision of position and orientation tracking in the HTC Vive Virtual Reality system for scientific research. In *Iperception*, volume 8, 3 2041669517708205. May 2017.
- [6] M. Windolf, N. Götzen, and M. Morlock. Systematic accuracy and precision analysis of the video motion capturing systems – Exemplified on the Vicon-460 system. In *Journal of biomechanics*, volume 2, pages 2776-2780. 2008.
- [7] P. Merriaux, Y. Dupuis, R. Boutteau, P. Vasseur, and X. Savatier. A study of Vicon system positioning performance. *Sensors*, 17:1591-1609. July 2017.
- [8] K.S. Arun, T.S. Huang, and S.D. Blostein. Least-Squares Fitting of Two 3-D Point Sets. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 9:698-700, May 1987.
- [9] J.M. Anglin, T. Sugiyama, and S.L. Liew. Visuomotor adaptation in head-mounted virtual reality versus conventional training. *Scientific Reports*, 4;7:45469, April 2017.
- [10] R. Gammeri, F. Turri, R. Ricci, and R. Ptak. Adaptation to virtual prisms and its relevance for neglect rehabilitation: a single-blind dose-response study with healthy participants. *Neuropsychol Rehabil*, 24:1-14. July 2018.
- [11] A.A. Ramos, E.C. Hørning, and I.L. Wilms. Simulated prism exposure in immersed virtual reality produces larger prismatic after-effects than standard prism exposure in healthy subjects. *PLoS One*, 24:14-19. May 2019
- [12] K. Yasuda, D. Muroi, M. Ohira, and H. Iwata. Validation of an immersive virtual reality system for training near and far space neglect in individuals with stroke: a pilot study. *Topics in Stroke Rehabilitation*, 24:533-538. 2017.
- [13] R. Salemme, A. Farnè, V. Gaveau, A. Belle, E. Koun, and F. Pavani. Methods and Apparatus for Acquiring a Spatial Map of Auditory Perception of a Subject. *European patent n°16305621.1*. January 2019.
- [14] V. Gaveau, F. Pavani, A. Coudert, A. Belle, M. Venturini, E. Koun, R. Salemme, C. Desoche, E. Tury, and A. Farnè. Active sound localisation in 3D space in unilateral and bilateral cochlear implant users: a new approach to spatial hearing assessment through immersive virtual reality. *10th International Symposium on Objective Mesasures in Auditory Implants*, October 2018, Tel Aviv (Israël).