



HAL
open science

Evaluation fondamentale des performances de localisation sonore de patients implantés cochléaires et mise en place d'une rééducation ciblée de la localisation dans cette population

Aurélie Coudert

► **To cite this version:**

Aurélie Coudert. Evaluation fondamentale des performances de localisation sonore de patients implantés cochléaires et mise en place d'une rééducation ciblée de la localisation dans cette population. Médecine humaine et pathologie. Université de Lyon, 2021. Français. NNT : 2021LYSE1211 . tel-03644549

HAL Id: tel-03644549

<https://theses.hal.science/tel-03644549>

Submitted on 19 Apr 2022

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

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



N°d'ordre NNT : 2021LYSE1211

THESE de DOCTORAT DE L'UNIVERSITE DE LYON

opérée au sein de
l'Université Claude Bernard Lyon 1

Ecole Doctorale N° 476 - NEUROSCIENCES ET COGNITION
Spécialité de doctorat : Neurosciences

Unité recherche : laboratoire IMPACT, UMR 5292 CNRS /1028 INSERM

Soutenue publiquement le 26/10/2021, par :
Aurélié COUDERT

Evaluation fondamentale des performances de localisation sonore de patients implantés cochléaires et mise en place d'une rééducation ciblée de la localisation dans cette population

Devant le jury composé de :

Pr. Jacques LUAUTE	PU-PH (UCBL1)	Président du jury
Pr. Paul AVAN	Professeur des Universités (UCA)	Rapporteur
Pr. Natacha TEISSIER	PU-PH (Paris 7)	Rapporteure
Pr Pascal BARONE	Professeur des Universités (Toulouse III)	Examineur
Dr. Anne CACLIN	Chercheure (INSERM)	Invitée
Pr. Jacques LUAUTE	PU-PH (UCBL1)	Examineur
Dr. Valérie GAVEAU	Maître de Conférences (UCBL1)	Co-Directrice de thèse
Pr. Eric TRUY	PU-PH (UCBL1)	Directeur de thèse

REMERCIEMENTS

La thèse est un long chemin fait de belles rencontres....

Mes remerciements s'adressent tout d'abord à tous les patients qui ont accepté de m'accorder leur confiance et leur intérêt dans cette démarche scientifique. Nos échanges ont été très riches et j'espère avec humilité que ces travaux de thèse permettront d'apporter un nouveau regard sur la spatialisation auditive.

Je remercie chaleureusement Mme Natacha Teissier et M. Paul Avan d'avoir accepté d'être rapporteurs de ce mémoire.

Je remercie également Mme Anne Caclin, M. Pascal Barone, et M. Jacques Luauté d'avoir accepté d'examiner mon travail.

Cette thèse n'aurait pas eu lieu sans le soutien initial de mes deux co-directeurs de thèse, Mme Valérie Gaveau et M. Eric Truy. Je vous remercie chaleureusement pour votre accompagnement bienveillant. Merci infiniment à la société Cochlear qui n'a pas hésité à me suivre dans cette aventure. Comme tu l'as justement dit Thierry, « les étoiles étaient alignées » pour que ce projet prenne vie.

Une thèse ne se construit pas seule...Valérie et Julie, vous avez été mes deux piliers au cours de ces trois années de recherche. Merci pour votre aide indéfectible à chaque étape de cette thèse.

Je remercie chaleureusement toute l'équipe des implants cochléaires d'Edouard Herriot. Votre intérêt et votre implication active dans la recherche clinique ont été une grande source de motivation pour avancer. Un merci tout particulier à Brigitte, Léa, Kathleen, Stéphanie, Armelle et Jordan. Florence et Marie-Hélène, qu'aurais-je fait sans votre connaissance affûtée de chaque patient ?!

Merci aux collaborateurs du 3^e étage d'HEH, Charles-Alexandre, Pierre-Antoine, et tous les étudiants de passage pour votre participation à ce grand projet.

Je remercie tous les membres de l'équipe IMPACT de m'avoir ouvert les portes des neurosciences. Merci Grégoire d'avoir eu la patience d'implémenter toutes mes requêtes pour

faire évoluer ce très bel outil de spatialisation auditive. Je tenais à vous remercier Karen, Francesco et Alessandro pour vos conseils avisés et constructifs sur mes travaux de recherche.

Je ne serais peut-être pas allée au bout de cette thèse sans le soutien des personnels médicaux et paramédicaux de l'HFME et de HEH. Merci d'avoir suivi avec intérêt et bienveillance mes aventures dans le monde de l'implant.

Le soutien et l'amour de ma famille ont été le moteur de ce grand projet. Mes mots sont peu de choses au regard de toute la reconnaissance et l'affection que je vous porte. Vous m'avez transmis depuis bien longtemps le goût de la recherche et je suis heureuse de partager cette thèse avec vous.

Ma pensée finale sera pour mon rayon de soleil de chaque jour, depuis 20 ans déjà... ma petite Fanny.

RESUME

La spatialisation auditive est un processus complexe qui participe à la construction de notre espace perceptuel. En collaboration avec d'autres afférences sensorielles telle que la vision, elle offre une expérience sensorielle unique. Son fonctionnement optimal passe par l'intégrité des structures auditives captant les sons en périphérie jusqu'à l'extraction et l'interprétation centrale. Toute surdité détériore ces étapes clés du traitement du signal sonore. L'innovation technologique a révolutionné la prise en charge des surdités sévères à profondes grâce aux implants cochléaires mais de nombreuses questions demeurent quant aux modalités de préservation et de réadaptation de la spatialisation auditive chez les patients implantés. Nos travaux de thèse se portent ainsi sur ces deux facettes de la localisation.

Dans une 1^{ère} étude, nous avons évalué dans les trois dimensions de l'espace les performances de localisation d'enfants bi-implantés cochléaires comparées à un groupe d'enfants normo-entendant. Dans une 2^{ème} étude, nous avons validé une forme courte en français du questionnaire pédiatrique « Speech, Spatial, and Qualities of hearing ». Dans une 3^{ème} étude, nous avons développé et testé la faisabilité d'un protocole d'entraînement de la spatialisation auditive à destination d'adultes bi-implantés cochléaires.

Chacune des études menées a permis de mieux caractériser les compétences spatiales dans chaque population testée. Les patients implantés conservent d'importantes difficultés soumises à variabilité inter-individuelle. Mais la mise en place d'un entraînement spatial apparaît comme une perspective d'avenir dans la modulation de ces difficultés.

Mots clés : spatialisation auditive, implant cochléaire, enfants, maturation auditive, rééducation spatiale

ABSTRACT

Spatial hearing is a complex process that contributes to the construction of our perceptual space. Together with other sensory modalities, like vision, it provides a unified sensory experience. Its optimal functioning relies on the integrity of the auditory structures that pick up sounds in the periphery as well as central regions involved in their extraction and interpretation. Deafness interferes with these key stages of sound signal processing.

Technological innovation has transformed the management of severe to profound deafness thanks to the development of cochlear implants, but many questions remain including how to preserve and rehabilitate spatial hearing in cochlear implant users. The work presented in this PhD thesis focuses on these two aspects of sound localization.

In a first study, we evaluated 3D sound localization performance of children with bilateral cochlear implants and compared it with that of normal hearing children. In a second study, we validated a short French form of the pediatric questionnaire "Speech, Spatial, and Qualities of hearing". In a third study, we developed and tested the feasibility of a spatial hearing training protocol for bilateral cochlear implant adults.

These studies allowed us to better characterize spatial hearing skills in cochlear implant users. Patients had significant difficulties and there was considerable inter-individual variability. Despite this, the development of spatial hearing training to improve individual performance is a promising perspective for the future.

Key words: spatial hearing, cochlear implant, children, auditory maturation, spatial rehabilitation

Table des matières

REMERCIEMENTS	iii
RESUME.....	vi
ABSTRACT	vii
REVUE DU CONTEXTE SCIENTIFIQUE.....	1
Chapitre 1 – Physiologie de la spatialisation auditive.....	2
1. Généralités	3
2. Indices acoustiques binauraux et monauraux	5
2.1. Indices binauraux	6
2.2. Indices monauraux	7
3. Quels indices pour quels espaces ?	8
3.1. Localisation en azimut	8
3.2. Localisation en élévation	9
3.3. Localisation en profondeur	10
4. Intégration centrale des informations périphériques.....	11
4.1. Intégration sous-corticale au niveau du tronc cérébral	11
4.2. Intégration au niveau du cortex auditif	13
5. Intégration multisensorielle : modulation de la localisation spatiale.....	14
5.1. Rôle des informations visuelles	15
5.2. Audition spatiale dynamique : influence des mouvements de tête	17
6. Lien entre spatialisation auditive et compréhension de la parole dans le bruit	20
7. Spécificités pédiatriques de la spatialisation auditive.....	22
8. Points-clés du chapitre	27
Chapitre 2 – Spatialisation auditive et implant cochléaire	28
1. Généralités sur les surdités	29
1.1. Classifications.....	29
1.2. Période critique et privation auditive chez l’enfant	30
2. Généralités sur l’implant cochléaire	31
2.1. Différentes parties de l’implant	31
2.2. Principes de codage du son par l’implant	32
3. Limites technologiques de l’implant.....	33
3.1. Dégradation des informations binaurales.....	33

3.1.1. Codage très limité de l'ITD	34
3.1.2. Compression de l'intensité des sons	35
3.1.3. Absence de fusion des indices binauraux	36
3.2. Dégradation des informations monaurales.....	37
4. Spécificités de l'implant en pédiatrie : quand immaturité et asymétrie auditives s'emmêlent.....	37
5. La spatialisation auditive est-elle au cœur du réglage audioprothétique des implants ?	39
6. Points clés du chapitre	40
Chapitre 3 – Outils d'évaluation et performances de spatialisation auditive.....	41
1. Outils d'évaluation quantitative.....	42
1.1. Localisation spatiale absolue	42
1.2. Spatialisation auditive dynamique.....	45
2. Développement d'un nouvel outil de spatialisation auditive : SPHERE.....	46
2.1. Atouts du système.....	46
2.2. Matériel utilisé	46
2.3. Grands principes du système	47
2.3.1. Tâche de localisation	47
2.3.2. Analyses des données	48
3. Performances de spatialisation auditive des sujets normo-entendant.....	51
4. Performances de spatialisation auditive des patients implantés cochléaires	52
4.1. Adultes mono et bi-implantés cochléaires	52
4.2. Enfants bi-implantés cochléaires	53
5. Evaluation de la qualité de vie.....	54
5.1. Généralités	54
5.2. Questionnaires spécifiques chez l'adulte.....	55
5.3. Questionnaires spécifiques en pédiatrie.....	56
6. Points clés du chapitre	58
Chapitre 4 – Rééducation de la spatialisation auditive	59
1. Grands principes de la rééducation	60
2. Place de la rééducation spatiale dans la prise en charge rééducative des patients implantés cochléaires	61
3. Protocoles d'entraînement de la spatialisation auditive en recherche	62
3.1. Chez des sujets normo-entendant	62
3.2. Chez des sujets atteints de surdit�	65

3.2.1. Patients implantés cochléaires	65
3.2.2. Patients appareillés avec une prothèse conventionnelle	67
4. Quels sont les éléments clés pour optimiser une rééducation de la spatialisation auditive à destination de patients implantés ?	69
4.1. Lieu de l'entraînement	69
4.2. Durée et fréquence de l'entraînement	71
4.3. Nature des stimuli délivrés	72
4.3.1. Phase de recherche du son	72
4.3.2. Feedback	74
4.4. Type d'espace rééduqué	75
5. Points clés du chapitre	77
TRAVAUX PERSONNELS DE RECHERCHE	78
Etude 1 – SPHERIC – Etude de la spatialisation auditive chez des enfants bi-implantés cochléaires	79
Partie 1 : Résumé de l'étude	80
1. Questions soulevées au regard du contexte scientifique	80
2. Justification du design de l'étude	81
3. Principaux Résultats	82
Partie 2 : Discussion générale de l'étude SPHERIC	83
1. Des performances de localisation limitées en condition « tête immobile »	83
2. Comparaison des capacités de spatialisation auditive de l'enfant et de l'adulte bi-implantés cochléaires	84
3. Impact de la bilatéralisation séquentielle retardée	86
4. Poursuite de la maturation auditive jusqu'à l'adolescence	88
5. Perspectives	89
Partie 3 : Article relatif à l'étude SPHERIC (<i>publié</i>)	91
Etude 2 – KID SSQ12 – Développement et validation d'une forme courte du questionnaire pédiatrique SSQ	105
1. Introduction	132
2. Matériel et méthode	132
3. Résultats	132
4. Discussion	132

Etude 3 – RECOVER – Rééducation de l’espace auditif 3D chez des adultes bi-implantés cochléaires	131
Partie 1 : Résumé de l’étude.....	132
1. Questions soulevées au regard du contexte scientifique.....	132
2. Justification du design de l’étude.....	132
3. Résumé de l’analyse des données	134
4. Principaux résultats.....	134
Partie 2 : Discussion générale de l’étude RECOVER.....	136
1. Bénéfice de l’entraînement centré sur la discrimination en azimut.....	136
2. Généralisation des effets de l’entraînement spatial	137
3. Variabilité des performances	138
4. Poids des interactions multisensorielles dans la rééducation.....	139
Partie 3 : Article relatif à l’étude RECOVER (<i>en révision</i>).....	141
 PERSPECTIVES GENERALES	 179
 REFERENCES.....	 182
 ANNEXES	 205
Annexe A : Validation du système VIVE comme outil de localisation spatiale.....	206
Annexe B : Active Hearing 3D	214

LISTE DES ABREVIATIONS

3D : tri-dimensionnel
BCI : implant cochléaire bilatéral
CAE : conduit auditif externe
HP : haut-parleur(s)
HRTF : head related transfer function
IC : implant cochléaire / implanté(s) cochléaire(s)
ILD : interaural level difference (différence inter-aurale d'intensité)
ITD : interaural time difference (différence inter-aurale de temps)
NE : normo-entendant
OD : oreille droite
OG : oreille gauche
OSL : olive supérieure latérale
OSM : olive supérieure médiane
PA : prothèse auditive
PEA : potentiels évoqués auditifs
RV : réalité virtuelle
SA : spatialisation auditive
SSQ : Speech, Spatial, and Qualities of hearing
UCI : implant cochléaire unilatéral
V : visuel
VA : visuo-auditif

REVUE DU CONTEXTE SCIENTIFIQUE

Chapitre 1

Physiologie de la spatialisation auditive

1. Généralités

Le système auditif humain supporte de nombreuses fonctions à l'origine de notre perception du monde extérieur et de nos interactions avec notre environnement. Citons à ce titre deux fonctions essentielles du quotidien : 1) « qu'est-ce que j'entends ? » (les anglo-saxons appellent ce traitement auditif le « What ? »), 2) « d'où vient le son perçu ? » (la voie du « Where ? »). La localisation de sources sonores, également appelée spatialisation auditive (SA), est un comportement ancestral hautement conservé entre les espèces (Grothe et al. 2010). Elle permet de détecter et d'appréhender rapidement la position d'un prédateur, même au-delà du champ visuel. Les études phylogénétiques ont d'ailleurs démontré que le traitement de la localisation spatiale suivait la même typographie chez les espèces. Nos capacités de localisation regroupent plusieurs facultés. Il est tout d'abord possible d'évaluer la position absolue d'une source sonore, que celle-ci soit statique ou en mouvement. Il est également possible de définir la position de cette source par rapport à une autre, on définit alors sa localisation relative. Dans la suite de notre exposé, nous nous intéresserons à la localisation absolue d'une source sonore dans l'espace tridimensionnel (3D).

De façon remarquable, l'espèce humaine est la plus précise en matière de performances en localisation absolue. Comme l'a montré Heffner (2004), il existe une balance entre les capacités du système visuel et celles du système auditif. Le champ visuel de l'Homme étant limité par rapport à celui d'autres espèces, ses capacités de localisation visuelle sont donc drastiquement diminuées, raison pour laquelle le système auditif vient contrebalancer cette faiblesse (Figure I.1).

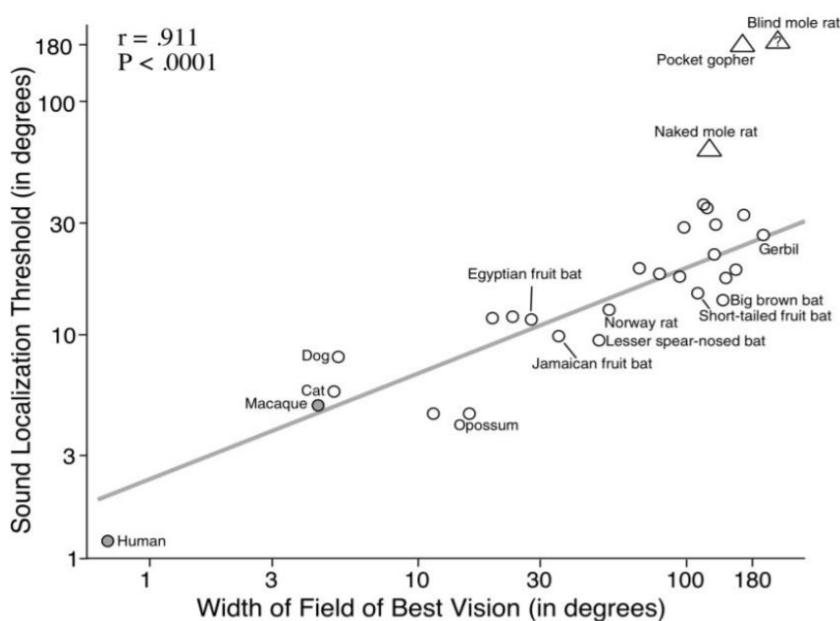


Figure I.1 : Capacités de localisation auditive et visuelle suivant les espèces, d'après Heffner 2004.

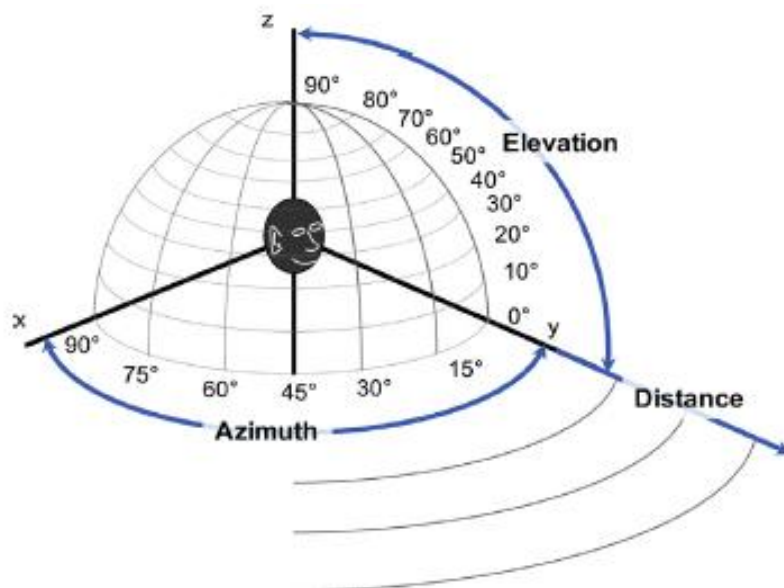
Au fil de l'évolution, le système de SA s'est avéré très utile dans les milieux bruyants pour augmenter le confort d'écoute. En effet, dans ces situations d'écoute difficile, il est important de pouvoir différencier les différentes sources sonores pour focaliser notre attention sur la source d'intérêt ; ce phénomène s'appelle le démasquage binaural (voir paragraphe 6).

Localiser un son nécessite de situer cette source dans l'espace. La position de cette source peut être représentée dans un système de coordonnées polaires dont l'origine est le centre de la tête de l'auditeur et définie par 3 coordonnées (2 angles et 1 distance, voir Figure I.2) :

- Position en **azimut** : elle s'obtient en projetant la position 3D de la source sonore sur le plan horizontal ($0 \pm 180^\circ$), permettant de situer un son provenant de notre héli-espace auditif gauche ou droit ; ou bien encore situé dans l'espace devant ou derrière nous.

- Position en **élévation** : elle s'obtient en projetant dans le plan transverse ($0 \pm 90^\circ$) la position sonore, permettant de localiser une source venant d'en haut ou d'en bas (par rapport à un plan horizontal passant par l'axe interaural).

- Position en **profondeur** ($0 \text{ cm} + \infty$) : elle permet d'estimer la distance de la source et de définir deux espaces – l'espace proche (à moins de 1 mètre de soi) ou l'espace lointain (à plus de 1 mètre).



Figures I.2 : Localisation spatiale en 3D

Coordonnées polaires utilisées pour définir la localisation spatiale d'une source sonore dans un repère orthonormé x, y, z dont l'origine est l'auditeur : azimut dans le plan horizontal, hauteur dans le plan vertical et distance en profondeur (extrait de Risoud et al, 2018)

De façon consensuelle dans la littérature, et pour la suite de cette thèse, nous considérerons que les sources localisées dans le plan horizontal et incluses dans un plan sagittal passant par le centre de la tête se trouvent à 0° droit devant et à 180° droit derrière. Selon un plan frontal passant par l'axe inter-aural, les sources situées à gauche sont à -90° et à droite à $+90^\circ$. Le plan axial passant à hauteur d'oreille définit la référence de 0° (Figure I.3) et permet de discriminer les sources à différentes élévations (vers le haut, les sources sont définies avec une angulation positive ; vers le bas avec une angulation négative).

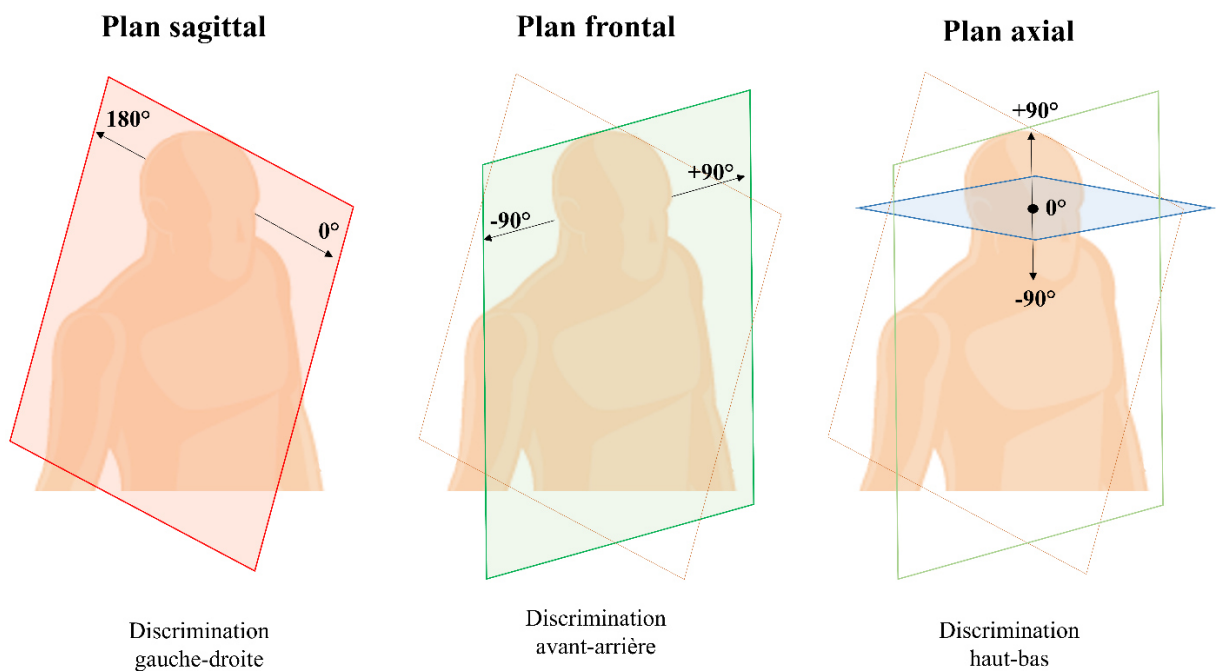


Figure I.3 : Définition des 3 plans sagittal, frontal et axial permettant de discriminer les sons en azimut et en élévation.

Dans ce chapitre, nous nous attacherons à développer les différents indices acoustiques utilisés dans la SA et pour chacun d'entre eux, leur rôle spécifique dans la localisation dans chaque plan de l'espace.

2. Indices acoustiques binauraux et monauraux

Lors de l'émission d'un son par une source, celui-ci subit un certain nombre de transformations acoustiques avant d'être transduit par le système auditif. Ces modifications

acoustiques dépendent des caractéristiques anatomiques de la tête, et sont notamment induites par l'effet de cône d'ombre de la tête et l'effet pavillonnaire (i.e. effet induit par le pavillon et le conduit auditif externe); ces modifications sont à l'origine d'indices acoustiques binauraux et monauraux.

2.1. Indices binauraux

La notion d'indices binauraux est un concept relativement récent, évoqué au début du XXe siècle par Lord Rayleigh. Lorsqu'une source sonore n'est pas située droit devant (i.e. à 0°) ou droit derrière le sujet (i.e. à 180°), celle-ci est perçue par chacune de nos oreilles avec un différentiel de temps et d'intensité. Ces différences sont ensuite analysées et comparées par les structures auditives centrales ce qui sert de base au développement et la mise à jour de cartes spatiales.

Prenons l'exemple de la Figure I.4, où la source sonore est située à gauche du sujet. Le son parvient à l'oreille droite avec un léger retard dans le temps. Ce délai est à l'origine d'un premier indice binaural, appelé différence inter-aurale de temps (ITD dans la littérature anglo-saxonne, pour Interaural Time Difference). Chez l'homme, la tête représente un obstacle acoustique pour des sons de haute fréquence, dont la longueur d'onde est trop petite pour la contourner. L'onde acoustique est également atténuée par l'épaisseur de la tête avant de parvenir à l'oreille droite ; l'effet du cône d'ombre de la tête est particulièrement important pour les hautes fréquences (Opstal 2016). Cette atténuation est à l'origine d'une réduction de l'intensité de l'onde, générant ainsi le deuxième indice binaural, appelé différence inter-aurale d'intensité (ILD dans la littérature anglosaxonne, pour Interaural Level Difference).

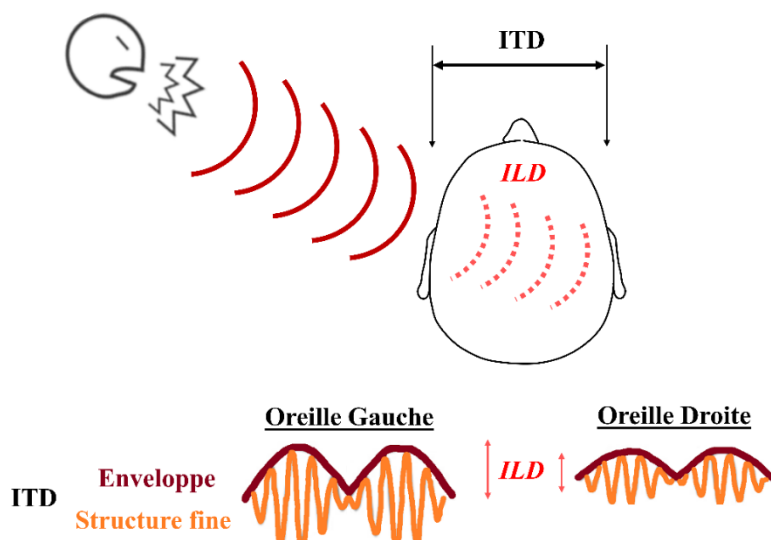


Figure I.4 : Schématisation des indices binauraux captés en périphérie

La différence inter-aurale de temps (ITD) se caractérise par l'enveloppe du signal (en marron) et sa structure fine (en orange).

La différence inter-aurale d'intensité (ILD) est représentée par la différence d'amplitude du signal arrivant à chaque oreille.

L'oreille humaine est capable de percevoir des différentiels de temps et d'intensité très faibles, de l'ordre de $10 \mu\text{s}$ pour l'ITD et de 1 dB pour l'ILD (Klumpp & Eady 1956, Mills 1958). La perception de ces indices binauraux est maximale pour certaines positions de sources sonores dans l'espace. Ainsi, ITD et ILD seront les plus importantes pour une source située à $\pm 90^\circ$ de sujet dans le plan azimutal, donc latéralement à sa gauche ou à sa droite (Hartmann 1999, voir Figure I.5).

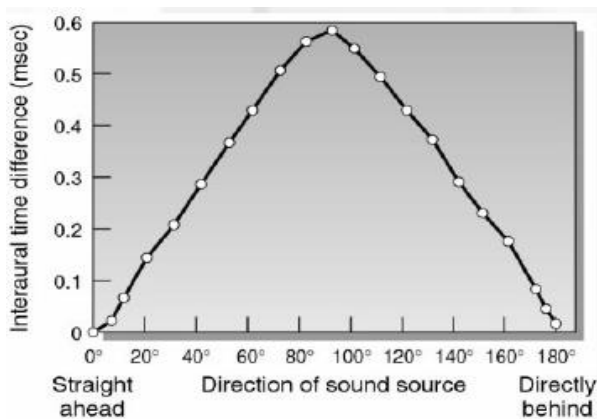


Figure I.5 : Différence inter-aurale de temps en fonction de la position azimutale du son

2.2. Indices monauraux

Les indices monauraux sont générés par des signaux arrivant à chaque oreille séparément. Le pavillon de l'oreille, la tête et dans une moindre mesure le buste jouent le rôle de filtre au passage d'une onde acoustique et modifient sa forme spectrale en fonction de son angle d'incidence et de sa fréquence. Ces changements peuvent être de 3 natures : diffraction de l'onde, absorption, ou réflexion ; ils se regroupent sous le terme de fonctions de transfert liés à la tête (en anglais head - related transfer functions – HRTF).

L'ensemble des indices acoustiques générés (binauraux et monauraux) ne sont pas codés dans un référentiel spatial directement en périphérie au niveau de l'oreille interne. En effet, il n'existe pas de système de cartographie spatiale précis au sein de la cochlée, à l'image du système visuel et de la rétine (e.g. la rétinopathie est une représentation spatiale de l'environnement visuel à travers l'organisation anatomique du système visuel, mais pour le système auditif, cette représentation est fréquentielle, c'est la tonotopie cochléaire). Au contraire, la SA fait intervenir différents effecteurs centraux, le premier relai ayant lieu dans le tronc cérébral (voir paragraphe 4.1).

3. Quels indices pour quels espaces ?

3.1. Localisation en azimut

Évaluer la directionalité d'un son, c'est-à-dire sa position projetée dans un plan azimutal, nécessite l'extraction et l'analyse correcte des indices binauraux. ITD et ILD sont rarement impliquées à part égale dans cette tâche. Il a ainsi été démontré que pour des sons de basses fréquences, l'ITD était la plus sensible, alors que pour les hautes fréquences, c'est l'ILD qui est majoritairement impliquée. Cela tient en partie au phénomène de cône d'ombre de la tête que nous avons mentionné précédemment. En effet, les ondes incidentes de hautes fréquences subissent d'importants phénomènes d'absorption et de réfraction par la tête, contrairement aux basses fréquences qui sont peu atténuées. Cela maximise les différences d'intensité perçues entre les 2 oreilles. Il est donc admis que les variations d'ILD sont plus informatives pour les hautes fréquences, au-delà de 1500 Hz (Mills 1958).

Par la suite, afin de mieux comprendre les difficultés de localisation éprouvées par des sujets malentendants, il convient de différencier deux entités au sein de l'ITD. Un son complexe se caractérise par sa structure temporelle fine (portée par les basses fréquences) et par son enveloppe (portée par les hautes fréquences). Ces deux entités de l'ITD véhiculent des informations différentes mais complémentaires pour la bonne localisation des sons. Globalement, l'ITD est l'indice binaural fondamental pour localiser des sons purs de basses fréquences, inférieures à 1500 Hz (Mills 1958). Pour des sons plus complexes à large spectre fréquentiel, l'ITD reste exploitable même pour des sons plus aigus grâce aux caractéristiques de son enveloppe.

De manière générale, la précision en azimut est meilleure pour les sources frontales que pour les sources arrières (Figure I.6), respectivement 3° et 6° (Blauert 1997). Mais elle peut varier de façon importante en fonction de la nature du son : la largeur du spectre fréquentiel impacte la précision en azimut (Butler 1986). Un bruit blanc, dont la gamme fréquentielle est très étendue offre une meilleure localisation qu'un bruit à spectre plus étroit (< 1 octave) ou à l'extrême un son pur. Plus un son pur est grave (< 1000 Hz), meilleure sera la SA (Stevens & Newman 1936, Mills 1958, Yost & Zhong 2014). A contrario, la gamme fréquentielle comprise entre 1000 Hz et 3000 Hz est la plus délétère car l'accès aux indices binauraux est très restreint, que ce soit les ILD ou les ITD (Risoud et al. 2018). L'intensité de la source sonore n'influe pas la précision de la localisation en azimut (Yost & Zhong 2014).

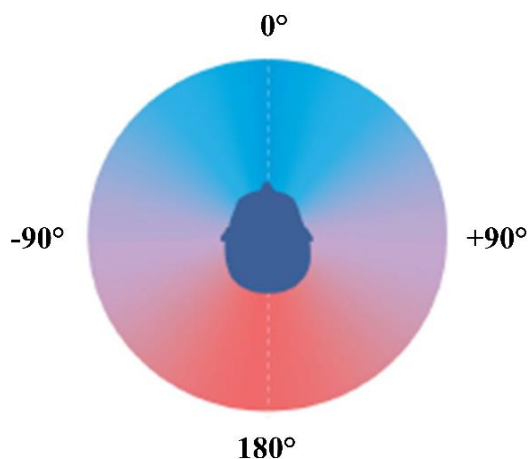


Figure I.6 : Acuité du système de localisation humain en azimut (d'après van der Heijden et al. 2019). La précision en azimut est meilleure pour les sources frontales (à 0°, matérialisées en bleu) et se détériore progressivement pour les sources latérales ($\pm 90^\circ$, en violet) et à l'arrière du sujet (à 180°, en rouge).

3.2. Localisation en élévation

Contrairement au plan azimutal, les indices binauraux ITD et ILD sont peu informatifs pour localiser en élévation. Angell et Fite (1901a,b) ont été les premiers à décrire le rôle des indices monauraux dans la SA. L'ensemble des indices spectraux générés par l'effet pavillonnaire est propre à la morphologie de chaque individu (Batteau 1967, Musicant & Butler 1984) et évolue avec la croissance du sujet. Au fil des années, le cerveau doit ainsi encoder une multitude d'HRTF pour différents types de sons, de positions données. Il existe donc un processus d'apprentissage dynamique de localisation sonore, particulièrement dans le plan vertical.

Au regard de l'effet important du pavillon sur les signaux de fréquence aiguë, il apparaît qu'un son à large spectre composé de fréquences supérieures à 7000 Hz est plus facile à localiser en élévation qu'un son à bande étroite centrée sur des fréquences plus graves (Butler & Humanski 1992, Roffler & Butler 1968). En l'absence de cet effet pavillonnaire, les indices spectraux se retrouvent dégradés et les HRTF ne reposent plus que sur le filtre effectué par le tronc. Ceci a toute son importance dans le cadre de surdité.

De manière générale, la précision du système auditif en élévation est bien moindre qu'en azimut (erreur absolue autour de 13° contre 6° en azimut, Tabry et al. 2013).

3.3. Localisation en profondeur

L'estimation de la distance d'une source sonore répond à un besoin primaire de protection face à un éventuel danger : cette source est-elle dans mon espace proche péri-personnel, c'est-à-dire à moins de 1 mètre et peut-elle mettre en jeu mon intégrité physique ? Ou bien est-elle située dans un espace plus lointain, à plus de 1 mètre ? Cette tâche est effectuée quasi-quotidiennement lorsque nous traversons une rue animée sans forcément regarder la route en première intention. Les variations de l'intensité sonore d'un véhicule en mouvement nous permettent d'estimer si celui-ci se rapproche (i.e. l'intensité augmente) ou s'éloigne de nous (i.e. l'intensité diminue). Ces variations ne sont pas les seuls indices exploitables pour évaluer la distance d'une source sonore. En effet, dans des espaces réduits où les sources sont très proches, les variations d'ILD sont les plus utiles pour localiser (Brungart et al. 1999a). A contrario, dans des espaces plus vastes et échoïques, la réverbération du son devient un indice acoustique précieux où le sujet reçoit deux informations : l'onde acoustique directe suivie des ondes réfléchies (Figure I.7). L'onde directe est la plus pertinente pour localiser dans des espaces de taille modeste (e.g. pièce de 10 m²) alors que les ondes réfléchies sont majoritairement exploitées dans de plus vastes espaces où le temps d'arrivée de l'onde réfléchiée par rapport à l'onde directe est supérieur à 30 millisecondes (e.g. supermarché) (Garas 2000).

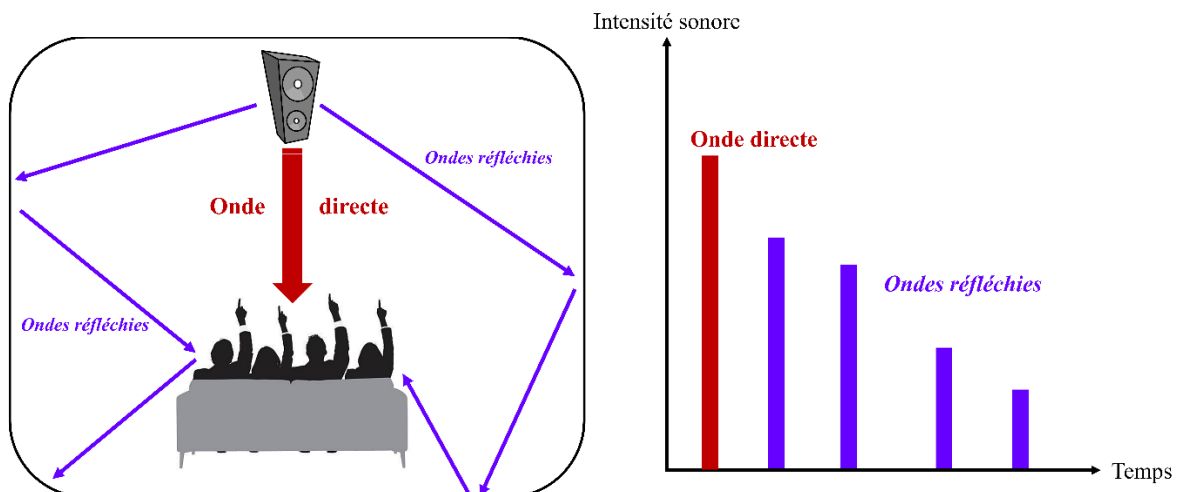


Figure I.7 : Principe de la réverbération d'un son dans l'appréciation de la distance d'un son. L'onde directe parvient en premier à l'auditeur avec une intensité maximale, elle n'est pas réfractée par les parois de la pièce contrairement aux ondes réfléchies qui arrivent plus tardivement et avec une moindre intensité.

De manière générale, les distances proches sont souvent surestimées et celles qui sont lointaines sont sous-estimées (Zahorik 2002). Evaluer avec précision la distance d'une source sonore statique s'avère plus complexe que d'estimer sa directionnalité, mais certains facteurs tendent à augmenter ce degré de précision, tels que les sons familiers ou la présence d'une pièce échoïque.

4. Intégration centrale des informations périphériques

4.1. Intégration sous-corticale au niveau du tronc cérébral

Le système auditif possède, de la cochlée en périphérie, jusqu'aux structures centrales, une représentation fréquentielle de l'espace auditif, c'est la tonotopie. La tonotopie cochléaire se projette sur toutes les structures centrales, y compris le cortex auditif primaire. Mais contrairement à la rétine, la cartographie spatiale de l'espace n'est pas représentée en périphérie. Celle-ci implique la mise en jeu de structures centrales qui, à différents niveaux, vont intervenir en comparant et en intégrant les indices acoustiques provenant de chaque oreille (Figure I.8). Les informations acoustiques sont véhiculées parallèlement, de chaque nerf auditif vers un premier relai central, les noyaux cochléaires. Au sein de la moelle allongée, chaque fibre du nerf auditif se ramifie pour envoyer 3 branches principales :

- Une branche ascendante vers le noyau cochléaire ventral antérieur
- Deux branches descendantes vers les noyaux cochléaires ventral postérieur et dorsal
(ceux-ci ne sont pas impliqués dans le processus de localisation)

Chaque noyau cochléaire a une fonctionnalité différente et conserve l'organisation tonotopique présente au sein de la cochlée (Purves et al. 2005).

De manière simplifiée, le noyau cochléaire ventral antérieur analyse spécifiquement les informations liées à l'ITD et envoie des projections nerveuses vers l'olive supérieure médiane (OSM). Les cellules latérales de l'OSM reçoivent des informations ipsilatérales tandis que les cellules médianes reçoivent des informations controlatérales. Les cellules de l'OSM agissent comme des détecteurs de coïncidence qui s'activent lorsque les stimuli ipsilatéraux et controlatéraux arrivent en même temps (Figure I.9). Chaque cellule a donc une sensibilité maximale spécifique de chaque ITD et donc d'une position donnée dans l'espace.

Ce mécanisme d'intégration centrale n'est valable que pour des basses fréquences, allant jusqu'à 3 000 Hz.

L'intégration des informations binaurales liées à l'ILD se fait d'abord au niveau de cellules spécifiques du noyau cochléaire ventral antérieur. Celui-ci envoie ensuite des projections ipsilatérales excitatrices et controlatérales inhibitrices (via des interneurons inhibiteurs du noyau médian du corps trapézoïde) vers l'olive supérieure latérale (OSL). Ainsi, pour une source sonore de 5 000 Hz située à la gauche du sujet, l'OSL gauche sera activée tandis que l'OSL droite sera inhibée. L'activation de l'OSL ipsilatérale à la stimulation auditive est d'autant plus importante que la source sonore est située latéralement, car l'ILD est alors maximale. Sachant que chaque OSL ne code que les sons de l'hémichamp ipsilatéral, l'intégrité des OSL gauche et droite est indispensable pour la représentation de l'ensemble de l'espace azimutal pour les fréquences aiguës.

Les informations monaurales ne transitent pas par le complexe olivaire supérieur et sont directement analysées dans les noyaux du lemnisque latéral controlatéral à la stimulation auditive.

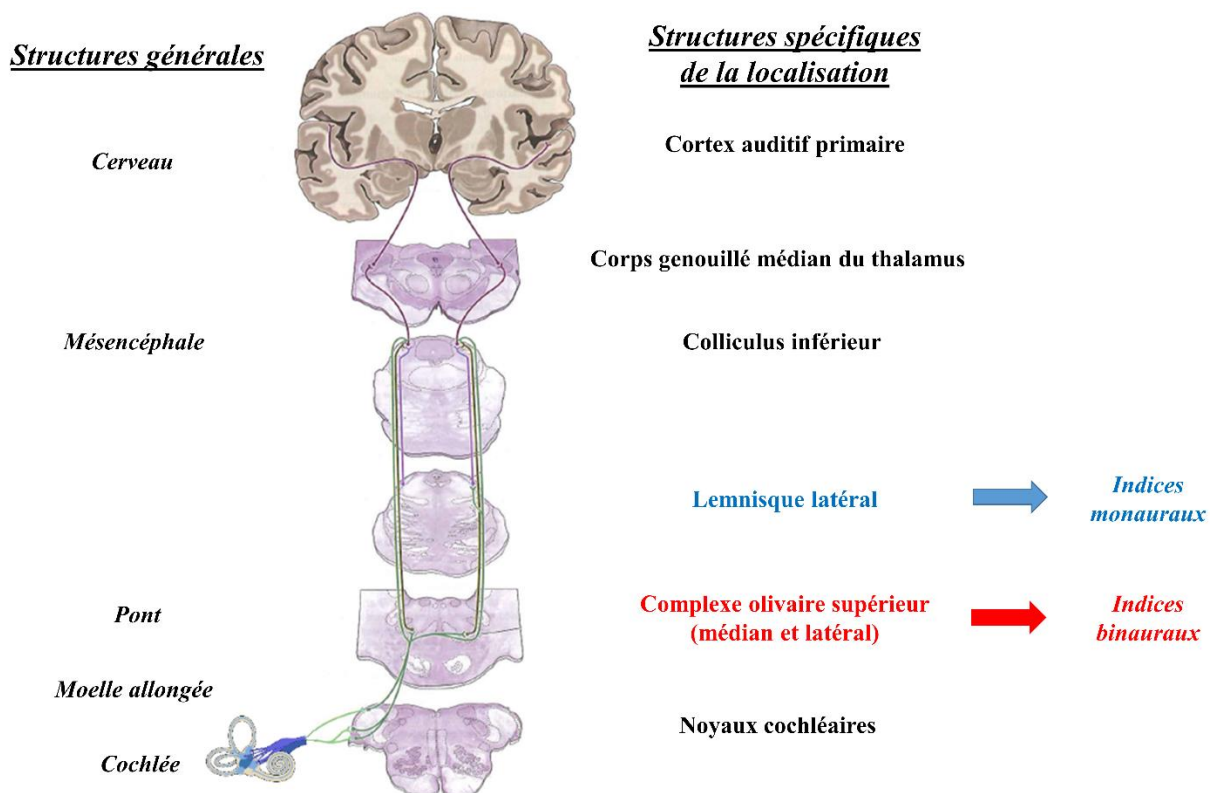


Figure 1.8 : Schéma modifié des principales voies auditives (d'après Purves et al., page 304). Les indices monauraux et binauraux sont traités dans des structures sous-corticales différentes, respectivement le lemnisque latéral et le complexe olivaire supérieur.

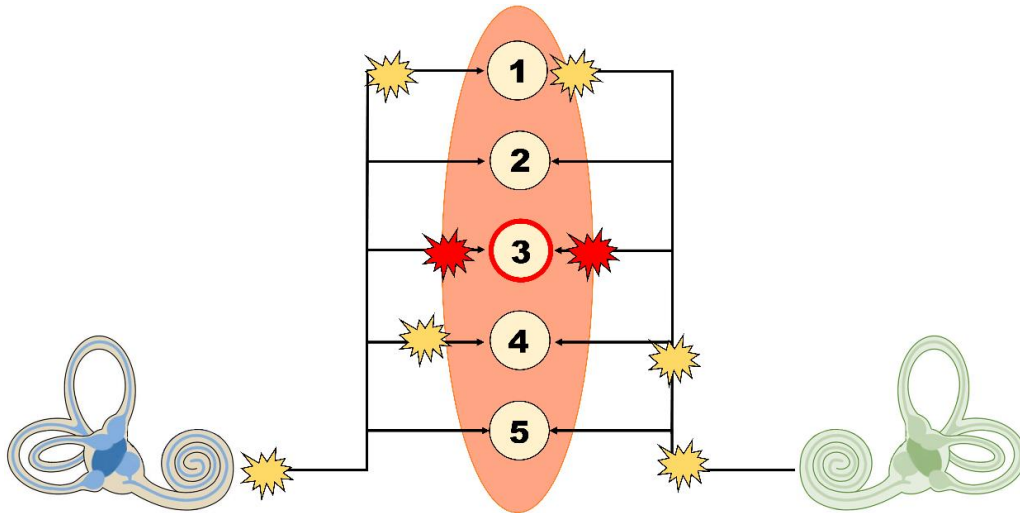


Figure I.9 : Intégration de l'ITD au niveau de l'olive supérieure médiane (schéma adapté de Jeffress 1948).

Chaque neurone de l'olive supérieure médiane (numéroté de 1 à 5) code pour une ITD spécifique. L'activation unitaire d'un neurone se produit lorsque celui-ci reçoit de façon concomitante l'influx nerveux des cochlées gauche et droite (en rouge). L'olive supérieure médiane joue donc le rôle de détecteur de concordance temporelle.

Au total, les informations sensorielles véhiculant l'ITD et l'ILD sont traitées par des centres spécifiques au niveau du pont du tronc cérébral, respectivement dans l'OSM et l'OSL. Les informations monaurales sont analysées juste au-dessus dans le lemnisque latéral. *In fine*, l'ensemble de ces voies auditives se rejoint au niveau du mésencéphale (Figure I.8).

Le centre auditif du mésencéphale est représenté par le colliculus inférieur. Les fonctionnalités de chaque population neurale ne sont pas complètement élucidées chez l'Homme mais si nous prenons l'exemple chez l'animal de la chouette effraie, il apparaît qu'une véritable carte de l'espace auditif se constitue au niveau du colliculus, chaque neurone répondant spécifiquement pour la position d'un son en azimut et en élévation (Purves 2005). Enfin, avant l'intégration corticale, la majorité des voies auditives ayant transité par les colliculi inférieurs fait relai au sein du corps genouillé médian du thalamus dont la spécificité est l'analyse spectrale et temporelle des sons.

4.2. Intégration au niveau du cortex auditif

Le cortex auditif reçoit l'ensemble des voies auditives centrales ascendantes ; il se divise en cortex auditif primaire (aire A1 située dans le gyrus temporal supérieur) et plusieurs aires

auditives secondaires. Chaque aire corticale est spécialisée dans une fonction auditive. Même s'il n'existe pas à proprement parlé de carte de spatialisation au sein du cortex auditif, différents travaux d'imagerie fonctionnelle ont permis de mettre en évidence les « voies auditives dorsales non primaires » responsables du traitement des informations liées à la SA. Celles-ci sont complétées par les « voies auditives ventrales » spécialisées dans l'identification d'un son (e.g. Alain et al. 2001, Arnott et al. 2004, van der Zwaag et al. 2011).

Il est intéressant de noter que des structures corticales non auditives ont été mises en évidence dans les processus de localisation auditive (Degerman et al. 2006, Antunes & Malmierca 2011), notamment les régions frontales et pariétales (Figure I.10). Ces dernières sont recrutées lors de tâches attentionnelles et décisionnaires (e.g. lors de la traversée d'une rue animée, il faut d'abord évaluer la distance à laquelle se trouve la voiture approchant puis, à partir de cet indice spatial, décider de franchir ou pas cette rue).

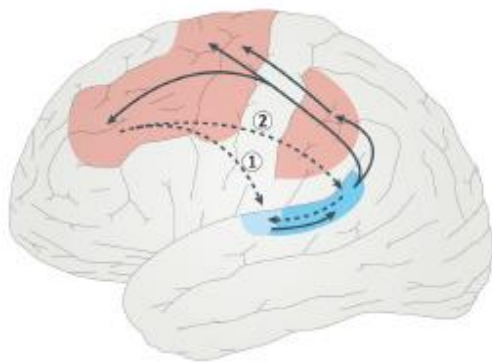


Figure I.10 : Voies auditives corticales dorsales impliquées dans le processus de spatialisation auditive (d'après van der Heijden et al. 2019).

Le cortex auditif primaire (en bleu) est rarement activé isolément dans l'analyse d'une scène auditive. Une interconnexion étroite existe avec les structures corticales préfrontales, voire pariétales (en rouge), celles-ci modulant par feedback (1 et 2) les régions auditives.

5. Intégration multisensorielle : modulation de la localisation spatiale

Au quotidien, nous évoluons activement dans un environnement riche en stimulations sensorielles, nécessitant une mise à jour constante de notre système de SA pour pouvoir adapter nos comportements à une situation donnée. Afin de limiter les ambiguïtés et d'augmenter la précision de ce système, le cerveau intègre, compare et synthétise en temps réel différentes informations sensorielles reçues indépendamment les unes des autres. Ce processus porte le nom d'intégration multisensorielle et mature progressivement au cours de l'enfance selon le degré de développement de chaque système sensoriel.

Dans ce paragraphe, nous nous focaliserons sur le rôle majeur de deux types d'informations sensorielles dans la modulation de la SA : les stimuli visuels et les informations auditives dynamiques induites par les mouvements de tête.

5.1. Rôle des informations visuelles

Plusieurs études réalisées chez l'animal et chez l'Homme ont montré le rôle crucial du système visuel dans le développement de la SA (Knudsen & Knudsen, 1985, Hofman et al., 1998). De manière surprenante, Campus et collaborateurs ont mis en évidence une activation très précoce du cortex visuel chez l'Homme en réponse à une stimulation sonore lorsque le sujet est engagé dans une tâche de localisation spatiale (Campus et al, 2017), reflétant des liens anatomiques fonctionnels entre les cortex auditif et visuel (Falchier et al, 2002 ; Clavagnier et al, 2004).

Dans des situations d'écoute difficile, les informations visuelles contribuent à recalibrer les indices auditifs afin d'harmoniser les correspondances entre les systèmes visuo-auditifs et d'augmenter in fine la précision de la localisation d'une source sonore (Heffner & Heffner 1992, 2014). Ainsi, de manière directe, le système visuel permet de confirmer la position d'un son perçu présent dans le champ visuel (e.g. être interpellé par un individu présent dans une foule). Il a d'ailleurs été démontré que la complémentarité des deux systèmes sensoriels augmentait également la rapidité d'orientation vers la source sonore (Shelton and Searle 1980, Bologni et al. 2007, Tonelli et al. 2015, Valzolgher et al. 2020b), ce qui s'avère particulièrement important en présence d'un danger imminent (e.g. voiture en approche). Dans des environnements plus restreints et réverbérants, l'intégration des informations visuelles offre également un confort d'écoute en supprimant l'effet d'écho induit par les caractéristiques acoustiques du milieu (Bishop et al. 2012).

De façon plus indirecte (i.e. lorsque la source sonore n'est pas visible), le système visuel peut soutenir la SA en permettant de prendre conscience des éléments présents dans l'environnement. Warren a d'ailleurs défini cette fonction sensorielle sous le terme de « facilitation visuelle » (Warren 1970). Au cours de ces travaux princeps, les sujets devaient localiser la position de haut-parleurs (HP) dissimulés derrière un écran selon deux conditions de test, yeux fermés ou yeux ouverts (Warren 1970, Shelton & Searle 1980). Les sujets étaient plus précis dans leurs performances lorsqu'ils avaient les yeux ouverts. A ce stade, il était impossible de faire la différence entre le bénéfice apporté par la perception de la structure même de l'environnement (i.e. dimensions d'une pièce, objets présents dans cette pièce) et par la position de l'écran masquant les HP orientant *de facto* l'attention du sujet vers un espace auditif réduit.

L'intégration multisensorielle et notamment l'influence d'une modalité sensorielle sur une autre, a souvent été étudié en faisant appel à des expériences d'illusions perceptives. Lors d'illusions, deux modalités sensorielles sont présentées simultanément mais dans un rapport conflictuel (chacune donnant une information différente). Ainsi, dans ces situations sensorielles conflictuelles, l'illusion apparaît quand il existe un biais intermodal en faveur d'une stimulation par rapport à une autre. Dans le cas de la SA, l'interaction visuo-auditive peut être étudiée à travers l'effet de la ventriloquie. Cette illusion révèle un biais intermodal en faveur des informations visuelles par rapport aux informations auditives sur la localisation d'une source sonore. Cet effet est exploité par les marionnettistes ventriloques qui font bouger leur marionnette tout en parlant sans bouger les lèvres, et ainsi la source sonore est localisée au niveau de la marionnette qui est animée. Expérimentalement, cet effet est obtenu quand les stimulations visuo-auditives sont délivrées avec un décalage spatio-temporel favorisant l'intégration multisensorielle (voir la revue de Recanzone 2009). De nombreuses études ont montré que l'intégration visuo-auditive dans la SA dépendait du poids attribué à chaque sens (suivant un modèle Bayésien). Ainsi, si les informations visuelles sont peu fiables (car présentées en vision périphérique ou floutées), alors le biais perceptif est plus sujet aux informations auditives (Burr & Alais 2006).

L'intégration récente de la réalité virtuelle dans les expériences de spatialisation a permis de préciser le poids des informations visuelles dans la SA. Lors d'une tâche de localisation auditive, Valzolgher et al. (2020b) ont manipulé l'environnement visuel virtuel en exposant les sujets à deux univers visuels de structures différentes : soit un écran gris (sans structure, Figure I.11A), soit une grille (structure simple, Figure I.11B). La condition de test avec la grille ne permettait pas d'améliorer significativement la précision des sujets mais elle augmentait leur rapidité de localisation. L'équipe de Ahrens (2019) est allée plus loin dans le réalisme de la scène visuelle en incorporant soit la structure de la pièce soit la position exacte de tous les HP (Figure I.11C). La présence des reliefs de la pièce était suffisante pour améliorer significativement les performances des sujets, principalement en azimuth, mais l'ajout des HP dans la scène visuelle apportait un bénéfice supplémentaire pour lever les ambiguïtés résiduelles (Ahrens et al. 2019).

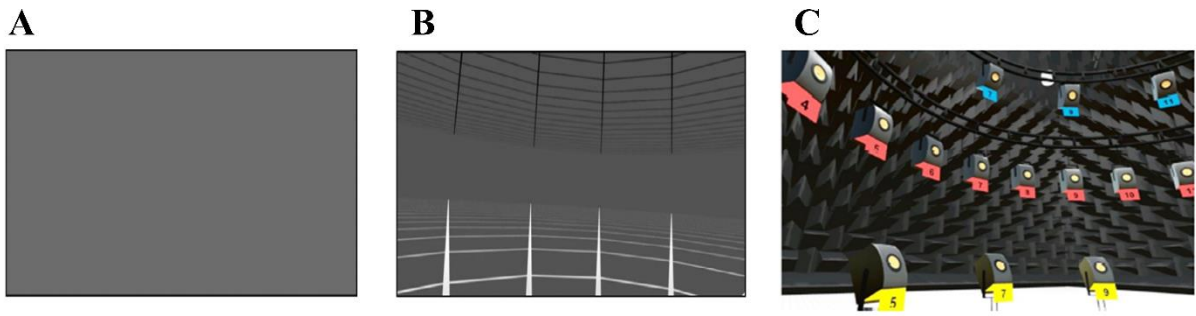


Figure I.11 : Différents environnements visuels utilisés en réalité virtuelle lors d'une tâche de localisation. (A) écran gris uniforme (Valzolgher et al. 2020b), (B) grillage (Valzolgher et al. 2020b), (C) haut-parleurs matérialisés (Ahrens et al. 2019).

Au total, la vision est un système pivot pour soutenir la spatialisation auditive, particulièrement dans les milieux bruyants, mais aussi lorsque le système auditif est déficitaire, comme nous le verrons dans la suite de cet exposé pour les patients implantés cochléaires.

5.2. Audition spatiale dynamique : influence des mouvements de tête

Comme nous l'avons mentionné dans le paragraphe 3.1, l'évaluation de la position d'une source sonore en azimuth est basée sur l'analyse des indices binauraux ITD et ILD. Dans certaines conditions, la source peut être située dans une zone de confusion où il devient très difficile de localiser avec précision l'endroit où celle-ci se trouve lorsque la tête du sujet reste immobile. La Figure I.12 illustre deux situations confusionnelles de discrimination spatiale de 2 HP. Dans chaque situation, les indices binauraux générés sont identiques, seuls les indices spectraux dans la situation A permettent de lever toute ambiguïté positionnelle.

Le moyen le plus efficace pour lever ces ambiguïtés et d'orienter la tête en direction, ou à l'opposé du son afin de faire varier ITD et ILD. Wallach (1940) a été le premier à rapporter le bénéfice des mouvements de tête dans ces situations à l'origine de confusions. Ainsi, les mouvements de tête, par phénomènes de rotation (tête inclinée vers le haut ou le bas, vers la gauche ou vers la droite, Lambert 1974, Perrett & Noble 1997b) génèrent de nouveaux indices acoustiques dynamiques qui facilitent la localisation des sons (Garas 2000).

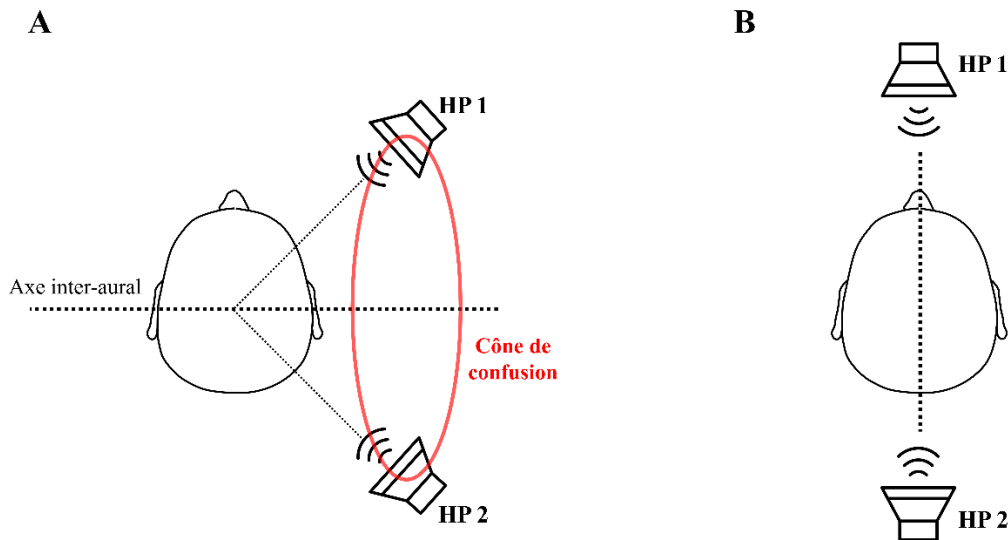


Figure I.12 : Schéma illustrant deux situations tête fixe où les positions des haut-parleurs (HP1 et HP2) ne peuvent pas être discriminées. Dans la situation A, les deux HP se trouvent dans un cône de confusion centré sur l'axe inter-aural (Wallach 1938, Blauert 1997), les indices acoustiques (ITD et ILD) parvenant aux deux oreilles sont identiques (les indices spectraux captés par l'oreille droite peuvent toutefois apporter quelques informations discriminantes). Dans la situation B, les deux HP sont situés sur la ligne médiane et les indices binauraux valent zéro (Wightman & Kistler 1999).

Les mouvements de tête s'inscrivent dans un comportement d'orientation naturelle en direction d'une source sonore (Kim C. 2013) et présentent de nombreux avantages. Tout d'abord, ils permettent de résoudre les situations confusionnelles citées ci-dessus (Figure I.12). En synergie avec le système visuel, ils permettent également de centrer le regard sur la source de sorte que la position du son soit encodée en coordonnées rétino-centriques (Bulkin & Groh 2006, Pavani et al. 2008) et que la source sonore soit définie en vision centrale (si les informations visuelles de la source sont disponibles). Cette association audiovisuelle offre un bénéfice certain dans les environnements multisensoriels complexes (Bologni 2005), et plus particulièrement dans les situations d'écoute difficile (Wightman & Kistler 1999). La compréhension du locuteur devient plus aisée en milieu bruyant en se tournant vers lui et en s'appuyant sur la lecture labiale (e.g. Grant & Seitz 2000, Munhall 2004).

De plus, les mouvements de tête modulent la précision de notre système de localisation pour déterminer avec finesse la position absolue d'une source. Mais le point de vue des auteurs diverge à ce sujet ; certains prônent les bénéfices des mouvements de tête (Wightman & Kistler 1999, Perrett & Noble 1997a,b) tandis que d'autres mettent en lumière une nette détérioration des performances lors de l'écoute dynamique (Cooper 2008, Leung 2018). Il semblerait que la durée du son émis au cours des de la phase de recherche de la source et la vélocité de la tête

soient les deux principaux facteurs dominants. Ainsi, pour des sons suffisamment longs (au moins 2 secondes), les mouvements de tête permettent de réduire significativement les confusions avant-arrière tandis que des sons plus courts (inférieurs à 500 ms) ne permettent pas au système auditif d'encoder avec précision les indices acoustiques dynamiques (Perrett & Noble 1997a, Iwaya 2003).

L'étude de la vitesse de déplacement de la tête (i.e. vélocité céphalique) permet de séparer deux comportements d'orientation : (1) un comportement très rapide (autour de 500°/seconde), réflexe, par exemple face à un danger ; (2) un comportement plus lent, de recherche ciblée. L'objectif des mouvements réflexes rapides n'est pas de localiser avec précision la source sonore mais de cerner instantanément l'espace d'où pourrait provenir un éventuel danger. Cette catégorie de mouvements de tête, de par leur vélocité, peut être comparée aux saccades oculaires ; elle s'inscrit donc dans un processus de protection et préservation des espèces au cours du temps. Ainsi, plusieurs auteurs ont montré expérimentalement que ces mouvements très rapides (de l'ordre de 200°/sec au cours des tests) détérioraient la précision des sujets pour localiser un son statique en azimut car ceux-ci généraient une distorsion de l'espace auditif (Cooper 2008, Leung 2018). Dans la recherche active et précise de l'origine d'un son (e.g. quelqu'un nous interpelle dans la rue), la seconde catégorie de mouvements de tête est mise en jeu. Ces derniers sont beaucoup plus lents et nous permettent d'échantillonner la scène auditive en la balayant progressivement. Ce phénomène offre au cortex auditif la possibilité d'encoder successivement plusieurs combinaisons d'indices acoustiques pour chaque position échantillonnée, de les comparer et ainsi de cibler avec plus de précision l'origine du son (Hofman & Van Opstal 1998). Plusieurs expériences chez le normo-entendant (NE) ont ainsi montré une amélioration des performances de localisation en azimut (en réduisant les confusions avant-arrière, Wightman & Kistler 1999) et en élévation (Perrett & Noble 1997a,b) lorsque les sujets étaient libres de bouger la tête pendant l'émission du son.

Il est important de comprendre que, même si les mouvements de tête de vélocité moyenne apportent une richesse dans l'exploration de la scène auditive, ils modifient complètement la réalité de cette scène en créant une illusion de mouvement à une source sonore qui pouvait être statique. Aussi, le cerveau doit être en mesure de stabiliser l'image auditive pour que celle-ci n'apparaisse pas hachée mais bien fluide (Brimijoin & Akeroyd 2014). Il doit également être capable d'imputer cette sensation de mouvement à la tête et non pas à la source sonore elle-même qui pourrait elle aussi être naturellement en mouvement. Tête fixe, les informations

auditives sont encodées dans un système centré par la tête. Par contre, lorsque la tête est en mouvement, le référentiel tête est remplacé par le référentiel corps (Goossens & van Opstal, 1999), et plusieurs informations somatomotrices extracochléaires sont intégrées et analysées en parallèles : les informations proprioceptives et les vestibulaires (Freeman et al. 2014, Carlile & Leung 2016).

Au total, au-delà du traitement des indices acoustiques, la localisation spatiale nécessite dans plusieurs situations du quotidien l'intégration d'informations multisensorielles qui permettent de faciliter la perception du monde environnement et d'affiner l'acuité des performances de localisation (Figure I.13). La SA est donc par essence un processus multisensoriel adaptatif du milieu dans lequel est plongé l'auditeur.

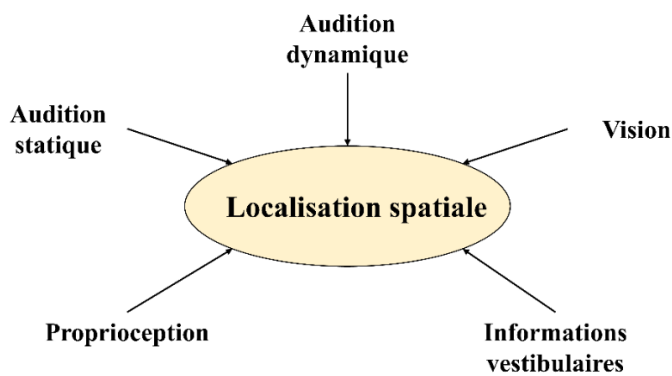


Figure I.13 : Rôles des informations multisensorielles dans la localisation spatiale

6. Lien entre spatialisation auditive et compréhension de la parole dans le bruit

Au-delà de la simple localisation d'une source sonore dans l'environnement, le système de SA joue un rôle clé dans le phénomène d'analyse de la scène auditive, concept défini par Bregman en 1990. Dans un environnement sonore complexe (e.g. cocktail party), le sujet doit être en mesure d'identifier la nature des différents sons qui l'entourent, puis il doit les localiser et les discriminer les uns par rapport aux autres (Kidd et al. 2005). Les sons compétitifs peuvent avoir un pouvoir masquant important s'ils sont de même nature que la source d'intérêt (e.g. de la parole émise par deux locuteurs ayant une fréquence fondamentale proche) ou bien ils peuvent être plus diffus et donc plus facile à isoler (e.g. ambiance instrumentale). Prenons l'exemple d'une scène au restaurant (Figure I.14) avec un bruit de fond généré par un orchestre et des conversations de groupe (effet cocktail-party). Pouvoir localiser correctement le locuteur d'intérêt offre un avantage non négligeable dans la compréhension du message délivré (Bronkhorst 2000).

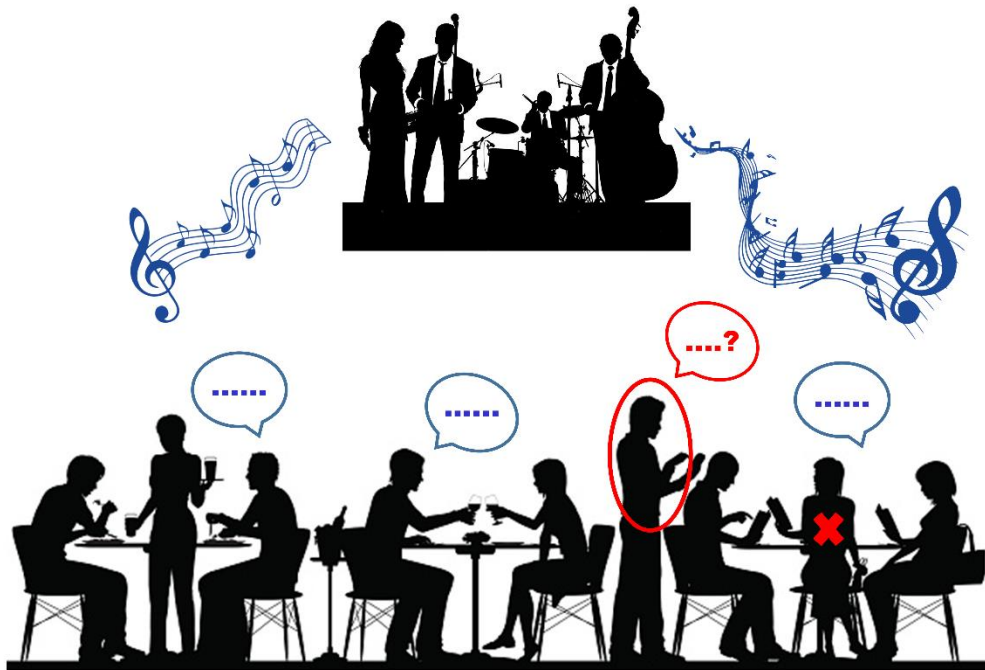


Figure I.14 : Exemple de scène auditive complexe au restaurant.
 Le locuteur (croix rouge) doit localiser la source d'intérêt (le serveur cerclé de rouge) parmi un certain nombre de sources compétitives (en bleu) autour de lui. Ces dernières sont de nature diverse (bruit de fond musical, autres locuteurs).

Toute perte auditive, même moyenne, entraîne d'importants troubles de compréhension de la parole dans le bruit, en partie du fait d'un manque de ségrégation des différentes sources sonores (Gatehouse & Noble 2004). Cette ségrégation des sources est bien connue dans la littérature anglophone sous le terme de « Spatial Release from Masking » (Edmonds & Culling 2005, Marrone et al. 2008) ; elle nécessite une extraction et une interprétation correcte des différentes ITD générées par chaque source (Middlebrooks 2015). Mais ce processus est complexe car la superposition des sons est à l'origine d'une distorsion importante des indices binauraux (Blauert 1997, Hawley et al. 1999, Langendijk et al. 2001, Best et al. 2004). Et les mécanismes "top-down" non auditifs permettant de focaliser l'attention sur la source d'intérêt ne sont pas suffisants pour lever toute ambiguïté restante (e.g. Figure I.15). Aussi, le processus de démasquage passe par l'isolation de chaque source séparément sur un temps très court lorsque chacune est émise seule (Faller & Merimaa 2004).

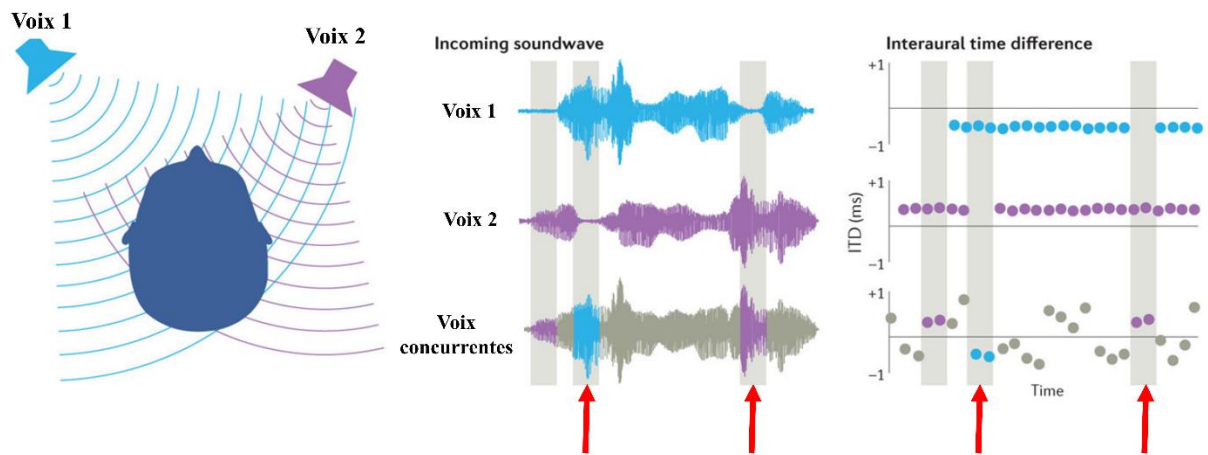


Figure I.15 : Spatialisation auditive en présence de deux voix concurrentes (d'après van der Heijden et al. 2019). Le panneau du milieu décrit 3 situations : Voix 1 isolée, Voix 2 isolée, et Voix concurrentes (les deux voix sont superposées dans le temps).

Le schéma de droite reprend ces 3 situations avec une simulation des ITD exploitables dans chaque cas. Analysées isolément, les ITD générées pour chaque voix sont facilement interprétables par le système auditif central (cercles bleus et violets). Mais lorsque les voix se superposent, les ITDs fluctuent dans le temps et subissent des phénomènes de distorsion qui les rendent ininterprétables par le système auditif (cercles gris). Dans le flux de parole, le système auditif doit arriver à capter dans de courtes fenêtres de temps (flèches rouges) les voix prises isolément afin de faciliter l'interprétation des ITD et donc la localisation des interlocuteurs.

Cette situation illustre bien les conditions d'écoute difficile du quotidien à l'origine d'une fatigabilité si la ségrégation des sources sonores doit être prolongée dans le temps.

7. Spécificités pédiatriques de la spatialisation auditive

Les différentes bases anatomo-physiologiques de la SA développées au début de ce chapitre valent pour un adulte NE. L'ensemble de ces pré-requis résulte de l'intégrité du développement auditif spatial, et celui-ci est progressif chez l'enfant. La perception que nous avons de notre environnement à l'âge adulte passe par une combinaison de changements génétiquement programmés sur le plan anatomique, physiologique mais aussi par l'expérience auditive pratiquée dès le plus jeune âge. L'objectif de ce paragraphe est de faire le point sur ces grandes étapes de maturation qui couvrent une large période de vie, de la petite enfance jusqu'à l'adolescence.

Le système auditif se développe très tôt au cours de la période fœtale, les cellules ciliées étant fonctionnelles dès le 3^e trimestre de grossesse (i.e. autour de la 27^e semaine) ce qui permet au fœtus de percevoir et de reconnaître la voix de sa mère (Galambos & Hecox, 1978). A la

naissance, l'ensemble des structures auditives sont présentes mais plusieurs phénomènes vont se succéder permettant une maturation progressive de la périphérie vers les structures centrales, comme le révèlent, entre autres, les études fonctionnelles avec les potentiels évoqués auditifs (PEA, voir Figure I.16) obtenus chez des sujets NE (e.g. Salamy 1984, Ponton et al. 2000).

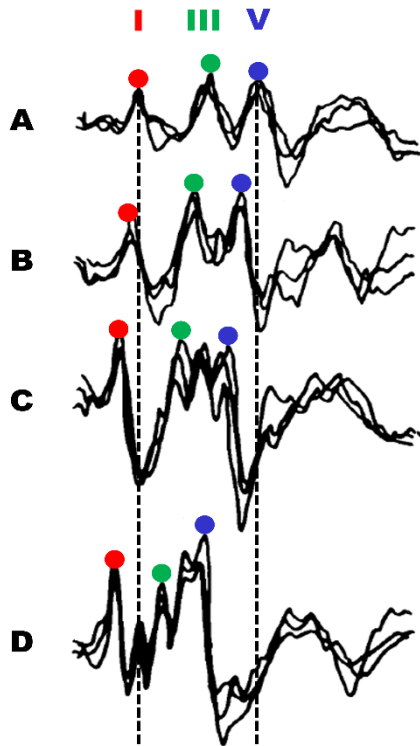


Figure I.16 : Evolution de la maturation auditive centrale sur les potentiels évoqués auditifs (PEA) précoces (adapté de Salamy 1984).

Les PEA ont été réalisés à différentes périodes : (A) en pré-natal à 35 semaine de gestation, (B) à la naissance, (C) à 3 mois de vie, (D) à 3 ans.

Les ondes I (rouge), III (vert), et V (bleu) correspondent respectivement aux réponses nerveuses provenant du nerf cochléaire, du complexe olivaire supérieur, et du colliculus inférieur.

La maturation auditive est visible à deux niveaux : augmentation de l'amplitude des ondes I, III et V avec le temps ; baisse de la latence d'apparition des ondes I, III et V avec le temps grâce aux processus de myélinisation.

Tout d'abord, l'oreille externe subit une croissance importante avec une augmentation de la taille du pavillon et du conduit auditif externe (CAE) conduisant à une modification importante et répétée dans l'enfance des HRTFs (par exemple, la fréquence de résonance maximale du CAE diminue progressivement jusqu'au début de la puberté, voir Figure I.17). Le système auditif central doit donc encoder régulièrement de nouveaux indices spectraux, interpréter plusieurs combinaisons de ces stimuli et créer de nouvelles cartes spatiales au fur et à mesure de la croissance de l'enfant.

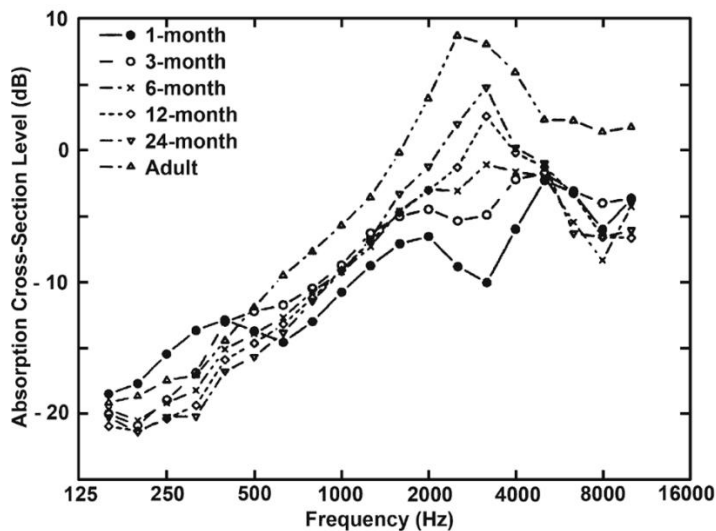


Figure I.17 : Evolution de la fréquence de résonance du conduit auditif externe en fonction de l'âge (d'après Abdala & Keefe 2011, page 22).

La fréquence de résonance maximale diminue avec l'âge et l'augmentation de la taille du CAE, passant de 5 000 Hz à 1 mois de vie à 2 500 Hz à l'âge adulte.

En parallèle de ces changements anatomiques, des connections synaptiques se mettent en place au niveau central ainsi qu'un processus de myélinisation les premières années de vie (Litovsky 2015).

A l'heure actuelle, l'ensemble du développement auditif humain n'est pas complètement élucidé mais les études comportementales associées à la neuroimagerie et aux extrapolations des résultats constatés chez l'animal permettent de définir les différentes capacités de localisation à un âge donné. Ainsi, dès l'âge de 3 mois, l'enfant est capable de s'orienter instinctivement vers une source sonore d'intérêt (Muir & Field 1979). Cette capacité est d'ailleurs facilement visualisable lors de la réalisation d'audiométrie comportementale avec mise en place d'un réflexe d'orientation conditionné. Par la suite, les compétences spatiales de l'enfant s'affinent et se diversifient. Il acquiert ainsi la capacité à discriminer spatialement deux sources mais aussi à localiser avec précision la position d'une source dans l'espace : c'est la localisation spatiale absolue.

La discrimination spatiale consiste à évaluer l'angulation minimale (the « Minimum Audible Angle » - MAA dans la littérature anglophone) pour laquelle un sujet arrive à identifier deux sources séparées. Dès l'âge de 2 ans, les enfants sont capables de discriminer deux sources antérieures espacées de 30° puis à partir de 5 ans, certains ont un seuil de discrimination proche de celui des adultes (environ 5°, pour revue voir Litovsky 2011 et Figure I.18).

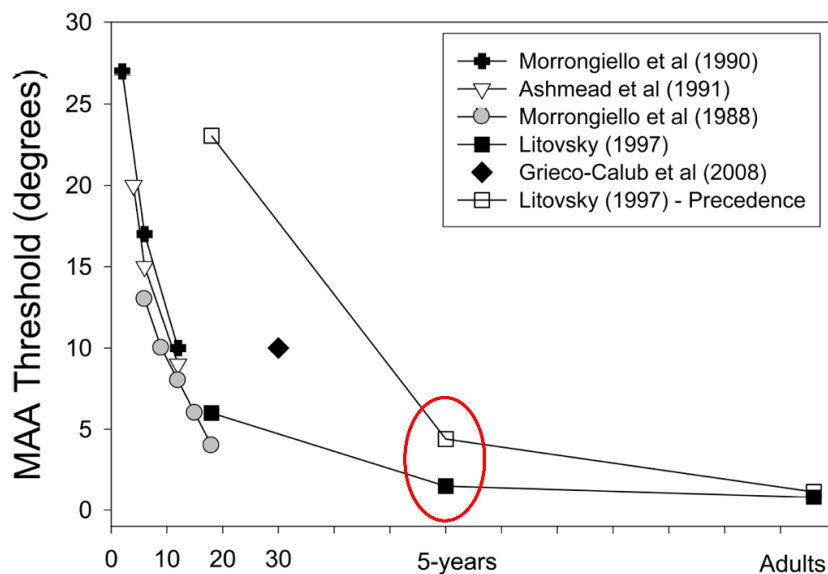


Figure 1.18 : Evolution de l'angle minimal audible (MAA) en fonction de l'âge (d'après Litovsky 2011). Les données ont été recueillies à partir de plusieurs études scientifiques. Plus le MAA est faible et plus les capacités de discrimination auditive sont bonnes. Le MAA diminue drastiquement les deux premières années de vie et atteint le seuil adulte (cercle rouge) vers l'âge de 5 ans.

La localisation spatiale absolue désigne la représentation exacte d'une source sonore, la plupart du temps statique, dans les trois plans de l'espace. Elle nécessite l'apprentissage de l'extraction des indices binauraux et monauraux ainsi que leur traitement par les voies sous-corticales et corticales. Plus le sujet est précis dans sa localisation, plus l'erreur absolue (i.e. l'angle qui sépare la source réelle et la position perçue) est faible. A partir de 5 ans, cette erreur peut être inférieure à 10° en azimuth chez certains enfants (e.g. Van Deun et al. 2019), ce qui est l'ordre de grandeur retrouvé chez l'adulte (e.g. Kerber & Seeber 2012). *Pour plus de détails, voir le paragraphe 3 du chapitre 3.*

Quelle que soit la tâche de localisation spatiale (discrimination ou évaluation de la position absolue d'une source), la plupart des auteurs s'accordent sur l'importante variabilité inter-individuelle retrouvée même après l'âge de 5 ans. Cette variabilité au cours des tests est multifactorielle. Tout d'abord, il existe certainement une différence de maturation corticale entre les enfants si bien que certains n'acquièrent les performances d'adultes qu'à l'adolescence (Litovsky 2015). De plus, au cours des tests, des facteurs confondants extra-auditifs sont susceptibles de fausser les performances réelles des enfants : niveau attentionnel, capacités de compréhension de la tâche, compétences mnésiques.

Au-delà de cette variabilité inter-individuelle, il existe également une variabilité intra-individuelle dans l'acquisition des compétences spatiales. En effet, à un âge donné, les capacités de discrimination spatiale ne sont pas forcément aussi matures que celles de la localisation absolue et inversement. Ces deux compétences ne font pas appel aux mêmes voies auditives centrales et leur maturation semble se faire indépendamment l'une de l'autre. Aussi, lors de la réalisation de l'un de ces deux tests, il n'est pas possible de prédire selon les résultats obtenus les capacités de l'autre fonction auditive (Grieco-Calub & Litovsky 2010).

Comme nous l'avons mentionné dans le paragraphe 5, de nombreuses tâches auditives complexes jalonnent notre quotidien (e.g. scène auditive avec plusieurs locuteurs) et leur analyse précise nécessite la mise en jeu de fonctions multisensorielles. Plusieurs auteurs ont montré que l'intégration multisensorielle mettait presque une décennie à être optimale chez l'enfant (Neil et al. 2006, Gori et al. 2008) mais ces premiers effets bénéfiques interviennent précocément dès les premières interactions sensorielles. Ainsi, à partir de 8 mois, l'intégration visuo-auditive permet d'augmenter la rapidité d'orientation oculaire et céphalique vers la source d'intérêt, comparé à la même expérience réalisée en condition unisensorielle (stimulus visuel ou auditif) (Morrongiello & Rocca 1987, Lewkowicz 2000,2002, Neil et al. 2006). Vision et audition sont étroitement liées depuis la naissance et certains auteurs émettent l'hypothèse que le système visuel sert à calibrer le système auditif au cours de la croissance car ce système sensoriel est, chez l'Homme, le plus précis pour cartographier l'espace environnemental (Gori et al. 2012).

8. Points-clés du chapitre

- a) Localiser une source sonore nécessite de connaître ses coordonnées spatiales en 3D : azimut – élévation – profondeur.
- b) Les indices binauraux (ITD et ILD) permettent de localiser un son en azimut et participent à la perception de la profondeur.
- c) Les indices spectraux induits par l'effet pavillonnaire permettent de localiser un son en élévation.
- d) Dans des scènes auditives complexes, la mise en jeu d'informations multisensorielles (vision, indices acoustiques dynamiques générés par les mouvements de tête) permettent d'optimiser la spatialisation auditive.
- e) La maturation de la spatialisation auditive chez l'enfant est un processus long et progressif couvrant au moins les 5 premières années de vie.
- f) Jusqu'à l'adolescence, les performances de spatialisation sont marquées par une importante variabilité inter-individuelle et intra-individuelle suivant la tâche auditive réalisée (i.e. discrimination spatiale versus localisation spatiale absolue).

Chapitre 2

Spatialisation auditive et implant cochléaire

1. Généralités sur les surdités

1.1. Classifications

La surdité est un problème de santé publique affectant tout autant la population adulte que la population pédiatrique. Les profils de surdités sont très hétérogènes d'un patient à l'autre, il est donc important de bien les classer selon plusieurs critères afin d'évaluer au mieux leurs conséquences. Ainsi, les surdités peuvent être classées selon différents facteurs :

- Leur étiologie : congénitale ou acquise
- Leur moment de survenue : dans l'enfance ou à l'âge adulte ; on définit alors 3 périodes dans l'enfance selon l'âge d'apparition de la surdité par rapport au développement du langage :
 - o Période pré-linguale, avant l'âge de 2 ans
 - o Période péri-linguale, entre 2 et 4 ans
 - o Période post-linguale, après l'âge de 4 ans
- Leur degré de perte auditive :
 - o Surdité légère : perte de 20 à 40 dB,
 - o Surdité moyenne de type 1 : perte de 40 à 55 dB,
 - o Surdité moyenne de type 2 : perte de 55 à 70 dB,
 - o Surdité sévère : perte de 70 à 90 dB,
 - o Surdité profonde : perte de 90 à 120 dB,
 - o Cophose : perte au-delà de 120 dB.
- Leur évolutivité : stable, fluctuante, dégradation progressive, voire amélioration spontanée

Les conséquences de la surdité seront d'autant plus sévères que :

- L'âge de survenue est précoce (avant 4 ans) : avant le développement du langage et la maturation des voies auditives,
- La surdité est sévère à profonde,
- La surdité s'accompagne de signes associés, notamment de troubles neurologiques ou vestibulaires,
- La réhabilitation auditive est tardive.

Dans le cadre des surdités sévères à profondes bilatérales, les prothèses auditives conventionnelles s'avèrent inefficaces pour restaurer convenablement la compréhension de la

parole dans le silence et dans le bruit. Ainsi, depuis les années 70, l'utilisation de l'implant cochléaire s'est répandue en France et dans le monde afin de restituer au mieux la communication chez les patients sourds sévères à profonds.

A l'heure actuelle, chez l'adulte, la bi-implantation est discutée au cas par cas, en tenant compte notamment du souhait du patient (Hermann et al. 2019). Dans la population pédiatrique, les sociétés savantes recommandent la bi-implantation cochléaire, avec un délai le plus court possible, afin de restaurer rapidement la binauralité et d'éviter les conséquences néfastes d'une privation auditive (SFORL 2018, Simon et al. 2019).

1.2. Période critique et privation auditive chez l'enfant

Alors que la cochlée humaine a achevé son développement à la naissance, les voies et centres auditifs vont se développer plus progressivement, depuis le tronc cérébral jusqu'au cortex. La période critique représente une période théorique au cours des 9 premiers mois de vie où le développement des structures auditives et les pré-requis du langage se mettent en place. De façon notable, des changements majeurs surviennent au niveau cortical avec une spécialisation de certaines aires polysensorielles en aires auditives sous l'influence de stimuli auditifs (e.g. Pascual-Leone et al. 2005). De plus, au cours de cette période, le développement de la spatialisation auditive (SA) est aussi influencé par la modalité visuelle, permettant aux cartes spatiales visuelles et auditives d'être cohérentes spatialement (pour revue, voir Opoku-Baah et al, 2021).

Toute surdité survenant en cours de cette période critique (i.e. surdité pré-linguale) risque de compromettre la maturation auditive de la périphérie jusqu'aux structures centrales (Kral et al. 2005, Trumpp & Kiefer 2018) et d'impacter la SA. Tout d'abord, quel que soit l'âge, on constate une mort neuronale progressive des cellules du ganglion spiral en l'absence de stimulations auditives prolongées (Kan & Litovsky 2015a, Todd et al. 2017). De plus, en l'absence de stimulations auditives précoces, les structures auditives cérébrales se réorganisent rapidement, avec à minima une perte de l'organisation tonotopique dans le cortex auditif primaire (e.g. Kral et al. 2009), voire une modification complète de l'activité de cette aire cérébrale au profit du traitement d'autres informations sensorielles, notamment visuelles (Crair & Malenka 1995, Barkat et al. 2011, Chun et al. 2013).

Ainsi, la privation auditive doit être raccourcie le plus possible afin de profiter des facilités de la plasticité cérébrale qui est optimale entre 0 et 2 ans mais finit par décroître après l'âge de 7 ans (Sharma et al. 2002). Un appareillage précoce offre toutes les chances de développer correctement le langage mais aussi les compétences de localisation spatiale (Figure II.1, Kral et al. 2019).

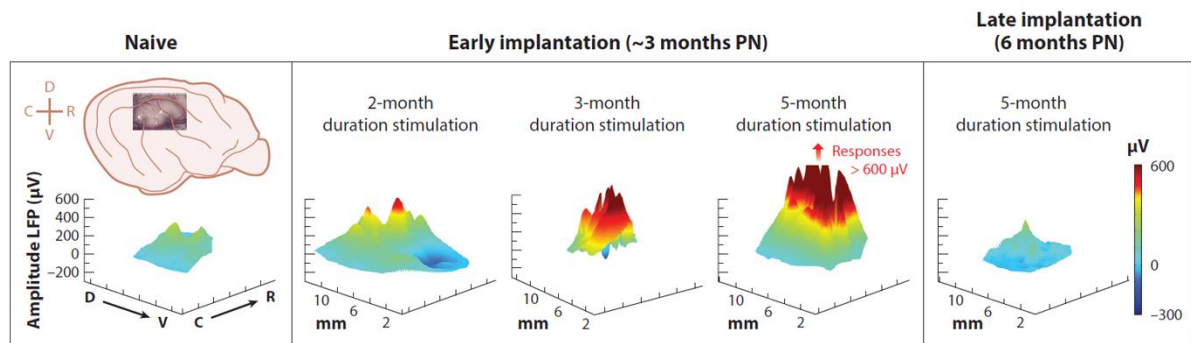


Figure II.1 : Effets de l'implantation cochléaire précoce et tardive sur la maturation du cortex auditif primaire du chat (d'après Kral et al. 2019).

Le panneau de gauche illustre l'activité corticale d'un chat sourd non appareillé, le panneau central après implantation précoce, et le panneau de droite après implantation tardive. Seule l'implantation précoce conduit à une réorganisation du cortex auditif avec augmentation et extension de l'activité neuronale au cours du temps.

2. Généralités sur l'implant cochléaire

2.1. Différentes parties de l'implant

L'implant cochléaire (IC) est une neuroprothèse implantable qui capte les sons acoustiques de l'environnement et les transmet, après transformations, sous forme d'impulsions électriques au nerf auditif. Il comprend deux parties distinctes, externe et interne (Figure II.2).

- La partie externe est composée d'un processeur relié à l'antenne par un câble. Les sons de l'environnement sont captés par le microphone situé dans la majorité des cas sur le processeur. Dans la population pédiatrique, il est toutefois possible d'activer le microphone sur l'antenne afin de positionner le processeur en déporté sur un vêtement (cela s'avère utile en cas d'absence de pavillon dans les aplasies d'oreille par exemple).
- La partie interne comprend un récepteur qui reçoit et numérise le son puis le transfère au porte-électrodes inséré chirurgicalement dans la cochlée. Suivant les marques d'implant, le porte-électrodes est composé de 12 à 22 électrodes réparties tout au long

de la cochlée afin de respecter l'axe tonotopique. Chaque électrode est en contact avec plusieurs fibres du nerf auditif qui sont sélectivement activées en fonction de la fréquence du signal d'entrée.

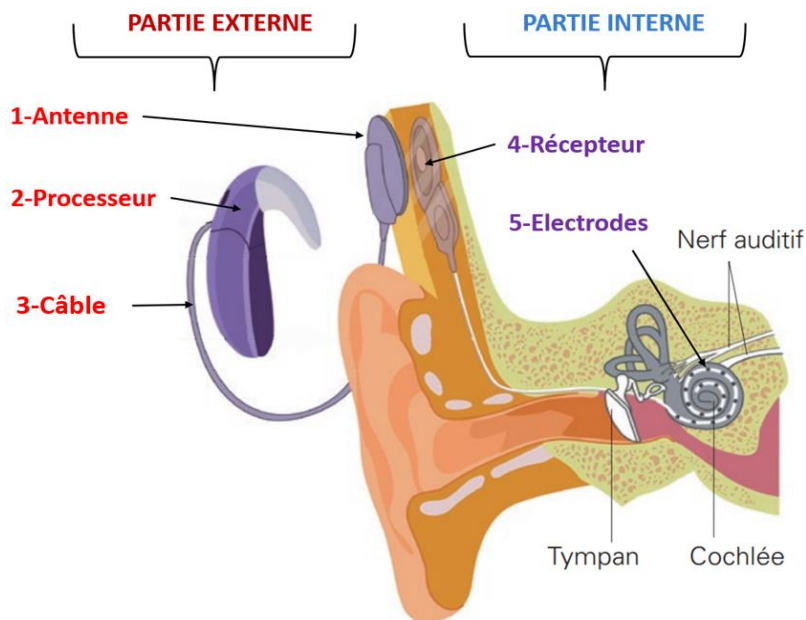


Figure II.2 : Schématisation de l'implant cochléaire.

L'implant cochléaire est composé de deux parties distinctes, la partie interne est insérée chirurgicalement tandis que la partie externe est facilement amovible par le patient, telle une prothèse conventionnelle.

2.2. Principes de codage du son par l'implant

Afin de s'approcher au plus près de la physiologie d'une oreille normo-entendante, l'implant cochléaire doit respecter trois étapes clés dans le traitement du signal acoustique entrant : analyser la nature du son (en fréquence et en intensité), le répartir le long de la cochlée en respectant l'axe tonotopique et le transformer en signal électrique exploitable par le nerf auditif (Figure II.3). Ainsi, le signal acoustique capté par le microphone est transmis au récepteur par radiofréquence. Celui-ci analyse le signal reçu et le démodule, c'est-à-dire qu'il découpe le signal en bandes de fréquences (12 à 22 suivant le nombre d'électrodes réparties le long de la cochlée) et l'énergie correspondante est distribuée sur chaque électrode sous forme de séquences d'impulsions. Les bandes de fréquences extrêmes (< 60 Hz et $> 6\,000$ Hz) sont filtrées car jugées peu informatives pour les activités du quotidien. En parallèle, un système de compression est appliqué sur chaque bande de fréquences (principalement pour les sons

ambiants de l'environnement et les sons brusques de très courte durée) afin que la dynamique acoustique soit adaptée à la dynamique électrique du nerf auditif. La dynamique électrique de chaque électrode est facilement modifiable au cours des réglages audioprothétiques pour assurer le confort auditif du patient. Elle est comprise entre le seuil auditif de confort (C) et le seuil de détection (T) du patient.

Suivant les marques d'IC, les stratégies de codage varient, de la fréquence de stimulation (également appelée « pulse ») aux nombres d'électrodes stimulées. Mais un principe général demeure, le codage de l'intensité est transmis sous forme de pulses au nerf auditif, tandis que le codage fréquentiel est assuré par la répartition spatiale des électrodes le long de la cochlée : les fréquences graves stimulent les électrodes apicales et les sons aigus activent les électrodes basales. Afin de conserver au mieux cette tonotopie cochléaire, les électrodes doivent être activées localement et séquentiellement, dans le cas contraire, des interactions inter-canaux peuvent conduire à une perte de la sélectivité fréquentielle.

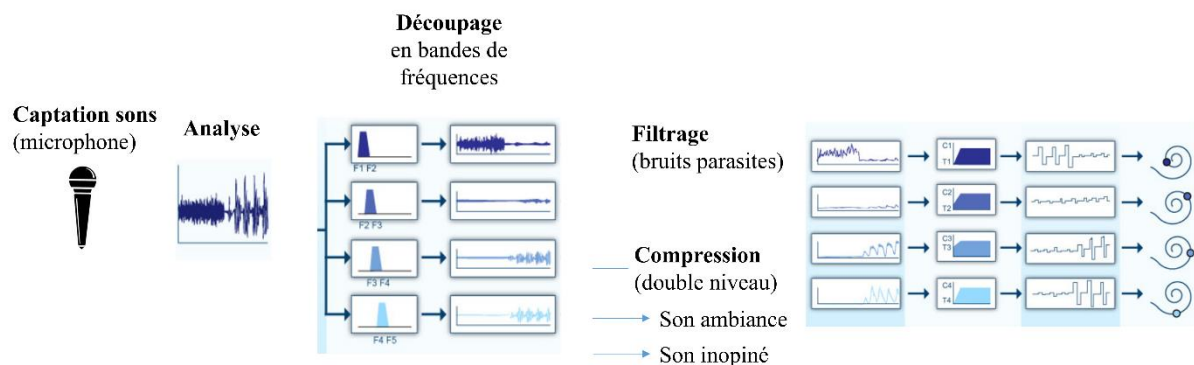


Figure II.3 : Différentes étapes de codage du son par l'implant cochléaire

3. Limites technologiques de l'implant

3.1. Dégradation des informations binaurales

Nous avons montré dans le chapitre 1 que le traitement et l'intégration des indices binauraux et monauraux étaient un point fondamental pour l'évaluation correcte de la position d'une source sonore en 3D. Or, chez les patients implantés cochléaires, plusieurs freins technologiques viennent impacter leurs capacités de localisation. Dans la suite de cette partie, nous détaillerons les principales limites de l'implantation cochléaire en termes de SA.

3.1.1. Codage très limité de l'ITD

Après traitement par le processeur, la structure-même du son est fortement modifiée, comme l'indique la Figure II.4. Ainsi, l'ITD est l'indice binaural le plus dégradé car la structure temporelle fine n'est plus codée mais est remplacée par plusieurs trains de pulses. Depuis quelques années, certaines stratégies de codage ont été développées par les industriels pour essayer de conserver une partie de la structure temporelle fine du son. Citons pour exemple MED-EL qui a implémenté une stratégie de stimulation pulsée (appelée FSP) visant à préserver au mieux la structure temporelle fine du son dans les basses fréquences. Plusieurs études suggèrent que cette stratégie améliore les capacités de compréhension de la parole dans le bruit (Lorens et al. 2010, Vermeire et al. 2010, Punte et al. 2014). Les bénéfices attendus en termes de localisation sont plus discutés car il semblerait que cette nouvelle stratégie de codage ne soit pas encore suffisante pour préserver de façon optimale l'ITD (Dorman et al. 2014, Ausili et al. 2020). Il est difficile d'expliquer cette absence d'effet sur la localisation car une étude récente a montré que les variations d'ITD perceptibles par les patients implantés utilisant le FSP étaient quasiment physiologiques pour la moitié d'entre eux (autour de 330 μ s, Eklöf & Tideholm 2018). Dans cette étude, la majorité de la population testée était des enfants implantés bilatéralement avant l'âge de 3 ans, donc à une période optimale pour le développement du système auditif. Dans l'étude de Ausili et collaborateurs (2020) réalisée chez des patients implantés à l'âge adulte, la sensibilité des patients aux ITD était beaucoup plus variable, d'une résolution médiocre à une perception d'ITD inférieure à la milliseconde. Cette importante variabilité entre les participants peut s'expliquer par les différents profils cliniques testés avec d'importants délais entre les deux implants et une majorité des patients implantés tardivement.

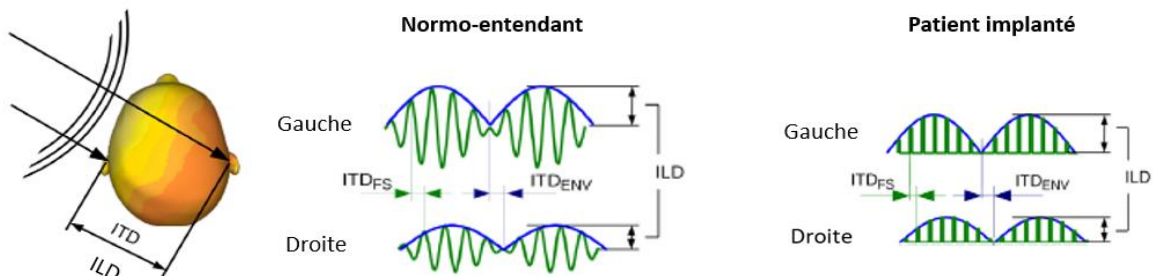


Figure II.4 : Comparaison des indices binauraux traités par un sujet normo-entendant et un patient bi-implanté cochléaire (adapté de Laback 2015)

L'ILD est l'indice binaural le mieux conservé par l'implant. À l'opposé, l'ITD est fortement dégradée avec perte de la structure fine.

Comme le souligne Ausili (2020), le manque de perception des ITD ne se résume pas uniquement à la perte de résolution du signal entrant filtré par l'implant. En effet, la perception des ITD est aussi soumise aux bonnes capacités d'intégration centrale du sujet. Celles-ci nécessitent une synchronisation parfaite des signaux captés par chaque oreille et analysés au niveau de l'olive supérieure médiane pour détecter une concordance temporelle de ces signaux. A l'heure actuelle, un patient porteur de deux IC ne bénéficie pas de cette synchronisation périphérique, ses deux implants traitant de manière indépendante les sons avec leurs propres caractéristiques de réglage. Cela renforce la distorsion des ITD perçues au niveau central. Enfin, une surdité congénitale non réhabilitée précocement présente le risque d'une détérioration fonctionnelle des voies corticales de spatialisation ; en effet, leur bon fonctionnement nécessite une expérience binaurale prolongée les premières années de vie (Grothe et al. 2010).

3.1.2. Compression de l'intensité des sons

Comme nous l'avons vu en début de chapitre, la dynamique électrique de l'IC est bien plus limitée que la dynamique acoustique de l'oreille humaine. Aussi, les IC actuels sont équipées à l'entrée d'un système de compression de l'intensité sonore appelé « Automatic Gain Control » (AGC). Cet AGC a pour rôle de compresser une large gamme d'intensité de notre environnement (allant de 0 à 120 dB HL) vers une gamme beaucoup plus étroite (autour de 10 à 20 dB HL) exploitable par la partie interne de l'implant. Ainsi, les sons faibles sont amplifiés et les sons forts sont atténués afin de préserver le système auditif d'un éventuel traumatisme sonore (Dillon 2001, Archer-Boyd & Carlyon 2019). La majorité des sons de l'environnement se retrouvent ainsi compressés autour d'un niveau confortable de 65 dB HL (Stöbich et al. 1999, Khing et al. 2013). Ce phénomène de compression induit deux conséquences majeures. Premièrement, l'intensité des sons de l'environnement étant compressée, il devient difficile d'utiliser les variations d'intensité pour évaluer avec précision la distance d'une source sonore. Deuxièmement, pour qu'une ILD soit perceptible pour un patient implanté, il faut que celle-ci soit déjà suffisamment grande avant d'être compressée par l'IC. Cela explique donc les difficultés ressenties par les patients pour déterminer avec précision la localisation d'un son proche de la ligne médiane (i.e. à l'endroit où l'ILD est la plus faible).

Il faut également noter que ces phénomènes de compression ne s'appliquent pas de manière uniforme aux deux processeurs car ceux-ci ne sont pas synchronisés. Ainsi, en fonction de la position de la source sonore et de son intensité, la compression peut s'appliquer à l'IC le plus

proche du son mais pas à l'IC controlatéral (parce que le niveau sonore perçu par celui-ci devient inférieur au seuil de compression avec l'effet d'ombre de la tête). Dans les conditions les plus délétères, ces phénomènes de compression asymétriques peuvent conduire à des ILD inversées et donc fausser complètement la localisation du sujet (Dorman et al. 2014, Archer-Boyd & Carlyon 2019).

3.1.3. Absence de fusion des indices binauraux

Physiologiquement, un son de notre environnement vient stimuler une zone fréquentielle identique au niveau de chaque cochlée puis chaque nerf auditif décharge de manière synchrone en véhiculant les informations binaurales de chaque oreille. Ainsi, les noyaux cochléaires du tronc cérébral reçoivent des informations ipsi et controlatérales qu'ils peuvent comparer et fusionner. En présence de deux IC, cette fusion des informations binaurales est compromise à plusieurs niveaux (Kan and Litovsky 2015). Tout d'abord, les deux processeurs ne sont pas synchronisés, chacun analyse les sons de l'environnement pour son propre compte donnant naissance à un signal singulier pour chaque IC. Cela signifie que *in fine*, chaque nerf auditif décharge dans la majorité des cas de façon asynchrone par rapport au côté controlatéral (Steel et al. 2015).

De plus, lors de l'implantation cochléaire, il n'existe pas à l'heure actuelle d'outils radiologiques et électrophysiologiques suffisamment précis pour réaliser une introduction à la même profondeur du porte-électrode dans chaque cochlée. Ainsi, lorsqu'un son de l'environnement est émis, même si des électrodes identiques sont activées sur chaque partie interne, celles-ci stimuleront des zones fréquentielles légèrement différentes entre les deux oreilles (Kan et al. 2013, Fitzgerald et al. 2015, Steel et al. 2015, Kan et al. 2019), et les structures centrales recevront des informations non congruentes de chaque oreille. Cette différence d'insertion des porte-électrodes impacte significativement l'intégration centrale de l'ITD mais assez peu l'ILD (Kan & Litovsky 2015). Rappelons qu'en pratique clinique, les patients bi-implantés sont parfois porteurs d'une partie interne différente au niveau des deux cochlées, du fait du délai inter-implant plus ou moins important et du bond technologique qui s'opère entre chaque génération d'implant. Cette différence technologique participe aux difficultés de fusion des indices binauraux. Certaines situations s'avèrent bien plus délétères lorsque les patients font le choix de changer de marque pour le deuxième implant. Cela renforce encore plus l'asymétrie des informations auditives qui parviennent aux structures centrales car

les implants des différentes marques n'ont pas forcément le même nombre d'électrodes ni la même philosophie de codage.

3.2. Dégradation des informations monaurales

Comme nous l'avons vu dans le chapitre 1, les indices monauraux sont en grande partie générés par l'effet pavillonnaire. Or, avec l'IC, le pavillon de l'oreille perd son rôle fondamental de filtres conduisant à un appauvrissement des indices spectraux. En effet, le son est directement capté par le microphone du processeur, et quelle que soit sa position (i.e. sur le processeur derrière le pavillon, ou en déporté au-dessus du pavillon pour les modèles Rondo ou Kanso, voir Figure II.5) la majorité des algorithmes ne permettent pas de restaurer les nuances spectrales induites physiologiquement. Mais certains auteurs suggèrent que suivant le positionnement du microphone, les capacités de SA seraient améliorées ou dégradées (Jones et al. 2016, Denk et al. 2019).

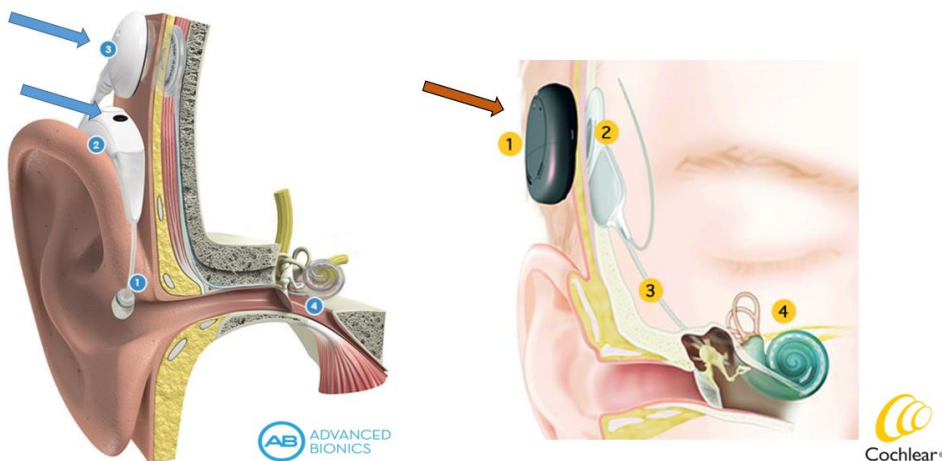


Figure II.5 : Exemples de positionnement des microphones sur le processeur de l'implant. Selon les marques d'implant, le microphone peut être activé à l'arrière du processeur (e.g. Advanced Bionics à gauche), ou sur l'antenne (e.g. Cochlear à droite).

4. Spécificités de l'implant en pédiatrie : quand immaturité et asymétrie auditives s'emmêlent

Chez l'enfant, la bi-implantation cochléaire fait office de référence depuis plusieurs années car elle a largement montré ses bénéfices pour le développement du langage, la localisation spatiale et l'insertion scolaire (SFORL 2018, Simon et al. 2019). Tout comme les

adultes, les enfants sont confrontés aux limites technologiques de l'IC. Mais d'autres facteurs peuvent directement impacter le neurodéveloppement de l'enfant : l'âge de survenue de la surdité, l'âge lors de la première implantation cochléaire, et le délai entre les 2 IC (appelé délai inter-implant). Une étude récente a montré que malgré la restauration précoce de l'audition binaurale, les enfants bi-implantés ne parvenaient pas à atteindre les mêmes performances que des enfants NE (Gordon et al. 2017). La surdité entraîne une réorganisation neuronale que l'implantation cochléaire, même précoce, ne parvient pas à inverser. En effet, l'absence de synchronisation entre les IC et leur décalage de positionnement dans chaque cochlée conduisent à une inadéquation dans les ILD et ITD générées par absence de fusion binaurale centrale (Kan et al. 2015, Steel et al. 2015). L'altération des indices binauraux touche largement les ITDs, et dans une moindre mesure les ILDs. Ainsi, après implantation, la perception de l'ILD s'établit rapidement après quelques mois d'expérience binaurale et ce, malgré une privation auditive longue (Gordon et al. 2014). Le traitement de l'ILD au niveau sous-cortical est un système rudimentaire hautement conservé au cours de l'évolution (Grothe et al. 2010) et son analyse implique des mécanismes neuronaux beaucoup plus simples que celui de l'ITD. L'expérience auditive par exposition précoce aux sons de l'environnement permet aux différentes zones sous-corticales et corticales de suivre une maturation proche de celle d'un enfant NE (Gordon et al. 2003).

A l'opposé, la réhabilitation précoce d'une surdité congénitale ne permet pas de rétablir un traitement adéquat de l'ITD au niveau du tronc cérébral (Tillein et al. 2010). En effet, les mécanismes physiologiques qui sous-tendent l'analyse de l'ITD sont complexes et plus évolués, ils nécessitent une expérience binaurale longue (Grothe et al. 2000). Plusieurs études ont ainsi mis en évidence les capacités de détection très limitées des ITD, dans des cas où la binauralité était restaurée rapidement (Gordon et al. 2014) et les enfants avaient une expérience binaurale supérieure à 2 ans (Salloum et al. 2010). Toutefois, les différences temporelles restaient exploitables par ces enfants si elles étaient 10 fois plus importantes que le seuil de détection d'un enfant NE (respectivement 150 μ s versus 10 μ s, Gordon et al. 2014). Finalement, l'accès limité aux ITDs chez l'enfant est double ; il est lié aux limites technologiques de l'implant et à l'altération du traitement auditif central dans la petite enfance.

Aussi, afin de minimiser les difficultés de traitement des indices binauraux, la pose du premier IC doit avoir lieu avant l'âge de 18 mois, voire idéalement avant 12 mois de vie (McConkey Robbins et al. 2004, Svirsky et al. 2004, Niparko et al. 2010, Nicholas & Geers 2013).

Le délai inter-implant est le troisième facteur clé dans la phase de développement et de maturation auditive du jeune enfant. Si le délai inter-implant dépasse 18 mois, les enfants implantés séquentiellement sont susceptibles de développer une asymétrie auditive centrale (Gordon et al. 2010, Gordon et al. 2013a, Litovsky & Gordon 2016). En effet, durant toute la durée de port d'un seul IC, le traitement des informations binaurales reste perturbé (Gordon et al. 2015). Celui-ci peut toutefois être optimisé par le port d'une prothèse auditive controlatérale (Mok et al. 2010, Polonenko et al. 2015). Cette asymétrie conduit notamment à une baisse de l'acuité spatiale et les enfants manifestent une réelle préférence envers le premier IC (Gordon et al. 2013b, Kral et al. 2013). Pour cette raison, l'implantation cochléaire séquentielle de moins de 6 mois est actuellement le gold standard souhaité en France (Truy et al. 2017).

5. La spatialisation auditive est-elle au cœur du réglage audioprothétique des implants ?

Dans un souci d'optimisation de la compréhension de la parole dans le bruit, de nombreuses innovations technologiques ont émergé ces dernières années dans les processeurs d'IC, tels que les réducteurs de bruit ambiant et les microphones directionnels. Ainsi, l'IC est en capacité de scanner l'environnement sonore du sujet et, lorsque de la parole est détectée, l'IC diminue drastiquement le bruit de fond présent dans l'environnement et oriente ses microphones vers le locuteur en train de parler. Ces systèmes automatisés sont très utilisés pour les conversations en milieu bruyant mais à contrario, ils diminuent drastiquement la richesse de la scène auditive en supprimant de précieuses informations pour la localisation des différentes sources sonores. Dans le centre d'implantation cochléaire de Lyon, ce type de système adaptatif est proposé aux adultes et aux adolescents, mais pas aux plus jeunes enfants afin de ne pas appauvrir dès l'enfance leur environnement auditif. Aussi, lorsque ces jeunes enfants se voient proposer à l'adolescence le passage à un système plus automatisé, il est fréquent de constater un rejet de cette technologie car ces enfants ont la sensation « de ne plus tout percevoir ». Cette constatation soulève une question importante : faut-il privilégier un confort d'écoute et favoriser la communication avec ses pairs, ou bien essayer de se rapprocher le plus possible d'un environnement sonore naturel et permettre au sujet de localiser l'ensemble des sources l'environnant, même les bruits de fond ? La réponse dépend très certainement de chaque patient et de ses activités quotidiennes. Nous pouvons cependant suggérer que dans une population pédiatrique implantée jeune, il semble dommageable d'appauvrir l'environnement auditif alors que la maturation des structures centrales n'est pas encore achevée et que les structures auditives conservent une importante plasticité.

6. Points clés du chapitre

- a) L'implant cochléaire dégrade significativement les indices binauraux et monauraux indispensables à une bonne localisation des sons en 3D.
- b) L'ILD est l'indice le mieux préservé chez les implantés.
- c) La sensibilité aux indices binauraux est sujette à une importante variabilité inter-individuelle.
- d) La variabilité inter-individuelle est plurifactorielle et fonction de la technologie de l'implant, des conditions chirurgicales d'insertion, de la durée de privation auditive, et de la maturation des voies auditives centrales.
- e) Une implantation cochléaire séquentielle avec un délai inter-implant supérieur à 6 mois risque d'entraîner une asymétrie auditive irréversible.
- f) Malgré la présence de nombreux facteurs limitant l'extraction et l'interprétation des indices acoustiques, les patients implantés conservent la capacité de modifier leurs cartes de spatialisation auditive.

Chapitre 3

Outils d'évaluation et performances de spatialisation auditive

Plusieurs outils d'évaluation des performances de spatialisation auditive (SA) ont été mis au point ces dernières décennies avec différents objectifs fondamentaux et cliniques : (1) mieux comprendre le développement de la SA pendant l'enfance, (2) évaluer les difficultés de localisation spatiale de patients déficients auditifs en comparaison de sujets normo-entendant (NE).

Dans ce chapitre, nous développerons les principaux outils comportementaux utilisés dans la recherche et la pratique clinique actuelle et nous nous attacherons à caractériser pour chaque outil les performances de SA de sujets NE et implantés cochléaires. Volontairement, nous ne développerons pas les différents outils d'imagerie fonctionnelle utilisés dans la spatialisation (e.g. IRM fonctionnelle, Tomographie par Emission de Positons, Electroencéphalographie, Potentiels Evoqués Corticaux, Near Infra Red Spectroscopy).

1. Outils d'évaluation quantitative

Les tests objectifs mesurant les performances de SA se sont beaucoup développés ces dernières décennies. Les dispositifs utilisés présentent de nombreuses variantes et permettent de caractériser des capacités de SA différentes comme : 1) localiser une source sonore fixe, 2) localiser une source sonore en mouvement, 3) ségréguer plusieurs sources sonores. La connaissance de ces outils est un pré-requis indispensable pour interpréter les performances individuelles des sujets. En effet, la précision de la localisation dans chaque dimension de l'espace est conditionnée par plusieurs facteurs dont la position des haut-parleurs (HP) et le type de son délivré. Ce paragraphe a pour but de résumer les principaux outils utilisés actuellement, leurs avantages et leurs limites.

1.1. Localisation spatiale absolue

La localisation spatiale absolue caractérise la capacité à définir dans les trois plans de l'espace la position exacte d'une source sonore fixe par rapport à un référentiel. Le référentiel utilisé pour localiser cette source est dans la majorité des cas le sujet réalisant l'expérience. Celui-ci a pour instruction de pointer la direction de la source sonore perçue, soit directement à l'aide d'un laser ou en orientant la tête, soit de manière indirecte en lisant le numéro correspondant au HP ou en sélectionnant un HP matérialisé sur tablette. La méthode historique pour évaluer les performances des sujets est de calculer le « root mean square » (RMS) error, qui correspond à la déviation standard de la réponse du sujet par rapport à la source

(Middlebrooks & Green, 1991). L'erreur RMS se calcule souvent en azimut et s'exprime en degrés.

Les conditions des tests de localisation varient en fonction de différents paramètres : la position des HP délivrant le son, la nature du stimulus auditif (fréquence, durée), la méthode d'exploration de la scène auditive (mouvements de tête autorisés ou non) et la méthode de réponse du sujet (e.g. pointeur).

Dans la littérature, les expériences pionnières de SA ont été réalisées en champ-libre, avec une rangée de HP disposés en arc de cercle autour du sujet, souvent dans l'espace avant (Brungart 1999a,b). Dans la plupart des cas, ces expériences étaient faites dans une chambre anéchoïque ou peu échoïque, permettant de diminuer de façon drastique la réverbération des sons. La tête était immobilisée pour assurer la reproductibilité des essais. Une alternative a été développée afin de s'affranchir des contraintes précédentes : la localisation de sons virtuellement spatialisés. Préalablement à l'expérience, les HRTF du sujet sont mesurées pour différentes positions de HP en azimut, élévation, voire profondeur. Puis au moment du test, le sujet est équipé directement d'écouteurs délivrant des signaux respectant les indices acoustiques mesurés au départ (e.g. Majdak et al. 2010). Cette technique est donc très intéressante pour simuler un son dans n'importe quelle position de l'espace mais elle reste majoritairement réservée à la recherche fondamentale pour plusieurs raisons : (1) les sons synthétiques sont générés à partir de portions définies de l'espace ce qui limite l'échantillonnage de l'espace auditif, (2) si le sujet bouge la tête pendant l'émission sonore, les sons doivent être mis à jour en temps réel, ce qui est un écueil technique, (3) si une ambiance sonore ou visuelle est implémentée, alors les sons doivent être réévalués pour tenir compte des nouveaux indices sensoriels de l'environnement, et enfin (4) cette technique reste assez éloignée de la réalité des ambiances sonores complexes éprouvées quotidiennement par les patients déficients auditifs.

En pratique, les services cliniques sont équipés de systèmes assez encombrants composés de plusieurs HP répartis tous les 10 à 15° devant le sujet, à une distance et une hauteur fixées (e.g. Figure III.1). Certains systèmes plus complets permettent de tester la discrimination avant-arrière avec une rangée de HP disposés derrière le sujet (e.g. Fischer et al. 2020a). Ces systèmes présentent plusieurs limites d'utilisation :

- Dans la majorité des cas, les HP sont visibles lorsque le sujet entre dans la pièce d'expérimentation voire pendant toute la durée du test. Cela donne des indices précieux

sur la position des HP (voir chapitre 1 paragraphe 5) ce qui peut conduire à sous-estimer les difficultés réelles du participant.

- Les HP étant fixés à distance du sujet, celui-ci ne peut que montrer indirectement la position du son perçu (i.e. sa direction en 2D), soit à l'aide d'un pointeur laser, soit par l'intermédiaire d'une tablette qui matérialise les différentes positions possibles des HP. Cette configuration offre une aide supplémentaire pour démasquer la position du HP. De plus, l'absence de réponse motrice directe à l'émission du HP modifie la spatialisation. En effet, une étude récente (Valzolgher et al. 2020a) a montré que l'interaction directe avec l'origine du son permettait d'augmenter l'acuité auditive en se rapprochant plus fidèlement des situations du quotidien (e.g. localiser son téléphone puis l'attraper ; localiser une trottinette en approche et stopper nette sa progression sur la chaussée).
- L'absence de mobilité des HP ne permet pas d'échantillonner facilement les 3 dimensions de l'espace et dans la plupart des expériences, seule l'azimut varie. Il existe donc peu de données cliniques sur les capacités de localisation en élévation et en profondeur.



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

Le type de son utilisé au cours des tests de localisation est un deuxième point critique à analyser lorsque l'on compare les performances de SA entre les études. En effet, l'ambiance sonore peut varier d'un modèle simple (e.g. émission d'un bruit blanc par un seul HP) à une immersion auditive complexe mais plus réaliste (e.g. un HP frontal émet de la parole tandis que plusieurs HP arrières envoient un bruit de cocktail party). Tout dépend du degré de difficulté

que l'on souhaite imposer au participant. Par exemple, pour des patients implantés, l'utilisation d'un bruit blanc ou d'un bruit rose assure la perception de la stimulation auditive par tous les sujets, et ce quels que soient leurs réglages, car toutes les fréquences sont représentées.

1.2. Spatialisation auditive dynamique

Plusieurs études ont utilisé l'orientation de la tête comme outil de pointage final dans les tests de localisation (Perrott et al. 1987, Wightman & Kistler 1989, Makous & Middlebrooks 1990, Recanzone et al. 1998) mais peu d'études ont exploré les comportements d'orientation de la tête depuis l'émission du son jusqu'au pointage final. Pourtant, l'exploration de la scène auditive par des mouvements de tête est un comportement naturel d'orientation qui permet d'enrichir les indices binauraux et spectraux générés (Wallach 1940). Ce phénomène est qualifié de localisation spatiale dynamique.

Les tests de localisation dynamique nécessitent la mise en place d'un certain nombre de contrôles sur le placement de la tête au départ de chaque essai mais aussi l'enregistrement des mouvements réalisés lors de la phase de recherche du son (i.e. temps de latence, durée, orientation du premier et dernier mouvement de tête).

Différents systèmes ont été développés pour répondre à cette demande, de la simple observation des mouvements de tête au cours de l'expérience (e.g. Honda et al. 2013) à leur monitoring précis à l'aide de système de motion tracking (Goossens & Opstal 1999, Populin 2008, Van Barneveld et al. 2011). Le système VICON (Figure III.2) est un exemple de motion tracking précis, utilisé dans plusieurs études récentes (Brimijoin et al. 2010, Brimijoin & Akeroyd 2016).

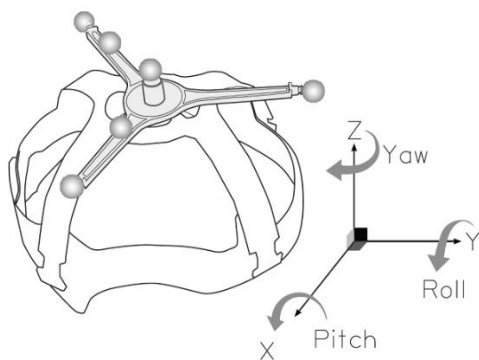


Figure III.2 : Système de localisation VICON (utilisé dans Brimijoin et al. 2010)

Lors du test, les sujets s'équipent du casque surmonté de boules réfléchissantes qui sont captées par des caméras infra-rouges. Ce système permet d'enregistrer chaque mouvement de tête et son orientation exacte selon les 3 plans de l'espace (X, Y et Z)

2. Développement d'un nouvel outil de spatialisation auditive : SPHERE

2.1. Atouts du système

Pour pallier les différents écueils méthodologiques cités ci-dessus, l'équipe IMPACT a développé un nouveau système de SA basé sur la réalité virtuelle (RV), adapté à une utilisation en pratique clinique (Verdelet et al. 2019, voir en Annexe A). Les atouts de ce système sont multiples :

- 1) Pouvoir contrôler les informations visuelles délivrées lors d'une tâche de SA : de l'absence d'indices visuels à l'immersion totale dans une scène animée,
- 2) Pouvoir dépasser les limites physiques d'un système de stimulation sonore encombrant, c'est-à-dire tester n'importe quelle position de source sonore de l'espace en manipulant les 3 dimensions (azimut, élévation, et distance),
- 3) Enregistrer en temps réel les comportements d'orientation du sujet face à une stimulation auditive, i.e. orientation du regard et mouvements de tête,
- 4) Avoir la possibilité d'offrir un retour de performance contrôlé (feedback) qui entretient la boucle sensori-motrice (détection-réponse).

2.2. Matériel utilisé

Le système de RV utilisé dans SPHERE est facilement transportable ; il est composé de 2 caméras infra-rouges, espacées d'au moins 4 mètres, enregistrant en temps réel la position de différents objets munis de capteurs (i.e. les objets sont dits « trackés ») : un casque de réalité virtuelle HTC VIVE équipé d'un système d'oculométrie, des manettes, et des trackers (de formes circulaires ceux-ci peuvent être posés sur n'importe quel objet devant être tracké, tel qu'un HP, Figure III.3). L'ensemble du dispositif est piloté par un programme personnalisé développé par la plateforme NeuroImmersion (<https://www.crnl.fr/fr/plateforme/neuro-i>).

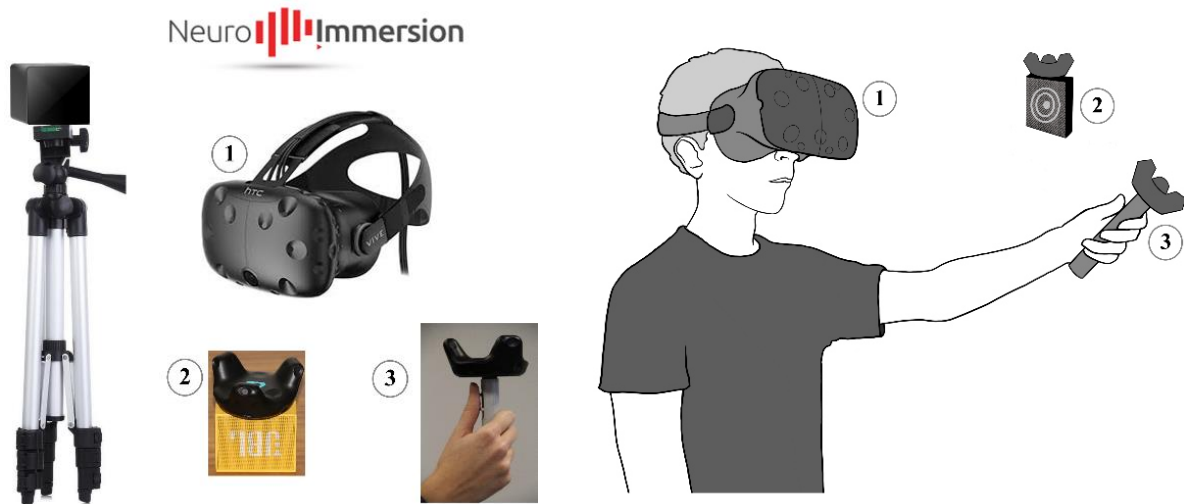


Figure III.3 : Matériel utilisé dans le système de localisation SPHERE

(1) Casque de réalité virtuelle ; (2) Haut-parleur surmonté d'un tracker ; (3) Manette de pointage surmontée d'un tracker.

2.3. Grands principes du système

2.3.1. Tâche de localisation

Le test SPHERE permet de réaliser plusieurs essais successifs afin d'échantillonner l'ensemble de l'espace auditif du sujet. A chaque essai, l'expérimentateur place manuellement un HP (mini haut-parleur modèle JBL GO Portable sans fil) dans une position 3D pré-déterminée à l'avance. Un son est alors émis par le HP et le sujet est libre de rechercher à sa guise l'origine du son par des mouvements de tête non contraints. La phase de recherche et l'émission sonore terminés, le sujet doit montrer l'endroit où il pense avoir localisé la source sonore. Suivant la tâche de localisation choisie, la réponse motrice peut être manuelle (i.e. le sujet montre à l'aide d'une manette la direction et la distance à laquelle le son a été perçu : la réponse motrice est en 3D) ou bien céphalique (i.e. l'orientation de la tête indique la direction du son perçu : la réponse motrice est en 2D). Après la phase de réponse, l'expérimentateur peut choisir d'afficher ou non un feedback dans le casque de RV afin de faire un retour de performance au sujet. Certaines variantes à ce protocole ont été utilisées par d'autres collaborateurs du projet en affichant directement la position de différents HP dans le casque de RV et en invitant les participants à nommer le HP perçu plutôt qu'à le montrer manuellement (Valzolgher et al. 2020a).

Afin que chaque essai soit initié de la même façon, plusieurs contrôles sont mis en place avant l'émission du son. Tout d'abord, le sujet doit positionner sa tête droit-devant lui (sans nécessité d'utiliser une mentonnière comme c'était le cas dans de précédentes études) en alignant deux croix, et regarder droit-devant. Cela permet d'enclencher l'étape finale de positionnement du HP, celui-ci étant toujours placé par rapport au référentiel centre de la tête (Figure III.4).

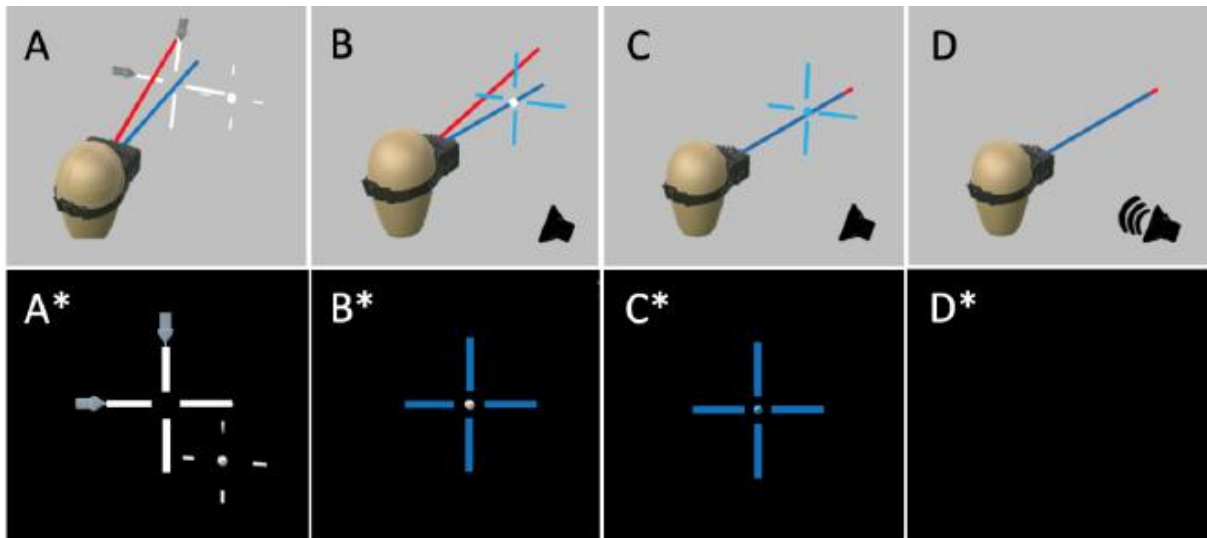


Figure III.4 : Succession des contrôles avant émission du son (Annexe B)

(A) Alignement de la tête (ligne bleue) sur la croix blanche ; (B) Placement manuel du haut-parleur par l'expérimentateur ; (C) Alignement des yeux (ligne rouge) sur la croix bleue ; (D) Emission du son après vérification des 3 contrôles requis (positionnement correct des yeux, de la tête, et du haut-parleur).

2.3.2. Analyses des données

Comme mentionné précédemment, le système SPHERE permet d'enregistrer en temps réel la position de tous les objets trackés, les mouvements de tête et de pointage du sujet. Ces mouvements sont explorés manuellement un à un par le logiciel MAIN développé au sein du laboratoire IMPACT et permettant de réaliser des analyses cinématiques. L'objectif de ces analyses hors-ligne est (a) de quantifier dans l'espace et le temps le déplacement de la tête et de la main essai par essai, et (b) de valider l'exploitation de chaque essai où certaines conditions doivent être respectées :

- 1) Le sujet ne doit pas montrer l'origine du HP avant la fin du son, sous peine de rentrer en collision avec l'expérimentateur et donc d'avoir un retour sur sa performance (Figure III.5B, essai validé ; Figure III.5D, essai rejeté par réponse motrice anticipée),

- 2) Selon les conditions d'expérimentation, les mouvements de tête peuvent être autorisés (Figure III.5A) ou proscrits (Figure III.5C) pendant la phase de recherche du son.

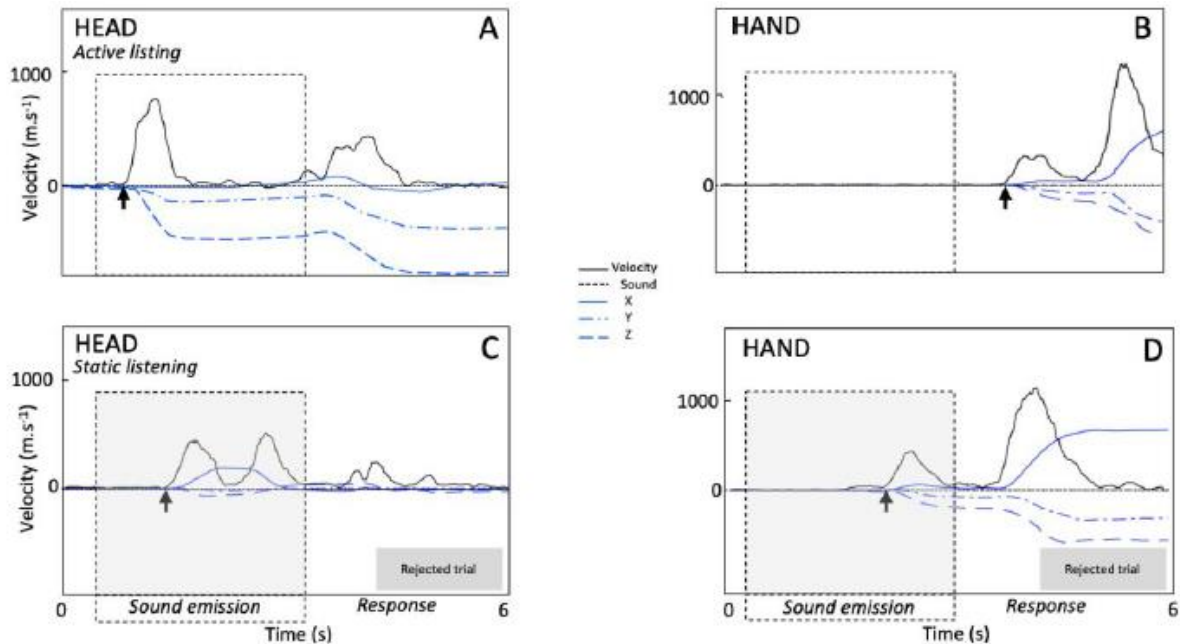


Figure III.5 : Analyse des comportements céphaliques et de pointage (Annexe B)

Les deux graphiques de gauche représentent les mouvements de tête, ceux de droite le pointage de la main. Les courbes bleues correspondent aux coordonnées du mouvement dans chaque plan de l'espace (X, Y, Z) et les courbes noires matérialisent la dérivée associée. Pour chaque graphique, la flèche correspond au début de l'essai et le rectangle gris à la durée d'émission du son.

Les cas A et B représentent un essai validé en condition « mouvements de tête autorisés » : le sujet a bougé la tête pendant l'émission du son (A) et il a bien attendu la fin de celui-ci avant de pointer avec sa manette (B).

Les cas C et D sont deux exemples d'essais rejetés, soit par la présence d'un mouvement de tête pendant l'émission du son alors que ceux-ci étaient interdits (C), soit par un pointage prématuré avant la fin du son (D).

Après suppression des essais invalidés, un calcul de performance est fait en prenant l'ensemble des essais retenus. Le système de RV nous permettant d'enregistrer une réponse motrice en 3D, une erreur de pointage en 3D est calculée (Figure III.6). Cette erreur 3D correspond à la différence dans l'espace entre la position de la source sonore (définie par ses coordonnées X-Y-Z) et le pointage final du sujet. Cette mesure est adaptée du travail de Rakerd and Hartmann (1986). De principe, l'erreur 3D est calculée pour chaque participant, soit pour chaque position de HP, soit pour une portion de l'espace auditif, selon la formule suivante : $\text{erreur 3D} = \sqrt{\underline{C}^2 + \underline{s}^2}$, où \underline{C} est la moyenne de la norme du vecteur entre la source et la position

de pointage, et \underline{s} est la déviation standard. L'intérêt de cette mesure est de pouvoir prendre en compte les 3 dimensions de l'espace (et pas une seule dimension isolée comme c'est souvent le cas avec le calcul en azimut de l'erreur RMS) ainsi que la variabilité de réponse de chaque participant.

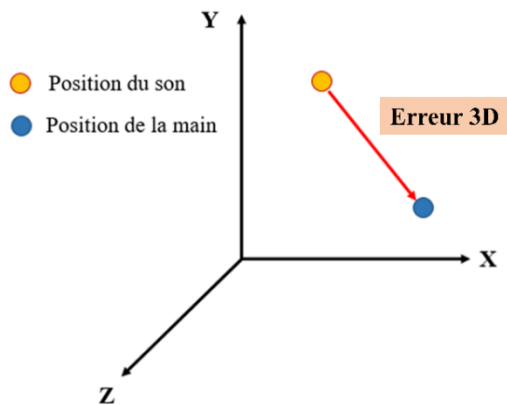


Figure III.6 : Erreur 3D entre la position de la source et la position de la main du sujet

Enfin, l'évaluation quantitative de l'erreur 3D est complétée par une visualisation graphique des performances de pointage de chaque sujet en fonction de la position de chaque HP (e.g., Figure III.7).

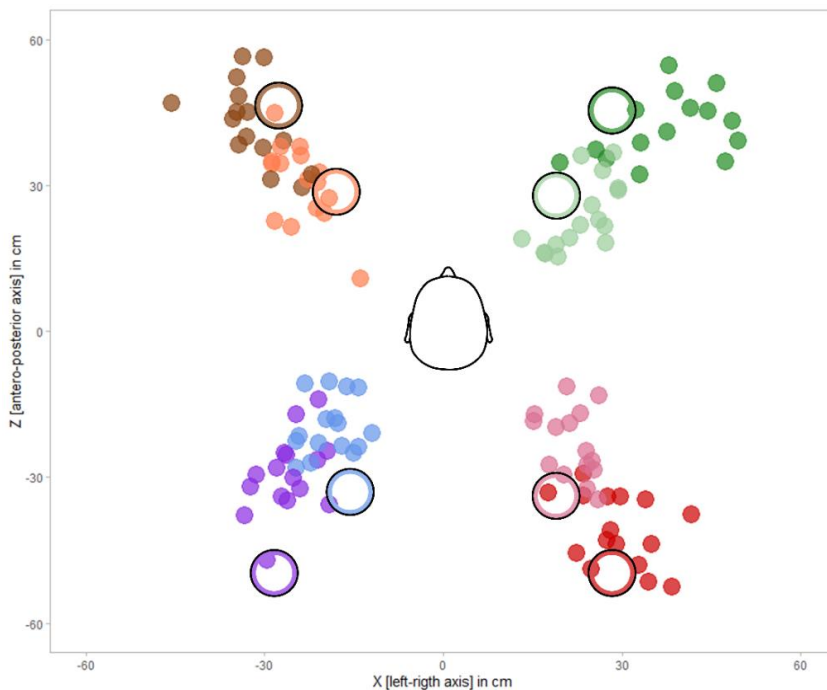


Figure III.7 : Exemple de performances de localisation d'un adolescent normo-entendant de 14 ans en fonction de la position des cibles, vue d'en haut. Les cibles sont représentées par des cercles noirs situées à hauteur d'oreille et à une distance de 35 ou 55 cm. Chaque cercle plein coloré correspond à un pointage du sujet en lien avec le HP d'émission (schéma réalisé à l'aide du logiciel R-Studio).

3. Performances de spatialisation auditive des sujets normo-entendant

Dans la littérature, de nombreuses études ont évalué les performances de spatialisation auditive d'adultes normo-entendant (NE) en azimut, en utilisant comme mesures l'erreur RMS et le pourcentage de confusions avant-arrière (i.e. un HP situé devant le sujet est localisé faussement derrière lui, et inversement). En moyenne, l'erreur RMS était de 5° (Middlebrooks & Green 1991, Blauert 1997, Kerber & Seeber 2012, Dorman et al. 2016). Avec notre système SPHERE, l'erreur RMS était plus élevée, de $26,1^\circ$ pour les HP avants à $17,1^\circ$ pour les HP arrières (Annexe B). Comparée aux précédentes études, cette différence de performances s'explique par l'existence d'une incertitude complète sur la position des HP (aucun indice visuel présent) et le cône d'ombre généré par le casque de RV. Ce phénomène a déjà été montré dans une autre étude utilisant la RV (Ahrens et al. 2019), et semble plus marqué pour les sources latérales en condition tête immobile pendant l'émission du son (Annexe B).

Selon les études, le pourcentage de confusions avant-arrière chez l'adulte NE peut varier de 0 à 30% suivant la position des HP et le type de son utilisé, les sons à large spectre étant plus facilement identifiables, de même que les sons durant plus de 500 ms (Middlebrooks & Green, 1991, Carlile 1997, Perrett & Noble 1997a, Fischer et al. 2020). La capacité à explorer activement une scène auditive varie d'une personne à l'autre, mais la plupart des adultes NE oriente naturellement leur tête vers une source sonore (Fuller 1992), leur permettant d'améliorer de façon notable leurs performances de SA (Thurlow & Runge 1967, Lambert 1974, Noble 1981, Perrett and Noble 1997a,b) avec diminution des confusions avant-arrière (Wightman & Kistler 1999, Brimijoin & Akeroyd 2012, Kim et al. 2013). Les bénéfices des mouvements de tête apparaissent seulement si le sujet est capable d'interpréter en temps réel les changements d'indices spectraux et binauraux, et de mettre à jour en continu ses cartes auditives spatiales.

Chez l'enfant, les performances de localisation spatiale absolue en azimut atteignent celles de l'adulte (i.e. erreur RMS moyenne de 10°) vers l'âge de 5 ans. Ce constat provient de plusieurs études qui ont comparé l'erreur RMS en azimut dans une population pédiatrique âgés de 1,5 à 8 ans (Lovett et al. 2012) ou 7 à 12 ans (Otte et al. 2013) versus de jeunes adultes NE (entre 20 et 34 ans). Il faut toutefois noter que, même si la moyenne de l'erreur RMS est proche de celle de l'adulte jeune, il existe une importante variabilité inter-individuelle chez l'enfant (Figure III.8) avec des erreurs pouvant osciller entre 9° et 29° (Grieco-Calub & Litovsky 2010). Cette variabilité peut en partie s'expliquer par la variété des tests de localisation, les HP étant espacés de 10° à 30° (Van Deun et al. 2019, Litovsky & Godar 2010, Grieco-Calub & Litovsky 2010, Kühnle 2012) ce qui diminue ou augmente la difficulté. Il est également possible que les

différentes tranches d'âge testées aient renforcé la disparité des résultats entre les études. En effet, deux catégories d'âges se détachent, les 4 à 6 ans (Van Deun et al. 2009, Litovsky & Godar 2010, Grieco-Calub & Litovsky 2010), et les 7-12 ans (Otte et al. 2013) voire les adolescents (Kühnle et al. 2012). Certains auteurs ont également constaté qu'à l'échelle individuelle, la variabilité des réponses entre les essais diminuait avec l'âge pour atteindre celle de l'adulte NE à l'âge de 8 ans (Kühnle et al. 2012, Freigang et al. 2015). Ces données suggèrent que la maturation de la SA surviendrait progressivement à plusieurs niveaux : sur le plan de la précision de la localisation mais aussi en terme de variabilité des réponses.

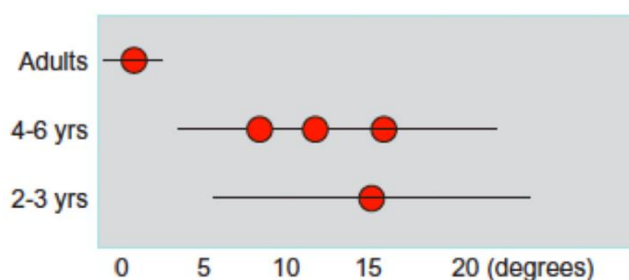


Figure III.8 : Erreur RMS en fonction de l'âge (d'après Litovsky 2015)

A l'heure actuelle, aucune étude pédiatrique n'a apporté de données sur l'acquisition de la discrimination des sources avant versus sources arrière ni étudié l'impact des mouvements de tête sur les performances de SA. Une étude ancienne (Morrongiello & Rocca 1987) a simplement constaté que la rapidité d'orientation de la tête en direction d'un son ne variait pas selon l'âge de l'enfant.

4. Performances de spatialisation auditive des patients implantés cochléaires

4.1. Adultes mono et bi-implantés cochléaires

Les capacités de SA des patients bi-implantés cochléaires (BCI) sont bien meilleures que celles des mono-implantés (UCI). Ainsi l'accès à la binauralité restaure aux BCI une discrimination des sons entre les héli-champs auditifs gauche et droit (Gantz et al. 2002, Tyler et al. 2002). De plus, face à une rangée de HP, l'erreur RMS en azimuth est plus faible chez les BCI, en moyenne de 30°, versus 50° chez les UCI (van Hoesel & Tyler 2003, Litovsky et al. 2004, Nopp et al. 2004, Verschuur et al. 2005, Grantham et al. 2007, Neuman et al. 2007, Litovsky et al. 2009). Néanmoins, l'accès aux indices binauraux ne permet pas aux BCI d'atteindre les performances des NE, notamment au cours des tests de localisation dans le bruit (Mosnier et al. 2009, Zheng et al. 2011, Kerber & Seeber 2012).

Il existe une importante variabilité dans les performances enregistrées dans la littérature des BCI. Dans la majorité des cas, celles-ci sont évaluées précocement après la pose du deuxième implant, de 1 mois à 9 mois (Litovsky et al. 2004, Nopp et al. 2004, Verschuur et al. 2005). Il est probable qu'au moment des évaluations précoces (i.e. moins de 12 mois après la bilatéralisation), les déficits de SA soient surestimés par le manque d'expérience binaurale. En effet, les capacités de spatialisation des implantés mettent au moins 12 mois avant de se stabiliser (Zhang et al. 2015).

Les bénéfices de mouvements de tête dans les capacités de SA d'adultes BCI ont été peu explorés dans la littérature (Mueller et al. 2014; Pastore et al. 2018). Dans ces études, les patients étaient libres de bouger la tête pendant l'émission d'un son dont la durée était de 0,5s, 2s, 3s ou 4,5s. L'exploitation des indices acoustiques dynamiques permettait aux patients BCI de diminuer leurs confusions avant-arrière de 25% à 5,5% pour l'étude de Mueller (2013), et de 41,9% à 6,7% pour l'étude de Pastore (2018). Ce bénéfice n'était visible que pour des sons suffisamment longs, strictement supérieur à 0,5s.

4.2. Enfants bi-implantés cochléaires

L'évaluation des performances de SA chez l'enfant IC a été réalisée pour la première fois par Litovsky et collaborateurs (2004) chez 3 enfants bi-implantés depuis 3 mois en utilisant une rangée de HP couvrant l'azimut de -70° à $+70^\circ$ tous les 15° . Les auteurs retrouvaient un bénéfice très limité de la bilatéralisation, très certainement du fait d'une évaluation très précoce. Par la suite, plusieurs études ont montré qu'après la pose du 2^e implant, l'erreur RMS en azimut passait en-dessous du seuil de la chance ($< 50^\circ$) avec une moyenne de 30° (Van Deun et al. 2010, Grieco-Calub & Litovsky 2010, Zheng et al. 2015, Choi et al. 2017) comme l'adulte BCI. A l'échelle individuelle, certains enfants BCI s'approchaient même des performances des enfants NE avec une erreur RMS comprise entre 10° et 15° (Van Deun et al. 2010, Zheng et al. 2015, Killan et al. 2018). Ce résultat était visible pour les enfants ayant reçu leur 1^e IC avant l'âge de 2 ans et ayant été stimulé par un appareillage auditif conventionnel les 18 premiers mois de vie (Van Deun et al. 2010), c'est-à-dire lors de la période critique de la maturation auditive.

Toutefois, contrairement à la littérature adulte, nous avons constaté que très peu d'études avaient comparé les performances de localisation spatiale absolue des enfants UCI vs. BCI (Beijen et al. 2007, Galvin et al. 2007). A notre connaissance, Galvin et collaborateurs (2007) ont été les seuls à monitorer l'évolution des capacités de SA chez les mêmes enfants avant et

après bi-implantation. A 6 mois post-implantation, la restauration de la binauralité avait permis de diminuer l'erreur RMS chez seulement 2 patients sur 8. Il est probable que l'émergence des bénéfices soit survenue plus tardivement pour les autres enfants.

L'ensemble de ces études illustre bien la latence du système de SA chez l'enfant pour recréer des cartes auditives spatiales en cas de privation auditive ainsi que de l'importante variabilité inter-individuelle dans les performances mesurées à un instant T.

Nous n'avons pas retrouvé dans la littérature pédiatrique de données sur les capacités de discrimination des sources avant et arrière, ni sur les bénéfices des mouvements de tête dans la réduction des erreurs de localisation.

5. Evaluation de la qualité de vie

5.1. Généralités

L'Organisation Mondiale de la Santé (OMS) a défini la qualité de vie comme « la perception qu'a un individu de sa place dans l'existence, dans le contexte de la culture et du système de valeurs dans lequel il vit, en relation avec ses objectifs, ses attentes, ses normes et ses inquiétudes » (WHOQOL Group 1993). La préservation de la qualité de vie prend tout son sens lorsqu'il s'agit de prendre en charge une pathologie chronique telle que la surdité. Dans son rapport de 2018, l'OMS a souligné qu'une personne sur 10 sera atteinte de déficience auditive en 2050, et elle a demandé aux acteurs des pays membres (professionnels de la santé, société civile et politiques) d'apporter une contribution à la promotion de diagnostic précoce et de soins spécialisés, et de veiller à l'accès et à l'utilisation des appareils auditifs et au recours aux services de réadaptation.

A l'heure où les dispositifs médicaux connaissent un essor technologique majeur, les patients malentendants doivent prétendre à une qualité de vie satisfaisante sur le plan physique, psychique et social. L'évaluation de leur qualité de vie passe par l'utilisation d'outils standardisés, les questionnaires de qualité de vie qui peuvent être génériques ou spécifiques d'une pathologie donnée.

Les questionnaires de qualité de vie générique ont le mérite d'être transposables à un large panel de pathologies et d'être normés selon différentes populations et tranches d'âge. Citons pour exemple le Health Utility Index 3 (Feeny et al. 1995) qui a été utilisé pour évaluer les bénéfices de l'implantation cochléaire (Ramakers et al. 2016). L'inconvénient majeur de ce type de

questionnaire est l'absence de spécificité (Wiebe et al. 2003). Aussi, dans le monde de l'audition, plusieurs questionnaires spécifiques ont été développés pour évaluer plus précisément à un instant T et dans le temps les capacités de communications orales, d'identifications sonores, de localisation spatiale mais aussi les efforts d'écoute fournis par les patients. Ces évaluations séquentielles sont primordiales pour impliquer les patients et leurs familles dans la prise en charge d'une pathologie chronique telle que la surdité. De plus, les questionnaires de qualité de vie offrent une vision complémentaire aux évaluations perceptives (e.g. audiométries) réalisées en pratique clinique. D'ailleurs, les dernières recommandations de la SFORL (2018) soulignent la nécessité d'utiliser de tels outils pour suivre les patients IC.

5.2. Questionnaires spécifiques chez l'adulte

Dans le cadre de cette thèse, nous nous intéresserons aux questionnaires validés traitant de la localisation spatiale des sons. Ces questionnaires sont très peu nombreux dans la littérature et montrent à quel point ce domaine de l'audition a été mis de côté ces dernières années au profit des bénéficiaires de la communication orale.

En 2004, Gatehouse et Noble ont créé le questionnaire SSQ (Speech, Spatial and Qualities of Hearing Scale) dédiant pour la première fois plusieurs items à la localisation d'une source sonore dans différentes situations du quotidien. Le SSQ est composé de 49 questions réparties en 3 sections : section A – Audition de la parole » (14 items), section B – Audition spatiale (17 items), et section C – Qualités d'audition (18 items). Chaque item est coté de 0 à 10 (10 correspondant à une absence de difficulté dans la situation citée). Une version française a été validée en 2015 (Moulin et al. 2015). Cet outil a montré que les scores des sections A et B étaient plus élevés chez des patients bi-implantés que chez des mono-implantés (van Zon et al. 2017).

Néanmoins, la passation du SSQ est chronophage et peu compatible avec la réalité de terrain en milieu hospitalier. Aussi, plusieurs équipes ont développé des formes raccourcies du SSQ en ne sélectionnant qu'une partie des questions de la forme longue. Citons comme exemples le SSQ5 mis au point par une équipe néerlandaise (Demeester et al. 2012), le SSQ12 par une équipe anglaise (Noble et al. 2013), et le SSQ15 par une équipe française (Moulin et al. 2019). Le deuxième questionnaire centré sur la spatialisation est le SHQ (Spatial Hearing Questionnaire). Il a été créé par Tyler et collaborateurs en 2009 (Tyler et al. 2009) et validé chez des patients bi et mono-implantés cochléaires. Le SHQ comprend 24 items ciblant spécifiquement les capacités de compréhension de la parole en milieux calmes et bruyants, ainsi

que la localisation spatiale. Chaque item est coté de 0 à 100 (100 correspondant à une absence de difficulté dans la situation citée). A noter qu'une forme courte du SHQ a récemment été développée (Ou et al. 2017) mais son utilisation ne semble pas encore démocratisée dans les études cliniques.

Plusieurs études ont montré que les scores obtenus aux questionnaires SSQ et SHQ étaient fortement corrélés (Tyler et al. 2009, Richard et al. 2014, Zhang et al. 2015).

5.3. Questionnaires spécifiques en pédiatrie

L'évaluation de la qualité de vie en pédiatrie soulève un écueil méthodologique principal : à quel âge peut-on considérer qu'un enfant est capable de remplir seul un questionnaire en faisant part le plus objectivement possible de ses difficultés du quotidien ?

Les premières hétéro-évaluations ont été réalisées chez des enfants âgés de 4 à 15 ans ayant bénéficié d'une bi-implantation séquentielle (Galvin et al. 2007). Pour les besoins de l'étude, les auteurs ont modifié le questionnaire SSQ de Gatehouse pour que celui-ci soit directement rempli par les parents, 6 semaines avant et dans les 6 mois suivant la deuxième implantation. Un temps d'observation d'une semaine était dédié afin que l'évaluation des parents soit la plus proche possible de la réalité du quotidien de leur enfant. Cette étude a montré que d'importants changements se produisaient les premiers mois post-implantation avec une nette amélioration de la qualité de vie en termes de compréhension de la parole en groupes et de localisation spatiale. Des études similaires ont mis en évidence les bénéfices de la bi-implantation cochléaire jusqu'à 2 ans post-implantation, particulièrement sur les items traitant de la localisation spatiale (Beijen et al. 2007, Lovett et al. 2010, Sparreboom et al. 2012, Galvin et al. 2016). La plupart de ces études ont apporté une vision nouvelle sur une tranche d'âge beaucoup plus basse, les moins de 5 ans (Beijen et al. 2007, Lovett et al. 2010, Sparreboom et al. 2012). Les bénéfices en termes de compréhension de la parole ne sont pas forcément retrouvés dans ces études quelle que soit la tranche d'âge testée. Chez les moins de 5 ans, cela peut s'expliquer par la difficulté pour les parents de coter certains items sur la capacité à suivre une conversation de groupes. Pour les plus grands, une grande majorité était implantée tardivement et évaluée après plusieurs années d'expérience binaurale, ce qui représente deux facteurs confondants pour mettre en évidence un éventuel bénéfice perceptif.

Chez les adolescents, il est légitime de s'interroger sur la fiabilité des réponses parentales par rapport à l'auto-évaluation réalisée par l'enfant. Ainsi, Galvin et Noble ont validé en 2013 une

version pédiatrique du SSQ comprenant 33 questions réparties dans les 3 sections sus-citées (section A : 9 items, section B : 6 items, C : 8 items). Plusieurs items de la version adulte de Gatehouse ont été supprimés car ils n'étaient pas adaptés au quotidien d'un enfant (e.g. conduire une voiture). Notons que 23 items sont communs avec la version SSQ pour les parents. Comme pour le remplissage du SSQ chez l'adulte (Gatehouse & Noble 2004), les auteurs recommandent de faire passer le SSQ en face à face plutôt que de laisser l'enfant le remplir seul. Cela permet de lever certaines ambiguïtés à la lecture des items, tout en prenant garde de ne pas influencer les réponses de l'enfant. L'utilisation d'une règle chiffrée permet à l'enfant de mieux visualiser les différents niveaux de difficulté, à l'image des échelles EVA utilisées dans l'évaluation de la douleur.

Les questionnaires d'auto-évaluation posent le problème d'âge minimum. Plusieurs auteurs suggèrent que ces systèmes de mesure ne doivent pas être utilisés chez les moins de 8 ans (Kessler et al. 1990, Arlinger 2001). Galvin et Noble (2013) considèrent même que certains items du SSQ sont difficiles à appréhender et que le temps de remplissage reste relativement long pour des enfants pouvant présenter des troubles attentionnels. Aussi, ils recommandent de ne pas proposer le SSQ pédiatrique aux enfants de moins de 10 ans. A l'image des travaux réalisés chez l'adulte, il semblerait intéressant de développer une forme courte du SSQ pédiatrique afin d'étendre son utilisation en pratique clinique et permettre l'auto-évaluation des 8-10 ans.

6. Points clés du chapitre

- a) La plupart des outils ayant été utilisés jusqu'à maintenant dans l'évaluation des performances de spatialisation auditive ne permettait qu'une appréciation partielle des compétences en azimut et en élévation des sujets.
- b) L'outil SPHERE, basé sur la réalité virtuelle, permet de monitorer avec précision les capacités de spatialisation auditive en 3D, tout en contrôlant les informations visuelles et l'impact des mouvements de tête.
- c) La précision de la spatialisation auditive de l'enfant NE rejoint celle de l'adulte vers l'âge de 5 ans (i.e. erreur RMS $\approx 10^\circ$) avec une nette diminution dans la variabilité des réponses au cours d'un même test dès l'âge de 8 ans.
- d) Les performances de localisation des adultes BCI sont bien meilleures que celles des UCI (i.e. erreur RMS $\approx 30^\circ$ et 50° respectivement) sans atteindre toutefois celles des NE.
- e) La recherche dynamique d'un son par des mouvements de tête orientés permet aux adultes IC de réduire leurs déficits de spatialisation auditive. Aucune donnée n'est actuellement connue chez les enfants IC.
- f) Les questionnaires spécifiques des compétences auditives sont très limités en pédiatrie. L'évaluation de la qualité de vie des enfants IC repose essentiellement sur l'hétéro-évaluation réalisée par les parents.

Chapitre 4

Rééducation de la spatialisation auditive

1. Grands principes de la rééducation

De manière générale, la rééducation d'un déficit sensoriel vise à prendre en charge une situation de handicap de manière non médicamenteuse dans le but de regagner une partie de la fonction perdue (restitution) et/ou de compenser par le biais d'autres afférences sensorielles intègres (compensation). Elle repose sur deux principes fondamentaux : la plasticité cérébrale et l'intégration multisensorielle.

Le phénomène de plasticité cérébrale a longuement été décrit dans la littérature pédiatrique (e.g. pour revues Greenough et al. 1987, Galvan 2010, Fandakova & Hartley 2020). Chez l'adulte, des processus similaires de réorganisation neuronale associée à des modifications comportementales ont été mis en évidence à la suite de contraintes internes (e.g. vieillissement) et externes (e.g. un entraînement spécifique) (e.g. pour revues Gage 2004, Lövdén et al. 2013). La plasticité cross-modale en est le parfait exemple ; ainsi les zones corticales non stimulées par la modalité sensorielle altérée sont recrutées par d'autres systèmes sensoriels intacts. Dans le cadre de la surdité profonde congénitale ou acquise plus tardivement à l'âge adulte, plusieurs auteurs ont montré que les structures auditives centrales étaient recrutées pour l'analyse des informations visuelles et somesthésiques (Campbell & Sharma 2014, Glick & Sharma 2017).

Comme nous l'avons mentionné dans le chapitre 1 (paragraphe 5), la perception optimale de notre environnement est basée sur l'intégration de plusieurs afférences sensorielles. Les capacités d'intégration multisensorielle ne sont pas innées et mettent en jeu des interactions entre structures corticales et sous corticales au cours du développement post-natal (pour revue, voir Perrault et al. 2012). L'expérience précoce des signaux intermodaux est essentielle pour le développement multisensoriel normal de toutes les espèces d'ordre supérieur. Lorsque l'une de ces afférences est déficitaire, l'intégration des afférences intactes dans le protocole d'entraînement permet de renforcer la réorganisation corticale en sollicitant les structures d'intégration multisensorielle, limitant aussi la fatigabilité liée à l'attention soutenue déployée vers le sens déficitaire.

Quel que soit le domaine, un protocole de rééducation ou d'entraînement doit répondre à plusieurs critères cliniques garantissant sa qualité : apparition d'effets bénéfiques, persistance de ces effets à moyen voire long terme (de quelques semaines à plusieurs mois selon les études, Mendonça 2014), et adhésion du sujet au protocole.

Les bénéfices attendus d'une rééducation suite à un déficit auditif sont quantifiables à plusieurs niveaux : (1) ils peuvent directement conduire à une amélioration des performances sensorielles

travaillées lors de l'entraînement (phénomène de restitution) ; (2) les compétences nouvellement ré-acquises peuvent également s'étendre à d'autres fonctions sensorielles non travaillées par le biais des interactions sensorielles survenant au cours de la rééducation. Certains auteurs parlent de transfert de compétences ou de généralisation de l'entraînement (Wright & Zhang 2009, Mendonça et al. 2013). Ces effets sont parfois difficiles à mettre en évidence et ne sont pas souvent mesurés dans les études cliniques de rééducation auditive. L'adhésion du sujet (i.e. maintien dans le temps de sa participation au protocole de rééducation) et sa motivation (i.e. facteur déclenchant pour participer au protocole) sont deux points critiques dans la réussite ou l'échec d'un programme d'entraînement, et ce d'autant plus que celui-ci s'étale dans le temps (Amitay et al. 2010, Henshaw et al. 2016).

2. Place de la rééducation spatiale dans la prise en charge rééducative des patients implantés cochléaires

Quel que soit l'âge du patient, la précocité de la prise en charge rééducative post-implantation cochléaire est un facteur clé de réussite (Truy et al. 2017). Cette prise en charge fait écho à la maturation des structures nerveuses et notamment celles impliquées dans l'intégration multisensorielle (voir chapitre 1). A l'heure actuelle, il n'existe pas de recommandation internationale décrivant précisément les étapes incontournables de prise en charge. En France, les orthophonistes ont un rôle central dans la rééducation auditive des patients implantés cochléaires (IC). Même s'il n'existe pas de consensus national, plusieurs centres français ont édité des guides à l'attention des patients et des orthophonistes exerçant en milieu libéral, afin de les accompagner au mieux dans la rééducation au long cours (e.g. en Ile de France, à Lyon). Ainsi, les séances sont centrées sur la réhabilitation de la parole avec un travail débutant sur la discrimination de sons puis des phonèmes. Par la suite, les exercices se complexifient avec reconnaissance de phrases délivrées directement par la voix de l'orthophoniste, ou plus difficile, via un média (e.g. radio, téléphone). Lorsque les performances du patient ont bien évolué (souvent après 1 an d'implantation) une place importante est attribuée à la compréhension de la parole dans le bruit, à l'utilisation du téléphone, et si le patient le souhaite, à la reconnaissance de la musique.

Une veille de la littérature et un sondage local auprès de plusieurs orthophonistes libérales ont mis en lumière un manque notable d'exercices centrés sur l'audition spatiale au cours du projet rééducatif des patients IC. Ce manque est très certainement plurifactoriel avec au premier plan des professionnels de la rééducation peu sensibilisés aux difficultés de spatialisation auditive

(SA) éprouvés par les patients. La plainte des patients n'émerge pas toujours spontanément ; il est donc important de les interroger sur des situations du quotidien par le biais de questionnaire de qualité de vie (voir chapitre 3, paragraphe 5.2.). A cela s'ajoute l'accès limité au matériel de localisation spatiale qui est coûteux et encombrant si l'on envisage une installation immersive avec plusieurs haut-parleurs (HP) fixes.

Au total, la place clinique de la rééducation spatiale reste anecdotique dans le parcours de soins du patient IC. Une prise de conscience tend à émerger ces dernières années grâce aux tests d'évaluation de la SA et ont conduit plusieurs équipes de recherche à développer des protocoles fondamentaux, principalement sur la base de modèles normo-entendant (NE).

3. Protocoles d'entraînement de la spatialisation auditive en recherche

Les protocoles d'entraînement de la SA ont émergé ces deux dernières décennies au sein des équipes de recherche et ont fait l'objet de nombreuses publications dans la population NE. En effet, c'est à travers les approches fondamentales réalisées avec des modèles humains de surdité (i.e. sujet NE pluggé donc porteur d'un bouchon d'oreille mimant une perte auditive) que les phénomènes de récupération spontanée et de plasticité auditive ont été appréhendés. A contrario, peu d'études se sont intéressées à la prise en charge de la SA des patients déficients auditifs alors que la nécessité de cette prise en charge est manifeste dans cette population. Ce décalage peut s'expliquer certainement par la difficulté de développer des protocoles qui s'adapteront à la majorité des patients (e.g. notamment à l'hétérogénéité des réglages des aides auditives, des étiologies, etc.).

Dans ce paragraphe, nous étudierons quelques exemples typiques de protocoles réalisés chez les sujets NE, puis nous reprendrons les différents protocoles de SA réalisés chez des patients atteints de surdité. L'ensemble des études discutées est synthétisé dans les tableaux IV.1 et IV.2.

3.1. Chez des sujets normo-entendant

De très nombreuses études ont été réalisées chez des NE pour apprécier les bénéfices de certains types d'entraînements auditifs. Dans cette approche, plusieurs techniques ont été utilisées pour dégrader les indices acoustiques perçus par les sujets et ainsi créer des modèles de surdité. La mise en place d'un bouchon d'oreille génère artificiellement une surdité unilatérale de transmission et altère les indices binauraux (ILD principalement et ITD dans une moindre mesure car le bouchon atténue mieux les hautes fréquences). La dégradation des

indices monauraux (e.g. appauvrissement spectral) est obtenue via des écouteurs de manière uni ou bilatérale suivant le protocole. Plus récemment, un système de vocodeur a été développé afin de mimer la dégradation acoustique des indices binauraux et monauraux générée par l'implant cochléaire. Ce modèle reproduit assez bien les déficits spatiaux expérimentés par les patients IC (Ausili et al. 2019) mais il ne reflète pas avec fidélité la qualité de la stimulation électrique (Dorman et al. 2017).

L'objectif de cette partie n'est pas de reprendre de manière exhaustive l'ensemble des protocoles d'entraînement de la SA développés chez les NE car ceux-ci sont très variés et testent des paramètres bien différents entre les études. Il nous paraissait cependant intéressant de se focaliser sur quelques exemples ayant permis de mieux comprendre les points forts et écueils à éviter dans le développement d'un entraînement spatial.

Les études reprises dans le tableau IV.1 ne concernent que des NE pluggés ou équipés d'écouteurs. Nous n'avons pas connaissance pour l'instant de protocoles d'entraînement spatial ayant utilisé le système de vocodeur ; les seules études retrouvées dans la littérature sont des entraînements auditifs perceptifs (Fu & Nogaki 2005, Stacey & Summerfield 2007, 2008).

En résumé, les entraînements proposés dans ces études ont montré un bénéfice perceptif lorsque l'entraînement spatial durait plus d'une journée, quand le feedback était multisensoriel, et lorsque le sujet était actif dans son environnement en interagissant spatialement avec la source sonore (i.e. mise en jeu de boucles sensori-motrices avec des mouvements céphaliques ou manuels).

Très récemment, Valzolgher et collaborateurs ont mené un entraînement de la SA dans l'espace azimutal chez des sujets NE pluggés unilatéralement (*article en cours d'écriture*). Lors de la séance d'entraînement, les sujets devaient « réapprendre » à spatialiser des sources sonores en azimut. Ainsi, des HP leur étaient présentés en RV et ils devaient indiquer manuellement celui qui était à l'origine du son délivré ; leur réponse était ensuite corrigée par un indice visuel qui leur indiquait la position spatiale de la source sonore. Les résultats obtenus ont montré que les sujets pouvaient re-spatialiser après une seule séance d'entraînement, et que cet effet pouvait se généraliser à une autre réponse motrice (tête) et d'autres cibles sonores frontales ; mais cet effet ne se généralisait pas à une autre tâche auditive spatiale. Au total, cette étude montre qu'un entraînement faisant suite à une altération auditive subite entraîne des modifications rapides de la SA. Mais les effets induits par l'entraînement ne sont peut-être pas assez robustes ou suffisants pour se généraliser à d'autres tâches auditives.

Tableau IV.1 : Exemples d'entraînement de la spatialisation auditive chez des sujets normo-entendant.

ETUDES		Kumpik et al. 2010	Strelnikov et al. 2011	Zahorik et al. 2006	Steadman et al. 2019	Valzolgher et al. 2020b	
Sujets		N=20 Age moyen 23 ans	N=18 20-40 ans	N=20 21-41 ans	N=36 18-38 ans	N=28 18-30 ans	
Entraînement	Groupes	N=3, GC=0	N=3, GC=0	N=2, GC=1	N=4, GC=1	N=2, GC=0	
	Variation de quels paramètres	durée d'entraînement	type de stimulation sensorielle	ajout ou non d'un feedback	type d'environnement visuel / exploration active ou non	type de réponse motrice	
	Généralités	Lieu Durée/fréquence	Laboratoire 1-7-8 jours	Laboratoire 5 jours	Laboratoire 2 jours	Laboratoire 3 jours	Laboratoire 1 jours
	Phase recherche	Uni/Multisensorielle Type espace rééduqué Son délivré MVT tête autorisés	auditif seul azimut 360° (12 HP) bandes larges non	visuo-auditif / ou auditif seul azimut avant (15 HP) bruit blanc non	visuo-auditif (visuel simple) azimut 360° + élévation à ±40° bruit blanc (son virtuel) oui	visuo-auditif (visuel complexe) azimut 360° + élévation à ±45° bruit rose + parole oui (1 seul groupe)	interaction motrice /lecture azimut avant (17 HP) bruit blanc oui
Feedback		visuel (position correcte du HP)	verbal (correct/incorrect) + visuel pour GC	audiovisuel (en RV : position exacte HP) OU aucun pour GC	visuel (position correcte HP)	audiovisuel (en RV)	
Résultats		Bénéfice si étalement des séances dans le temps	Bénéfice différent de l'entraînement multisensoriel et du feedback visuel	Bénéfice du feedback sur les confusions avant-arrière ; Pas d'effet sur l'élévation ; Performances non modifiées dans le GC	Pas d'effet de l'aspect jeu vidéo mais bénéfice additionnel des MVT de tête Performances non modifiées dans le GC	Bénéfice si interaction motrice avec le son	

GC : groupe contrôle ; HP : haut-parleur ; MVT : mouvements ; N : nombre

Dans les études de Zahorik et collaborateurs (2006) le stimulus visuel est qualifié de « simple » dans la phase de recherche du son car il consistait en un point lumineux fixe situé droit devant le sujet ; le groupe contrôle de cette étude ne recevait pas de feedback.

Dans l'étude de Steadman et collaborateurs (2019) le stimulus visuel est qualifié de « complexe » dans la phase de recherche du son car la scène visuelle virtuelle comportait de nombreux détails réalistes (géométrie, objets, etc) ; le groupe contrôle de cette étude ne réalisait aucun entraînement.

De manière générale, les résultats issus des modèles « sujet sain pluggé » doivent être interprétés avec précaution. En effet, la surdité induite par la pose d'un bouchon d'oreille induit une plasticité immédiate qui conduit à des réorganisations fonctionnelles rapides des circuits existants (modèle par levée d'inhibition GABA, voir pour revue Pascual-Leone et al. 2005). A contrario, la plasticité des patients implantés se met en place sur le long terme car elle est basée sur une réorganisation en profondeur des circuits cérébraux (e.g. Strelnikov et al. 2015). De plus, l'IC appauvrit drastiquement les indices acoustiques nécessaires à la SA (voir chapitre 2) de manière incomparable au bouchon d'oreille.

3.2. Chez des sujets atteints de surdité

3.2.1. Patients implantés cochléaires

Les entraînements de la SA n'ont fait l'objet que de cas reportés chez des patients bi-implantés (Tyler et al. 2010) et chez un patient mono-implanté dans le cadre d'une surdité unilatérale (Nawaz et al. 2014).

Le système d'entraînement à deux sources sonores (ou HP) utilisé par Tyler et collaborateurs (2010) a subi plusieurs évolutions avant d'être éprouvé au domicile des patients IC. Leur objectif était de promouvoir la bonne localisation et compréhension de la parole dans le bruit. Dans un travail pilote, 3 adultes bi-implantés cochléaires (BCI) suivaient un programme d'entraînement en laboratoire avec 8 sources sonores différentes (pendant 1 mois) suivi par un relai au domicile (pendant 2 mois). Ce programme était composé de deux modules principaux : localisation auditive spatiale et reconnaissance de la parole dans le bruit. Les patients étaient libres dans le choix des modules et le niveau de difficulté appliqué lorsqu'ils s'entraînaient. Cette étude pilote a montré la faisabilité de ce type d'entraînement chez des patients IC mais ces derniers ont critiqué la difficulté d'utilisation de ce système à 8 HP à domicile. De plus, les bénéfices escomptés étaient modérés puisqu'un seul participant a amélioré ses performances de SA. L'absence d'amélioration pour les deux autres patients était certainement plurifactorielle car ces derniers n'ont pas suivi le programme dans son intégralité et il est possible (même si cela n'est pas relevé par les auteurs) que le temps d'entraînement passé avec le module de localisation était moindre que celui passé avec le module de compréhension de la parole. Par la suite, les auteurs ont réalisé la même étude avec un système à 2 HP au domicile de 9 patients BCI (voir Tableau IV.2 pour un résumé des différents paramètres utilisés). Sur les 3 patients ayant suivi l'entraînement, 2 ont amélioré leurs performances de SA. Il est intéressant de noter

que les auteurs ont utilisé un test 2 en 1 incluant la bonne localisation de la source et la reconnaissance du mot délivré par le HP. Les 6 autres patients étaient dans le groupe contrôle et ne réalisaient aucun entraînement ; leurs performances ne se sont pas modifiées entre les évaluations pré et post-entraînement.

Dans l'étude de Nawaz et collaborateurs (2014), un patient de 57 ans mono-implanté à la suite d'une surdité unilatérale brusque d'origine inconnue a bénéficié d'un entraînement spatial à domicile dans les suites immédiates de l'implantation. Les modalités de l'entraînement étaient peu détaillées mais après 3 mois de pratique, le patient réussissait à discriminer les sons venant de la droite et de la gauche. L'imputabilité de l'entraînement en termes de bénéfices constatés reste discutable car plusieurs auteurs ont montré que les mois suivant une implantation cochléaire sont marqués par une amélioration spontanée des performances grâce à l'expérience auditive du quotidien (Chang et al. 2010, Zhang et al. 2015). Il est toutefois possible que l'entraînement ait potentialisé les capacités d'apprentissage naturelle du patient.

En résumé, ces publications de cas cliniques ont montré la faisabilité d'un protocole d'entraînement de la SA au domicile de patients IC mais la généralisation des bénéfices à l'échelle de la population d'implantés reste limitée. On peut notamment relever une grande variabilité inter-individuelle des performances initiales des patients et le manque de contrôle sur leur compliance au traitement (i.e. exercices réalisés).

En complément de ces deux études cliniques, une étude animale a été réalisée chez des furets rendus sourds après la naissance et implantés unilatéralement (UCI) ou bilatéralement (BCI) à l'âge adulte (Isaiah et al. 2014). Les auteurs voulaient tester les effets d'un entraînement multisensoriel (i.e. visuo-auditif) sur les performances de SA. Un bruit blanc de 2s (intensité comprise entre 62 et 66 dB SPL pour que l'AGC de l'implant ne soit pas activé) était émis par l'un des 7 HP répartis dans l'espace avant du furet (à $\pm 90^\circ$), de manière concomitante à un indice visuel spatialisé (flash lumineux situé au-dessus du HP émetteur). Les animaux réalisaient la tâche de localisation à l'aide de mouvements de tête autorisés pendant l'émission sonore et monitorés en temps réel (via une caméra infra-rouge posée sur leur tête). Dès l'émission du son, les furets étaient conditionnés pour se diriger vers la source sonore perçue et recevaient une récompense lors de la réalisation correcte de la tâche. Quatre furets (2 UCI et 2 BCI) ont suivi cet entraînement. Les animaux BCI ont significativement amélioré leurs performances de SA tandis que les UCI n'ont pas modifié leurs performances. Ces résultats

suggèrent que les furets UCI n'avaient pas suffisamment d'indices acoustiques exploitables pour tirer profit de cet entraînement spatial.

Un groupe contrôle composé de 4 autres furets BCI ne suivaient aucun entraînement multisensoriel ; leurs performances de SA restaient inchangées lors des évaluations.

3.2.2. Patients appareillés avec une prothèse conventionnelle

Même si les difficultés de SA sont majorées chez les patients IC en comparaison de patients appareillés (car les indices acoustiques sont moins altérés par une prothèse auditive conventionnelle), il nous paraissait intéressant de faire le point sur deux études cliniques ayant inclus un nombre conséquent de patients appareillés et soulevant les points essentiels d'un entraînement pour la re-spatialisation auditive.

La première étude a été menée par Kuk et collaborateurs en 2014. Quinze adultes appareillés bilatéralement depuis plusieurs années (1 à 20 ans) ont été divisés en 3 groupes afin de tester l'efficacité d'un protocole d'entraînement de la SA à domicile versus en laboratoire. Les deux premiers groupes suivaient un programme d'entraînement sur 2 mois (schéma de cross-over : 1 mois en laboratoire puis 1 mois au domicile pour le groupe 1, inversement pour le groupe 2) tandis que le troisième groupe (groupe contrôle) ne réalisait que les tests pré et post-entraînement. Lors de l'entraînement, la source sonore émise était de durée variable (0,3 à 3s) et aucun indice visuel concomitant à l'émission du son n'était délivré. Les patients devaient localiser le HP émetteur perçu en le sélectionnant sur une tablette représentant l'ensemble de l'installation de HP (voir tableau IV.2 pour plus de détails sur les modalités d'entraînement dans chaque groupe).

L'entraînement de la SA proposait deux aspects originaux par rapport aux autres études. (1) En fonction du pourcentage d'erreurs réalisées au cours d'un entraînement, le programme adaptait sa difficulté en modulant la nature du stimulus auditif (durée du son, atténuation du bruit de fond) : par exemple, la difficulté était accrue lorsque la durée du son diminuait et l'intensité du bruit de fond augmentait. (2) Les patients recevaient un retour de performance sous la forme d'un feedback visuo-auditif (i.e. émission du son et visualisation concomitante du HP émetteur) d'une durée illimitée, ce qui leur permettait d'explorer librement avec des mouvements de tête les nuances acoustiques non perçues initialement lors de phase de recherche du son.

Plusieurs résultats se dégagent de cette étude. Tout d'abord, les deux groupes ayant suivi l'entraînement de SA ont significativement amélioré leurs performances de SA, sans que le lieu de l'entraînement n'influe sur leurs performances. Cette amélioration se maintenait 1 mois

après la fin de l'entraînement. Aucun changement de performance n'était constaté dans le groupe contrôle qui n'avait pas suivi d'entraînement ; il n'y avait donc pas d'apprentissage spontané des tests réalisés en pré et post-entraînement. Ces données sont en accord avec d'autres études ayant également inclus un groupe contrôle dans leurs travaux (Tyler et al. 2010, Strelnikov et al. 2011, Steadman et al. 2019). De façon plus anecdotique, il semblerait que l'entraînement en laboratoire ait tendance à majorer les performances des patients, certainement du fait d'une plus grande attention portée aux exercices proposés et du meilleur encadrement qu'à domicile (i.e. adhésion et motivation). Cette étude montre donc qu'une prise en charge régulière et adaptée aux capacités du patient permet de faire émerger des effets bénéfiques sur la SA et ces effets peuvent se maintenir dans le temps.

Une deuxième étude a été menée par Firszt et collaborateurs (2015) chez 11 patients atteints de surdité profonde unilatérale non appareillée. Indépendamment des performances individuelles, la difficulté du programme d'entraînement augmentait progressivement au fur et à mesure des 5 sessions d'entraînement par modulation du contenu des informations multisensorielles délivrées. Lors de la première session d'entraînement, le patient muni d'une tablette voyait s'afficher parmi tous les HP matérialisés celui qui allait émettre le son (indice visuel de localisation auditive). Ensuite, un son de 500 ms était délivré et le stimulus visuel s'éteignait. Au fil des séances, le pourcentage d'essais débutant avec cet indice visuel diminuait pour ne plus du tout être présent à la séance 5, laissant le patient sans indice visuel sur la position spatiale de la source sonore. Huit patients sur 11 amélioraient leurs performances de SA ; cette amélioration était d'autant plus importante que le patient avait un déficit de spatialisation marqué lors de l'évaluation pré-entraînement. Ce résultat suggère que les patients expérimentant le plus de difficultés ont une marge de progression plus élevée que des patients ayant initialement un meilleur traitement des indices auditifs spatiaux. On notera que l'utilisation de la tablette comme système de réponse a pu biaiser les performances réelles des participants. En effet, les patients devaient apprendre à translater les coordonnées auditives 3D des HP vers un espace réduit en 2D. Il est possible que les effets obtenus soient liés à un apprentissage dépendant du mode d'entraînement, et ne soient pas généralisables à d'autres espaces auditifs non entraînés.

En résumé, ces deux études ont mis en avant des points essentiels dans la rééducation de patients atteints de surdité. Un programme d'entraînement de difficultés croissantes permet aux patients de trouver les ressources nécessaires (motivation, adhésion, attention, etc) pour réaliser

pleinement les exercices. Néanmoins, l'ajustement du protocole aux capacités individuelles du patient (Kuk et al. 2014) est un facteur confondant dans les études de groupe et doit surtout être mis en avant dans une prise en charge individuelle.

De plus, il paraît essentiel que le patient soit actif dans l'exploration de son environnement visuo-auditif lors d'un entraînement car les retours de ses actions sont autant de signaux d'erreurs qui lui permettent de recalibrer son environnement sensoriel. Mais comme le souligne Firszt, les bénéfices induits peuvent difficilement être extrapolés dans le quotidien car celui-ci est beaucoup plus riche que l'environnement confiné d'un laboratoire, et les patients trouvent dans leur environnement des indices sensoriels leur permettant de minorer leur handicap (e.g. support de la lecture labiale pour la compréhension du langage). Aussi, l'ajout d'une échelle de qualité de vie dans l'évaluation des effets induits par l'entraînement permettrait de mieux cerner les bénéfices directs dans le quotidien des patients.

4. Quels sont les éléments clés pour optimiser une rééducation de la spatialisation auditive à destination de patients implantés ?

L'expérience des différentes équipes de recherche nous a amenés à nous interroger sur les éléments clés devant être intégrés à un protocole d'entraînement spatial à destination de patients implantés cochléaires. Dans ce paragraphe, nous proposons de reprendre point par point ces éléments ; ceux-ci ont été à la base du développement de notre propre protocole d'entraînement RECOVER (voir Etude 3).

4.1. Lieu de l'entraînement

Les protocoles d'entraînement peuvent avoir lieu dans un laboratoire de recherche ou directement au domicile du patient. Le choix du lieu dépend des contraintes matérielles de l'expérience mais doit également s'envisager en fonction du projet de soin du patient.

Le laboratoire est un environnement contrôlé qui présente plusieurs avantages techniques et humains. En effet, il permet de maîtriser l'ensemble des stimuli délivrés (auditifs, visuels, proprioceptifs), ainsi que l'acoustique de la pièce (notamment la réverbération). Une pièce d'expérimentation est souvent dédiée et permet de conserver un espace identique lors des différentes passations, elle garantit donc la cohérence des mesures répétées. De plus, elle permet d'agencer une structure de localisation spatiale complexe (e.g. plusieurs rangées de haut-parleurs à des hauteurs différentes). Enfin, la présence de l'expérimentateur permet l'interaction

Tableau IV.2 : Exemples d'entraînement de la spatialisation auditive chez des sujets appareillés et implantés cochléaires.

ETUDES		Kuk et al. 2014		Firszt et al. 2015	Tyler et al. 2010	Nawaz et al. 2014	
Sujets		N=15 28-83 ans S. bilatérale modérée à sévère appareillée		N=11 23-73 ans S. unilatérale profonde non appareillée	N=9 63-77 ans BCI	N=1 57 ans UCI (surdité unilatérale)	
Entraînement	Groupes		N=3, GC=1		N=1, GC=0	N=2, GC=1	N=1, GC=0
	Variation de quels paramètres		lieu entraînement (cross over), type stimulation (durée, atténuation)		% essais avec indice visuel lors de la phase de recherche	type d'entraînement (spatial / non spatial)	NA
	Généralités	Lieu	Laboratoire	Domicile	Laboratoire	Domicile	Domicile
		Durée/fréquence	6 séances de 2h sur 1 mois	30min/j, 5j/semaine sur 1 mois	5 sessions de 1h tous les 3 à 7 jours	minimum 30 min/j tous jours, pendant 1 à 3 mois	1h/j, 5h/semaine, pendant 3 mois
	Phase de recherche	Uni/Multisensorielle	auditif seul	auditif seul	visuel puis auditif	auditif seul	auditif seul
Type espace		azimut 360° (12 HP)	2 HP (1 devant 0°, 1 derrière 180°)	azimut avant (15 HP)	azimut avant (2 HP)	NA	
Son		parole + bruit	30 sons différents (dont parole)	large spectre	sons du quotidien (sans parole)	phrases dans le bruit	
MVT tête autorisés		non	non	NA	oui	NA	
Feedback		audiovisuel	audiovisuel	visuel (via écran)	NA	NA	
Résultats		Pas de différence de performances entre les lieux d'entraînements : bénéfiques en termes de SA constatés dans les 2 groupes entraînés Intérêt de réaliser un entraînement adaptatif Performances non modifiées dans le GC		Amélioration des performances pour 8 patients sur 11 Les patients ayant les erreurs de SA les plus grandes avant entraînement s'améliorent le plus	Importante variabilité inter-individuelle Performances non modifiées dans le GC	Acquisition de la discrimination gauche-droite	

BCI : patients bi-implantés cochléaires ; GC : groupe contrôle ; HP : haut-parleur ; MVT : mouvements ; N : nombre ; S : surdité ; SA : spatialisation auditive ; UCI : patients mono-implantés cochléaires

directe avec le patient en s'assurant de la bonne compréhension des consignes et la réalisation correcte de la tâche. Ce lieu est surtout dédié à définir les modalités optimales de la SA tout en limitant au maximum les biais expérimentaux.

L'entraînement au domicile est moins contraignant et permet au patient de réaliser les séances au rythme de ses activités journalières. Toutefois, la plupart des séances suivant ce modèle utilisent des protocoles adaptatifs où le patient est libre de choisir son entraînement parmi un panel varié ainsi que la durée des séances, ce qui va à l'encontre de la reproductibilité d'un protocole de recherche. En effet, la reproductibilité des résultats est nécessaire pour la comparaison d'effets entre groupes, ce n'est qu'un fois la prise en charge validée dans un environnement contrôlé (dans un centre ou dans plusieurs centres simultanément) que cette dernière peut être évaluée à domicile. Si la prise en charge à domicile semble plus viable sur la durée pour les patients, cette démarche nécessite un contrôle étroit à distance des entraînements effectués par le patient (voir le système de contrôle utilisé par Kuk et al. 2014).

En résumé, chaque lieu d'entraînement présente ses avantages et inconvénients mais il se dégage une complémentarité qu'il serait intéressant d'exploiter de façon plus systématique à l'avenir. Dans l'idéal, les patients pourraient débiter leurs séances d'entraînement en laboratoire et poursuivre sur le long terme à domicile ou en cabinet d'orthophonie avec un matériel de localisation simplifié afin de maintenir dans le temps les bénéfices apportés par un entraînement en laboratoire.

4.2. Durée et fréquence de l'entraînement

L'organisation temporelle d'une rééducation est un point crucial en termes de faisabilité (i.e. adhésion du patient sur la durée, motivation à réaliser la tâche demandée) et de bénéfices escomptés. Trois principaux points doivent être considérés : la durée totale du protocole d'entraînement (e.g. 1 semaine versus 1 mois) avec le nombre de séances prévues, la fréquence des séances (e.g. 5 séances successives sur 1 semaine versus 1 séance par semaine), et la durée de chaque séance. En effet, il faut trouver un compromis entre le nombre maximal de séances acceptables pour que le patient suive le protocole dans sa totalité et le nombre minimal pour espérer l'émergence d'une amélioration des performances. En effet, une prise en charge de courte durée va entraîner des changements rapides mais labiles (Pascual Leone et al, 2015) donc aux bénéfices limités pour le patient (mais pas forcément dans la cadre d'une recherche fondamentale). Les rares études réalisées chez des patients appareillés (Kuk et al. 2014, Firszt

et al. 2015) et implantées (Tyler et al. 2010, Nawaz et al. 2014) étalaient les séances sur au moins 1 mois.

Concernant la fréquence des séances, les travaux de Kumpik et collaborateurs (2010) ont montré que la répétition de courtes sessions d'entraînement sur plusieurs jours permettait un meilleur apprentissage qu'une seule session plus longue. Mais contrairement aux sujets normo-entendant capables de suivre un rythme soutenu de séances (4 à 5 d'affilée par semaine), les patients déficients auditifs sont exposés à une importante fatigabilité auditive qui impose d'espacer beaucoup plus les séances. Ce modèle rééducatif a longuement été éprouvé en orthophonie libéral où les patients ont au maximum 2 séances par semaine, chaque séance durant en moyenne entre 30 et 45 minutes.

4.3. Nature des stimuli délivrés

De façon consensuelle, l'entraînement spatial se divise en 2 parties : la phase de recherche et de localisation de la source sonore, suivie d'un retour de performance – le feedback. Ce dernier n'est pas forcément présent dans tous les protocoles de la littérature, or sa nature multisensorielle est à la base de la recalibration des cartes spatiales auditives. Pour chaque partie, il est possible de faire varier la nature des stimuli délivrés (i.e. type de son émis, informations uni ou multisensorielles, écoute passive ou active) ; ces choix conditionnent en grande partie les capacités d'amélioration des sujets.

4.3.1. Phase de recherche du son

Nous avons vu dans le chapitre 1 que la nature du son à localiser impactait directement les performances de SA (e.g. un son aigu est plus facile à localiser en élévation qu'un son grave), d'autant plus lorsque les sujets ont un déficit auditif et ne peuvent pas exploiter l'ensemble des indices binauraux et monauraux. Pour une population de patients implantés présentant des marques d'IC et des réglages différents, l'idéal est d'utiliser des sons qui sont perçus par tous (i.e. éviter des sons à bandes de fréquences étroites). Deux choix peuvent alors se discuter : privilégier un son à larges bandes type bruit blanc / bruit rose ou bien un son plus saillant du quotidien comme la parole (Firszt et al. 2015). On notera que beaucoup d'études en audiologie privilégient des stimulations sonores adaptées aux capacités auditives de chaque participant (afin de limiter des biais entre sujets) ; cependant cette approche n'est pas pertinente dans la cadre d'études sur la SA.

Quelle que soit la nature du son choisi, il faut également être attentif à sa durée. En effet, si le choix se porte vers un entraînement spatial dynamique (i.e. avec une exploration libre de la scène auditive avec la tête), le son délivré doit être suffisamment long (supérieur à 500 ms) pour que les indices acoustiques générées par les mouvements de tête soient exploitables par le sujet (voir chapitre 1, paragraphe 5.2).

L'ajout d'une scène visuelle (réelle ou virtuelle), de manière concomitante à l'émission du son, permet de promouvoir les interactions multisensorielles qui sont naturellement facilitatrices dans le développement de la spatialisation. L'équipe de Strelnikov (2011) a ainsi démontré qu'un entraînement visuo-auditif (i.e. stimulation auditive de 50 ms associée à un flash lumineux spatialisé à l'endroit du HP) permettait une plus grande amélioration des performances qu'un entraînement auditif seul. Cet indice visuel donnait un indice spatial permettant de déplacer l'attention exogène du sujet vers la source sonore, et ainsi réaliser un traitement accru des indices auditifs spatiaux.

L'apport de la RV a été bénéfique pour manipuler et contrôler de façon plus rigoureuse la teneur des informations visuelles délivrées. Il est ainsi possible d'apporter un ancrage visuel minimaliste au sujet (e.g. Zahorik et al. 2006, Valzolgher et al. 2020b) ou bien une scène virtuelle plus riche reprenant les volumes et objets de la pièce (e.g. Ahrens et al. 2019, Steadman et al. 2019). Ce choix dépend notamment du niveau de difficulté que l'on souhaite appliquer au cours de l'entraînement. Plus les caractéristiques spatiales de la scène visuelle est en adéquation avec la scène auditive et plus le sujet aura des facilités à localiser l'origine du son (Ahrens et al. 2019). Nous savons que les patients IC conservent des capacités d'intégration visuo-auditives supérieures à celles des sujets NE (notamment en privilégiant la lecture labiale dans des environnements bruyants, Rouger et al. 2007). Il semblerait donc plus opportun d'intégrer un environnement réaliste pour débiter un protocole d'entraînement avec des patients (et ensuite le faire évoluer en fonction des performances dans le cadre d'une prise en charge individuelle). Enfin, la RV offre l'opportunité d'immerger le patient dans un environnement interactif, voir ludique, de contrôler la complexité des stimulations délivrées, rendant le patient actif de sa rééducation.

Un autre facteur prédictif a été manipulé lors des phases d'entraînement : l'exploration active et spontanée de la scène auditive par des mouvements de tête. Steadman et collaborateurs (2019) ont montré que le groupe de NE exposés à un entraînement visuo-auditif passif (i.e. exploration céphalique non permise pendant l'émission du son) avait de moins bonnes

performances de SA que le groupe suivant un entraînement visuo-auditif actif (i.e. mouvements de tête autorisés). Selon un protocole similaire, ces résultats ont été confirmés par l'équipe de Bermejo (2020), où ils ont également mis en évidence une modification du comportement d'orientation céphalique en fin de rééducation. En effet, les sujets ont diminué leur temps d'exploration et allaient plus directement vers la source sonore. Ces résultats suggèrent que la génération d'indices acoustiques dynamiques par des mouvements de tête libre (et non contraint comme c'est le cas dans certaines études, e.g. Kumpik et al. 2010, Irving et al. 2011) permet de palier la distorsion des indices acoustiques induite dans ces deux études par un système d'écouteurs. Les patients IC subissant naturellement cette distorsion, il semblerait intéressant d'inclure de manière systématique l'écoute active dans les protocoles d'entraînement, d'autant plus que certaines études ont montré le bénéfice des mouvements de tête dans cette population lors de tâches ponctuelles de SA (Mueller et al. 2013, Pastore et al. 2018, Fischer et al. 2020).

A la fin de l'émission du son, le sujet doit indiquer la position du son perçu. Sa réponse peut être réalisée de deux manières : (1) directement en pointant la zone de l'espace avec la main (réponse 3D) ou la tête (réponse 2D, c'est-à-dire en direction de la source sonore), ou bien (2) indirectement en nommant à haute voix le HP source ou en sélectionnant ce HP matérialisé sur un écran (e.g. tablette). Au cours de l'entraînement, l'interaction motrice directe avec la source sonore conduit à une amélioration plus importante des performances que la réponse indirecte (Valzolgher et al. 2020a). Ce constat renforce l'idée que la richesse des interactions sensori-motrices réalisées dans un espace commun promeut le développement d'une nouvelle carte spatiale.

4.3.2. Feedback

Le terme « feedback » est issue de la littérature anglophone et signifie littéralement « retour d'information ». Ce concept est connu depuis longtemps pour promouvoir l'apprentissage d'une tâche en donnant la possibilité au sujet de constater ses erreurs puis de se corriger (Bauer 1966). Lors d'un entraînement spatial, la présence du feedback n'est pas obligatoire mais plusieurs auteurs ont montré qu'il permettait aux sujets d'améliorer significativement leurs performances de SA comparé à un groupe contrôle ne recevant aucun feedback (Zahorik et al. 2006, Strelnikov et al. 2011).

Le feedback peut varier selon sa nature sensorielle (e.g. retour visuel, auditif ou proprioceptif) et la richesse des informations qu'il véhicule (d'un simple « réponse vraie ou fausse » à un

retour précis de la position exacte du son, c'est-à-dire avec une correspondance spatiale entre la réponse et le feedback). A l'heure actuelle, aucune étude n'a combiné ces deux variables et comparé des feedbacks de nature différente. Il est donc difficile de tirer des conclusions claires sur le feedback le plus efficace pour promouvoir la SA chez des patients déficients auditifs. Nous pouvons toutefois supposer qu'un feedback visuel binaire non spatialisé (e.g. Irving and Moore 2011, Strelnikov et al. 2011) versus un feedback visuel spatialisé (e.g. Kumpik et al. 2010, Steadman et al. 2019) limite un peu plus l'apprentissage. En effet, dans le premier cas, seul un voyant lumineux situé droit devant le sujet lui indique s'il a pointé dans la bonne direction à 5 degrés près (i.e. en rouge si l'erreur est supérieure à 5°, en vert si elle est inférieure). Le sujet n'a donc aucun retour sur la position exacte du HP qui a émis ni sur la magnitude de son erreur si le voyant est rouge. De plus, avec un feedback non spatialisé le sujet reçoit un indice dans un référentiel non spatial qui est différent de celui qui lui a permis de localiser le son et de générer sa réponse (il doit donc apprendre une correspondance entre un indice non spatial et un comportement spatialisé).

D'autres études ont utilisé des feedback plus complexes, sur le modèle de l'interaction multisensorielle de la phase de recherche (Zahorik et al. 2006, Majdak et al. 2013, Valzolgher et al. 2020a). Ainsi, une fois que le sujet avait donné sa réponse, il était possible de lui indiquer visuellement la position exacte du HP tout en continuant d'émettre le son. En complément, il était demandé au sujet de se réorienter en direction de la source afin d'associer et d'encoder correctement les deux informations sensorielles simultanément. Nous faisons l'hypothèse que ce mode de feedback est le plus efficace car il combine dans un même référentiel spatial une stimulation sonore, une réponse motrice et un retour de performance multisensorielle qui s'enrichissent à travers l'interaction du sujet avec son environnement.

Enfin, nous avons constaté que certaines équipes avaient rajouté des éléments de jeux vidéo pour rendre les séances plus ludiques. Cette variante n'a pas montré de bénéfices pour des sujets NE (Steadman et al. 2019) mais pour des patients appareillés : Kuk et collaborateurs ont montré que la visualisation des performances sous forme de code couleur (sans détailler le pourcentage exact d'erreurs pour ne pas décourager les patients) leur permettait de mieux focaliser leur attention sur l'espace le plus déficitaire.

4.4. Type d'espace rééduqué

Dans la plupart des études, seule une dimension spatiale est entraînée, principalement l'azimut (voir tableaux IV.1 et IV.2). Citons toutefois quelques contre-exemples dans une

population de NE où l'élévation a également été travaillée (Hofman et al. 1998, van Wanrooij & van Opstal 2005, Zahorik et al. 2006, Steadman et al. 2019). La restriction des conditions d'entraînement à un seul espace s'explique en grande partie par la complexité des dispositifs de localisation à installer, ceux-ci prenant beaucoup de place et pouvant générer des problèmes d'écho (e.g. si plusieurs sources sonores se trouvent l'une derrière l'autre pour tester différentes profondeurs à une même position azimutale).

La combinaison simultanée de deux espaces soulève également un problème d'apprentissage car cela complexifie de façon non négligeable la tâche de SA, le sujet devant intégrer dans le même temps des indices binauraux et monauraux, potentiellement altérés. Dans le cadre d'un déficit sensoriel, l'approche progressive semble donc plus adaptée. Initialement, l'isolement de chaque espace permettrait au patient de travailler isolément chaque indice auditif (i.e. les indices binauraux pour l'azimut et les indices monauraux pour l'élévation). Puis, lorsque ces pré-requis sont validés, le mélange des espaces (changement simultané de la position en azimut et en élévation d'un haut-parleur) permettrait au patient d'expérimenter une situation plus proche de la réalité de son environnement.

5. Points clés du chapitre

- a) La rééducation repose sur les phénomènes de plasticité cérébrale et d'intégrations multisensorielles afin que les représentations sensorielles non déficitaires modulent celle qui fait défaut.

- b) Il n'existe pas de protocole standardisé pour réduire les déficits de spatialisation auditive (SA) des patients implantés cochléaires.

- c) Les rares protocoles d'entraînement spatial développés en recherche pour des patients appareillés ou implantés ont montré des effets bénéfiques sur les performances de SA.

- d) Plusieurs critères permettent d'optimiser les effets de l'entraînement de la SA chez les patients IC :
 - Entraînement prolongé dans le temps (au moins 1 mois) à un rythme moyennement soutenu (maximum 2 séances par semaine)
 - Combiner dans un même référentiel spatial la stimulation sonore, la réponse motrice et le feedback multisensoriel tout en incitant le sujet à interagir activement avec son environnement (e.g. par l'exploration céphalique)
 - Entraînements de difficultés croissantes (e.g. réduire initialement chaque dimension spatiale séparément, diminuer progressivement la durée du son et la richesse des informations visuelles, etc)

TRAVAUX PERSONNELS DE RECHERCHE

- ❑ Etude 1 : SPHERIC - Etude de la localisation en 3D chez des enfants bi-implantés cochléaires

- ❑ Etude 2 : KID SSQ12 – Développement et validation d’une forme courte du questionnaire pédiatrique SSQ

- ❑ Etude 3 : RECOVER - Rééducation de l’espace auditif 3D chez des adultes bi-implantés cochléaires

Etude 1

SPHERIC

Etude de la spatialisation auditive chez des enfants bi-implantés cochléaires

**Publiée dans Ear and Hearing
2021 Jul 1. doi: 10.1097/AUD.0000000000001090**

Partie 1 : Résumé de l'étude

1. Questions soulevées au regard du contexte scientifique

Ces dernières années ont été marquées par l'essor de la bi-implantation cochléaire dans le monde pédiatrique afin de restaurer au mieux la binauralité chez des enfants souffrant de surdité profonde bilatérale. L'enjeu premier de la bi-implantation est d'optimiser les compétences de communication orale de l'enfant afin de promouvoir son développement global et son insertion socio-éducative (Truy et al. 2017, Simon et al. 2019). L'importance de la bi-implantation dans la localisation des sons est un sujet émergent et les études dans le domaine restent limitées à quelques équipes sensibilisées par le sujet (e.g. Grieco-Calub & Litovsky 2010, Gordon 2014 et al. ; Litovsky & Gordon 2016). De plus, l'évaluation pédiatrique de la spatialisation auditive (SA) est un défi à part entière car les tests qui sont essentiellement destinés à évaluer la spatialisation chez l'adulte doivent être adaptés aux compétences langagières et motrices de l'enfant.

Plusieurs questions restent actuellement en suspens :

- Est-il possible de quantifier aisément les capacités de SA dans les 3 dimensions de l'espace chez l'enfant ? En effet, il n'existe pas de tests de localisation spatiale dans le domaine pédiatrique qui répondent aux exigences expérimentales de la recherche (e.g. contrôle des entrées visuelles, possibilité d'échantillonner tout l'espace auditif...)
- Les enfants normo-entendant (NE) et bi-implantés cochléaires (BCI) sont-ils en mesure d'exploiter les indices acoustiques dynamiques induits par les mouvements de tête ? Cette stratégie céphalique leur permet-elle d'améliorer leurs performances de localisation ?
- Les performances de SA en profondeur n'ayant jamais été évaluées en pédiatrie, les enfants NE et BCI peuvent-ils discriminer des distances proches (i.e. à moins d'un mètre de la tête) ?
- La maturation de la SA est-elle comparable pour tous les enfants NE passé l'âge de 8 ans ?
- Existe-t-il un lien entre les résultats objectifs des tests de localisation spatiale et le ressenti des enfants sur leurs difficultés au quotidien ?

Nous proposons dans cette première étude d'adapter notre système de SA en réalité virtuelle (RV) pour quantifier en 3D les performances d'enfants NE et BCI.

2. Justification du design de l'étude

Trente-six enfants (18 dans le groupe NE et 18 dans le groupe BCI) ont été inclus dans l'étude SPHERIC. Un âge minimal de 8 ans a été requis afin de limiter les problèmes de compréhension au test, les troubles attentionnels et un important retard de maturité auditive (Litovsky 2015). La limite supérieure était de 17 ans, ce qui permettait de suivre la poursuite d'une éventuelle maturation auditive à l'adolescence.

Chaque enfant réalisait un test de localisation spatiale sur la plateforme immersive de Neuro-immersion d'une durée maximale de 30 minutes pour éviter l'effet d'épuisement. Huit positions différentes de haut-parleurs (HP) étaient évaluées afin d'évaluer l'ensemble de l'espace auditif de l'enfant. Nous faisons varier les positions azimutales sur 360° ainsi que la distance des HP dans l'espace proche, à 35 et 55 cm.

Tout au long de l'expérience, les enfants étaient équipés d'un casque de RV où ils n'avaient aucun support visuel pouvant les orienter sur la position des HP. Chaque essai débutait par l'émission d'un bruit blanc de 3 secondes. Ce son était à large spectre et suffisamment long pour être interprétable par l'ensemble des participants et permettre aux enfants d'échantillonner leur espace à l'aide de mouvements de tête. A la fin de l'émission sonore, les enfants devaient localiser directement avec une baguette la position du son perçue ce qui permettait d'enregistrer une réponse motrice en 3D. Ils ne recevaient aucun feedback sur leurs performances.

Nous souhaitons évaluer l'impact des mouvements de tête sur les performances de SA des enfants. Aussi, pour la moitié des essais, les enfants étaient libres d'explorer avec leur tête la scène auditive pendant les 3 secondes d'émission du son (i.e. condition « tête libre »), pour l'autre moitié les mouvements de tête n'étaient pas autorisés avant la phase de réponse (condition « tête immobile »).

Dans le groupe des enfants implantés, parents et enfants complétaient séparément le questionnaire SSQ pédiatrique adapté par Galvin and Noble (2013). Cela permettait d'apprécier spécifiquement les difficultés de SA grâce aux items de la section B du questionnaire.

Pour plus de détails sur la méthodologie, voir l'article soumis en partie 3.

3. Principaux Résultats

Le système breveté SPHERE (Annexe B) qui utilise la technologie de la RV nous a permis de quantifier de manière aisée les performances de SA de deux populations pédiatriques, des NE et des BCI. En effet, les enfants qui ont été inclus dans l'étude ont tous réalisé le test de localisation en respectant les consignes (ne pas bouger la tête pendant l'émission sonore pour certains essais, et attendre la fin du son pour pointer la source). Moins de 5 % des essais ont été retirés des analyses pour non-respect de ces consignes.

Les difficultés de localisation des enfants BCI ont pu être quantifiées en 3D, leur erreur de pointage 3D était deux fois plus importante que celle des NE (respectivement 55 et 24 cm). Chez les BCI, l'erreur 3D se réduisait significativement de 8,6 cm en condition « tête libre » tandis que pour les NE, les performances n'étaient pas modifiées (baisse non significative de l'erreur 3D de 0,5 cm). L'amélioration des performances des BCI était directement corrélée avec le pourcentage d'essais où les enfants bougeaient la tête pendant le son. Les bénéfices de la localisation spatiale dynamique étaient également visibles à l'échelle des confusions avant-arrière chez les BCI avec un passage de 44,6% à 25,1%. Ces confusions restaient anecdotiques chez les NE (passage de 2,1 à 1,4%).

La discrimination des deux distances de l'espace proche dans l'étude (35 et 55 cm) était possible pour les NE mais très difficile pour les BCI. En moyenne, les NE percevaient les deux distances avec un écart de 11 cm contre 3 cm pour les BCI. La condition « tête libre » n'améliorait pas les performances en profondeur quel que soit le groupe.

Une variabilité inter-individuelle a émergé dans les deux groupes d'enfants, particulièrement sur la perception de la distance et les capacités d'amélioration grâce aux mouvements de tête. Pour les enfants BCI, nous avons retrouvé une corrélation positive entre leurs erreurs de localisation et le délai inter-implant : plus le délai était important et plus l'erreur 3D était élevée. Pour les enfants NE, les erreurs 3D étaient corrélées négativement avec l'âge au moment du test : plus l'enfant était âgé, plus l'erreur 3D était faible. Enfin, nous n'avons pas retrouvé de corrélation entre l'erreur 3D et le score moyen à la section B du SSQ (comportant les items de localisation spatiale).

Partie 2 : Discussion générale de l'étude SPHERIC

1. Des performances de localisation limitées en condition « tête immobile »

Dans cette étude, nous avons évalué et comparé les performances de SA dans les 3 dimensions de l'espace d'enfants BCI par rapport à celles d'enfants NE. Notre système nous a permis de nous affranchir de plusieurs limites rencontrées avec les précédentes études pédiatriques. Tout d'abord, notre test comportait 96 essais (contre 150 en moyenne, Zheng et al. 2015 ; Choi et al. 2017) ce qui limitait le risque de biais attentionnel survenant lors de tests longs. De plus, nous avons supprimé tout indice visuel qui aurait pu biaiser et faciliter la perception spatiale, comme cela était le cas avec l'utilisation de tablette permettant à l'enfant de visualiser la position des HP (Grieco-Calub & Litovsky 2010 ; Zheng et al. 2015 ; Killan et al. 2018) ou l'envoi d'un feedback visuel qui permettait à l'enfant d'adapter ses réponses (Zheng et al. 2015).

Dans la condition où les mouvements de tête n'étaient pas autorisés pendant l'émission du son, nous avons constaté que tous les enfants implantés avaient acquis la capacité à discriminer les sons venant des héli-espaces gauche et droit (Figure V.1). Ce résultat est en accord avec les constatations de Zheng et collaborateurs (2015) et confirme que la latéralisation d'un son dans l'espace est un processus auditif rudimentaire qui ne nécessite pas une grande finesse dans le traitement des indices binauraux.

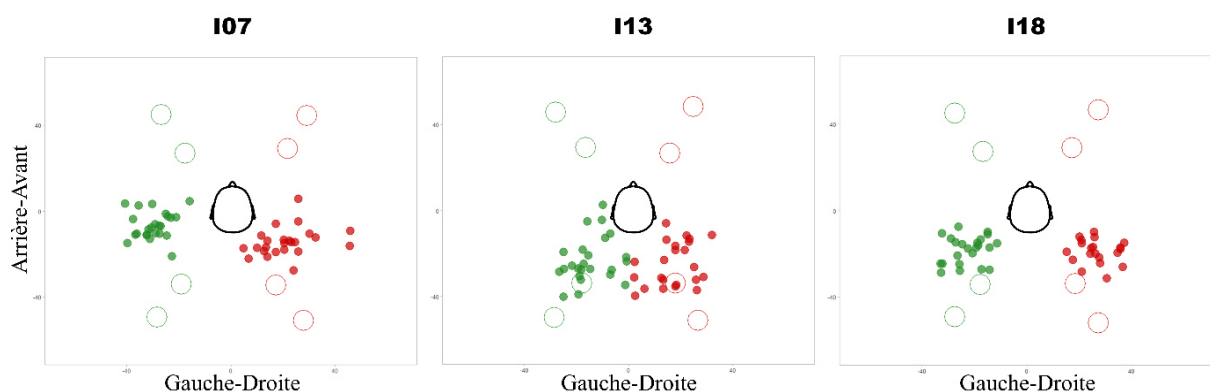


Figure V.1 : Discrimination gauche-droite en condition statique de 3 enfants bi-implantés cochléaires (I07, I13, I18).

Chaque point coloré correspond à un pointage (vert à gauche, rouge à droite). Les haut-parleurs sont matérialisés par un cercle.

A l'inverse, nous avons constaté que la discrimination avant-arrière était un véritable challenge car les performances des BCI étaient proches de la chance. De plus, une grande majorité des enfants BCI avaient des réponses très stéréotypées en pointant dans un espace très restreint autour de leurs implants et plus particulièrement proche du microphone situé sur la partie arrière du processeur (Figure V.1). Les difficultés rencontrées étaient certainement majorées par la restriction des informations sensorielles disponibles dans cette tâche : absence d'indice visuel et de feedback, accès très restreint aux ITDs (Salloum et al. 2010 ; Gordon et al. 2014 ; Ehlers et al. 2017).

Nous avons fait le choix d'utiliser comme critère de localisation le dernier pointage de la main des participants. Dans une étude récente, Zheng et collaborateurs (2015) insistaient sur l'intérêt d'étudier l'orientation du premier mouvement de tête (élément facilement analysable avec notre système de mesure). En effet, les auteurs ont constaté qu'un certain nombre d'enfants s'orientaient spontanément du bon côté de la source sonore puis finalement pointaient à l'opposé dans la mauvaise direction. Dans certaines conditions, il semblerait qu'un conflit survienne entre les réponses réflexes guidées par les structures sous-corticales (i.e. orientation de la tête) et la prise de décision modulée par les structures corticales sous l'influence de plusieurs facteurs (e.g. niveau attentionnel, mémoire, stratégie). Il serait intéressant dans une future étude de voir si les erreurs que nous avons enregistrées en condition « tête immobile » se maintiennent en condition « tête libre » en analysant le premier mouvement d'orientation de la tête.

2. Comparaison des capacités de spatialisation auditive de l'enfant et de l'adulte bi-implantés cochléaires

Il est intéressant de constater que les données acquises chez les enfants BCI ont certains points communs mais aussi des différences avec celles recueillies lors de la même expérience chez des adultes BCI (Figure V.2, Gaveau et al., *article en cours d'écriture*). Dans cette étude en préparation, Gaveau et collaborateurs ont quantifié la SA d'adultes NE et BCI à l'aide du même dispositif que SPHERIC et dans les mêmes conditions méthodologiques : pas de retour visuel pendant le test, localisation dans l'espace proche à l'aide de la main, test réalisé avec et sans mouvement de tête, son durant 3 secondes. La seule différence notable était que les cibles sonores étaient réparties selon 3 profondeurs différentes (i.e. 35, 55, 75cm) alors que dans l'étude SPHERIC, seules 2 profondeurs étaient testées. Les adultes BCI avaient autant de

difficultés à discriminer les sons venant de l'avant et de l'arrière (i.e. erreur moyenne de 46%) mais leur erreur-3D était bien plus grande (69 cm versus 55 cm chez l'enfant), de même que les confusions gauche-droite (18% vs. 8%). Lorsque le test le permettait, les enfants et les adultes exploraient activement l'espace avec leur tête lors de l'émission du son. Les bénéfices de cette stratégie céphalique étaient toutefois plus limités chez les adultes BCI puisque leurs confusions avant-arrière ne diminuaient que de 7% (vs. 19,5% chez les enfants) et l'erreur 3D de 3,5 cm (vs. 8,6 cm). Ce bénéfice restreint peut s'expliquer par le fait que la majorité des adultes BCI bougeaient moins fréquemment la tête en condition « tête libre » (11/17 sujets bougeaient dans moins de 50% des essais vs. 5/18 enfants BCI). Nous pouvons émettre l'hypothèse que le comportement naturel d'orientation de la tête face à un son est mieux conservé et/ou plus facilement rééduqué chez un enfant dont la plasticité cérébrale est plus importante (Fandakova et al. 2020).

Plusieurs études antérieures suggéraient que les capacités de discrimination des ILDs et des ITDs des enfants BCI étaient beaucoup plus altérées que celles des adultes IC (Litovsky et al. 2010, Gordon et al. 2014), pouvant ainsi se répercuter sur leurs performances de localisation. Nos résultats vont à l'encontre de cette théorie. Nous faisons l'hypothèse qu'en étant implantés précocément, les enfants conservent mieux le réflexe naturel d'orientation de la tête et gardent de meilleures capacités d'extraction et d'interprétation des indices acoustiques dynamiques.

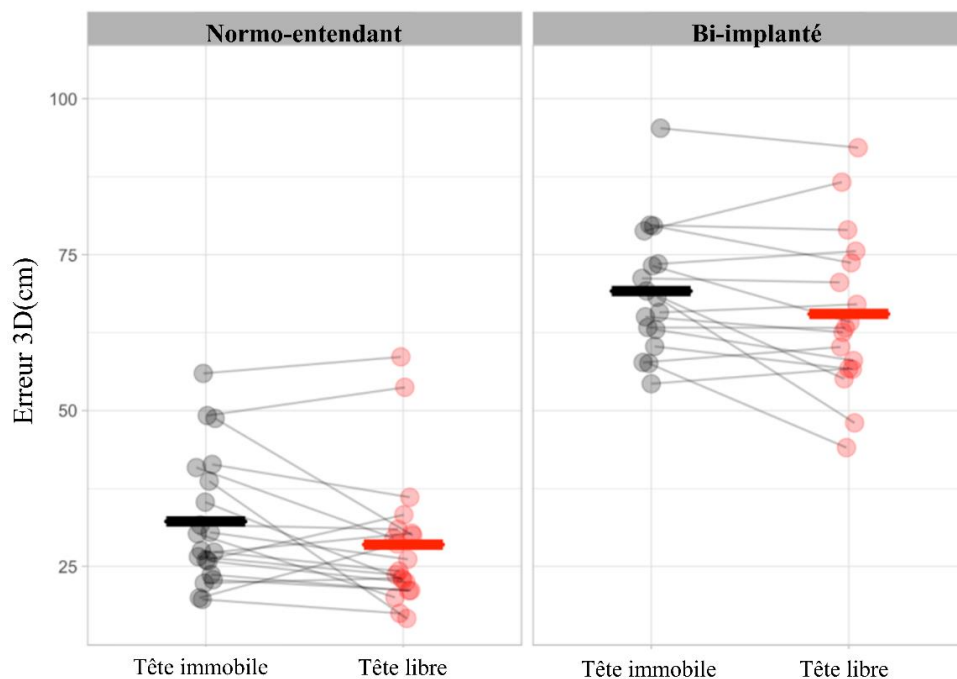


Figure V.2 : Erreur 3D en fonction des conditions de test (« tête immobile vs. « tête libre ») chez des adultes normo-entendant et bi-implantés cochléaires (Gaveau et al. article en cours d'écriture). La barre épaisse représente la moyenne par groupe, chaque point correspond aux données individuelles.

Malgré les bénéfices constatés en condition « tête libre », nous faisons l'hypothèse que les performances enregistrées ne sont pas encore le reflet du quotidien des enfants. En effet, dans la réalité, les interactions multisensorielles prédominent, et les enfants peuvent s'appuyer sur les informations visuelles pour lever toute ambiguïté lors de la localisation d'un son. Ces informations visuelles sont d'ailleurs reconnues pour guider et dominer la perception spatiale au cours du développement de l'enfant (King 2009, Gori et al. 2012). De plus, la présence de voix familières dans un environnement connu facilite la mise en place de cartes spatiales chez l'enfant (Gori et al. 2012). Notre hypothèse de départ est renforcée par les résultats du questionnaire SSQ de notre étude. Ainsi, hormis quelques cas individuels, aucun parent ni enfant n'a manifesté de réelles difficultés de localisation dans des situations fréquentes du quotidien. Il serait intéressant à l'avenir d'évaluer le poids des informations visuelles dans la SA des enfants NE mais aussi BCI en manipulant les indices visuels délivrés dans le casque de RV, de la visualisation précise des volumes de la pièce à un feedback sur le positionnement exact des HP. Nous faisons l'hypothèse que les performances des enfants BCI seront meilleures que celles mesurées dans notre étude. Nous présumons également que la variabilité de performances entre les sujets sera conservée car physiologiquement, l'intégration multisensorielle n'atteindrait sa maturité que tardivement, entre 8 et 12 ans, du fait de la maturation décalée de chaque système sensoriel (Lewkowicz 2002, Gori et al. 2008, Gori et al. 2012). Les adolescents devraient donc être plus performants dans cette tâche multisensorielle.

Enfin, toute la difficulté de l'évaluation pédiatrique repose sur l'utilisation d'un outil adapté à la localisation 3D dans un environnement contrôlé qui puisse reproduire le plus fidèlement possible le quotidien des enfants. Les évaluations actuellement menées en recherche permettent d'identifier les déficits de SA sans pour autant préjuger des adaptations mises en place par l'enfant BCI dans son environnement qui par essence est multisensoriel.

3. Impact de la bilatéralisation séquentielle retardée

Ces quinze dernières années ont été marquées par l'élargissement des indications de bilatéralisation chez l'enfant. La meilleure compréhension des conséquences physiologiques de la privation auditive a poussé le corps médical à éditer des recommandations récentes prônant un délai inter-implant le plus court possible, dans l'idéal inférieur à 6 mois (SFORL 2018, Simon et al. 2019). Les enfants évalués dans notre étude font partie des premières générations

d'enfants ayant bénéficié d'une bilatéralisation séquentielle retardée où les délais inter-implant étaient encore importants.

Nous avons ainsi mis en évidence une corrélation positive entre le délai inter-implant et l'erreur de localisation 3D. Plus ce délai était élevé et plus l'erreur 3D était importante. Passé un délai de 36 mois, tous les enfants de notre étude avaient une erreur 3D supérieure à 50 cm. Ce cut-off concorde avec les données de la littérature. Plusieurs auteurs ont ainsi montré qu'au-delà de 18 mois entre deux implantations, il persistait une asymétrie de traitement des indices acoustiques au niveau sous-cortical, même après 9 mois d'expérience binaurale (Gordon et al. 2011, Gordon et al. 2013). Au cours du neurodéveloppement de l'enfant, cela conduit à terme à une préférence auditive vers la première oreille implantée (Gordon et al. 2008, Gordon et al. 2013, Illg 2013, Gordon et al. 2015, Kral et al. 2019).

Il n'existe pas à notre connaissance dans la littérature pédiatrique d'études ayant directement mis en évidence cette préférence aurale d'un point de vue comportemental. Néanmoins, Gordon et collaborateurs (2014) ont évalué les capacités de perception des ILDs et des ITDs chez 79 enfants BCI aux profils cliniques différents (variation des durées d'expérience binaurale et délai inter-implant). Lors de la tâche, les enfants devaient indiquer si le stimulus était perçu unilatéralement à gauche ou à droite, ou s'ils l'entendaient bilatéralement. Quel que soit le profil clinique, tous les enfants percevaient les ILDs ; à l'opposé, la sensibilité aux ITDs était fortement altérée et les enfants ayant eu une expérience monaurale prolongée (supérieur à 18 mois) percevait plus facilement le son du côté du premier implant.

Nous avons repris individuellement nos résultats pour vérifier l'hypothèse d'un pointage préférentiel vers le premier IC chez les enfants ayant un délai inter-implant supérieur à 36 mois. Aucune asymétrie nette des réponses n'a émergé (Figure V.3). Il est possible que l'asymétrie auditive ait été compensée pour partie par l'expérience auditive (Easwar et al. 2018). En effet, tous les enfants de notre étude avaient un minimum d'expérience auditive de 24 mois.

Finalement, l'asymétrie auditive induite précocément au cours du développement par le port prolongé d'un seul implant ne s'oppose pas à la bonne intégration des indices binauraux au sein du tronc cérébral (Gordon et al. 2012). Par contre, elle crée une rupture dans leur traitement cortical conduisant notamment à une baisse du seuil de détection des ITDs et un remaniement des cartes spatiales (Litovsky et al. 2010, Salloum et al. 2010) partiellement compensé par une expérience binaurale prolongée.

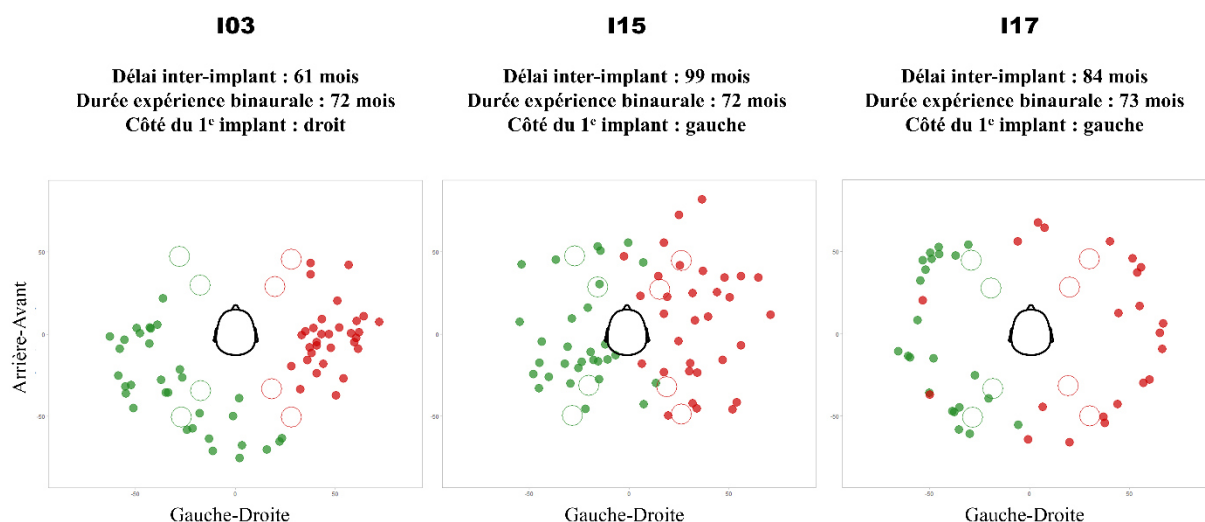


Figure V.3 : Discrimination gauche-droite en condition « tête immobile » de 3 enfants bi-implantés cochléaires (I03, I15, I17) selon différents délais inter-implant et durées d'expérience binaurale. Chaque point coloré correspond à un pointage (vert à gauche, rouge à droite). Les haut-parleurs sont matérialisés par un cercle.

Dans un prochain travail, il serait intéressant d'explorer les phénomènes de plasticité cérébrale dans une étude longitudinale en évaluant de manière séquentielle et prolongée dans le temps les performances de SA d'enfants mais aussi d'adultes BCI avec des délais inter-implant plus ou moins longs. Nous faisons l'hypothèse qu'à délai inter-implant identique, les enfants BCI auront une asymétrie de spatialisation auditive moins marquée que les adultes car celle-ci sera mieux contrebalancée par la plasticité cérébrale.

4. Poursuite de la maturation auditive jusqu'à l'adolescence

Nous avons constaté une certaine variabilité dans les performances des enfants NE avec l'émergence d'une corrélation entre l'âge au moment du test et l'erreur 3D. Les études princeps de SA pédiatrique suggéraient que la maturation des voies auditives était terminée à l'âge de 5 ans. Cela se vérifie pour des tâches simples de discrimination spatiale (e.g. héli-espace gauche vs. héli-espace droit). Mais il semblerait que la discrimination de deux distances proches nécessite une maturation plus importante qui se poursuit après l'âge de 8 ans. En effet, deux des enfants testés âgés de 8 ans et demi percevaient une différence de 5 cm tandis que deux enfants de 14 ans percevaient une différence de 15 cm. Il existerait donc une maturation différentielle entre les trois plans de l'espace et peut-être même entre l'espace proche et l'espace

lointain (i.e. au-delà de 1 mètre du sujet) puisque les indices acoustiques disponibles dans chaque espace diffèrent. En effet, dans l'espace proche, le poids des ILDs est majeur pour évaluer la distance de la source, tandis que dans l'espace lointain il est nécessaire d'interpréter les différentes ondes réfléchies dans un milieu réverbérant. A notre connaissance, il n'existe pas d'études ayant montré l'évolution du développement de la SA en profondeur chez l'enfant. Il serait intéressant de tester un plus large panel d'enfants NE par tranches d'âge différentes et d'évaluer la temporalité d'acquisition de la localisation spatiale absolue dans chaque plan de l'espace (e.g. variations de l'azimut et de l'élévation dans les espaces proche et lointain car les indices auditifs diffèrent grandement en fonction de ces paramètres). Cela nécessiterait de permuter notre système de localisation sonore (bague) par la manette du système VIVE qui peut contrôler en RV un objet virtuel en 3D ; ainsi, cela permettrait d'indiquer précisément la profondeur perçue sans contrainte physique.

Dans notre expérience, les enfants étaient exposés à un environnement calme et les seules informations sensorielles perçues étaient auditives. Il est probable que si nous refaisions le même test de localisation avec un bruit de fond, type cocktail party, la variabilité inter-individuelle constatée serait plus importante car la localisation précise d'une source dans un environnement auditif complexe nécessite la mise en jeu de processus corticaux qui continuent de mûrir jusqu'à l'adolescence (Freigang et al. 2015, Litovsky 2015).

5. Perspectives

Actuellement, la course à l'innovation pousse plus que jamais les firmes d'implants cochléaires à développer de nouveaux systèmes offrant un confort d'écoute aux patients. Ces nouvelles technologies n'ont pas pour but premier d'améliorer les capacités de SA mais au vue de l'élargissement des indications d'implantation pédiatrique et de la période critique de maturation spatiale, il semblerait important d'identifier les paramètres optimisant la spatialisation. Tout d'abord, le positionnement du microphone sur le processeur (derrière, déporté sur l'antenne, ou plus rarement dans le conduit auditif externe) et sa directionnalité (omnidirectionnel vs. directionnel) sont deux paramètres identifiés dans la littérature adulte comme modifiant significativement les performances de SA (Jones et al. 2016, Dorman 2018, Denk et al. 2019, Fischer et al. 2020a). Nous n'avons retrouvé aucune étude comparable en pédiatrie. De plus, selon leurs stratégies de codage, certaines marques d'implants (i.e. Med-EL

et Advanced Bionics) restituent plus fidèlement la structure fine des sons qui est nécessaire à la détection des ITDs (Eklöf & Tideholm 2018). Il serait intéressant de comparer, à âge identique d'implantation, deux groupes d'enfants porteurs d'implants aux stratégies de codage différentes et voir si sur le long terme, l'IC préservant le mieux la structure fine des sons apporte un réel avantage pour la maturation cérébrale de l'analyse des ITDs.

Enfin, vu les déficits de SA constatés dans cette étude, il nous semblerait intéressant de proposer un programme d'entraînement spatial ludique à destination d'enfants BCI dans un premier temps. En effet, les recommandations actuelles prônant un délai inter-implant le plus court possible, il ne nous paraît pas pertinent pour l'instant de débiter une rééducation avec un seul implant.

Shamugha Priya	06/25/21	4 Color Fig(s): F2-5	14:30	ARTEANDH-D-20-00153
----------------	----------	----------------------	-------	---------------------

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 XXX;XX;00-00)

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 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.,

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).

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 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

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-ringing	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.

(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 ± SD: 12.1 ± 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 ± 32.4 months). BCI users were all bilateral CI daily users and had excellent monosyllabic word recognition performance (mean with left CI: 96.8 ± 3.3%, and with right CI: 95.6 ± 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 ± SD: 12.3 ± 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 × 3 m, reverberation time RT₆₀: 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 × 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 (SensoriMotoric Instruments, Berlin, Germany; www.

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	Crystalis XDP	Omnidirectional	91	Digisonic SP EVO, SAPHYR SP, R	Crystalis 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	Crystalis XDP	Omnidirectional	106	Digisonic SP EVO, SAPHYR SP, R	Crystalis 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	Synchrony 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.

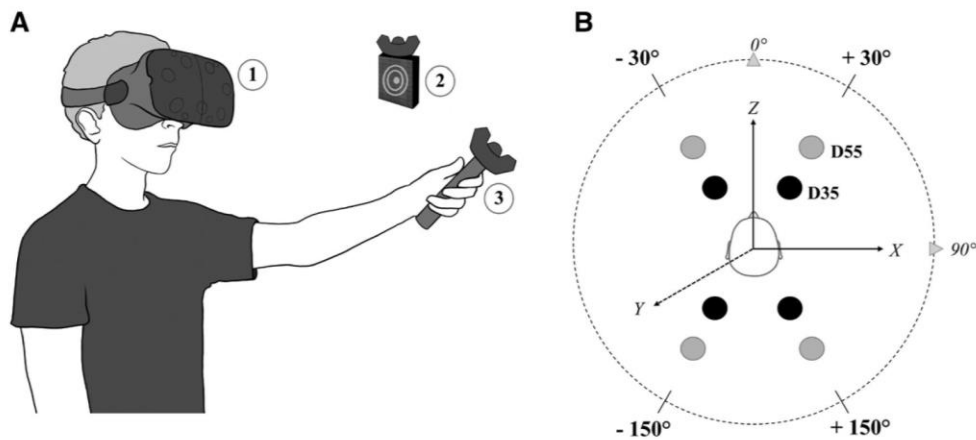


Fig. 1. Experimental setup. A, Apparatus based on the virtual reality system, comprising (1) a head-mounted display (HTC VIVE), (2) a VIVE tracker mounted on a loudspeaker, and (3) another tracker mounted on a hand-held pointer. Head and trackers positions were recorded in real time by two cameras, and 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.

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; $68.3 \times 82.7 \times 30.8$ 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 $+30^\circ$, $+150^\circ$, -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 ($13 \frac{3}{4}''$) from the participant's head, and D55 at 55 cm ($21 \frac{5}{8}''$). 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: $3D\ error = \sqrt{\bar{C}^2 + \bar{s}^2}$, where \bar{C} is the mean of the vector norm between the sound source and the hand pointing position, and \bar{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 (compared to 0.5% for NH peers). This figure also shows that their hand pointing was more lateralized, without a clear distinction

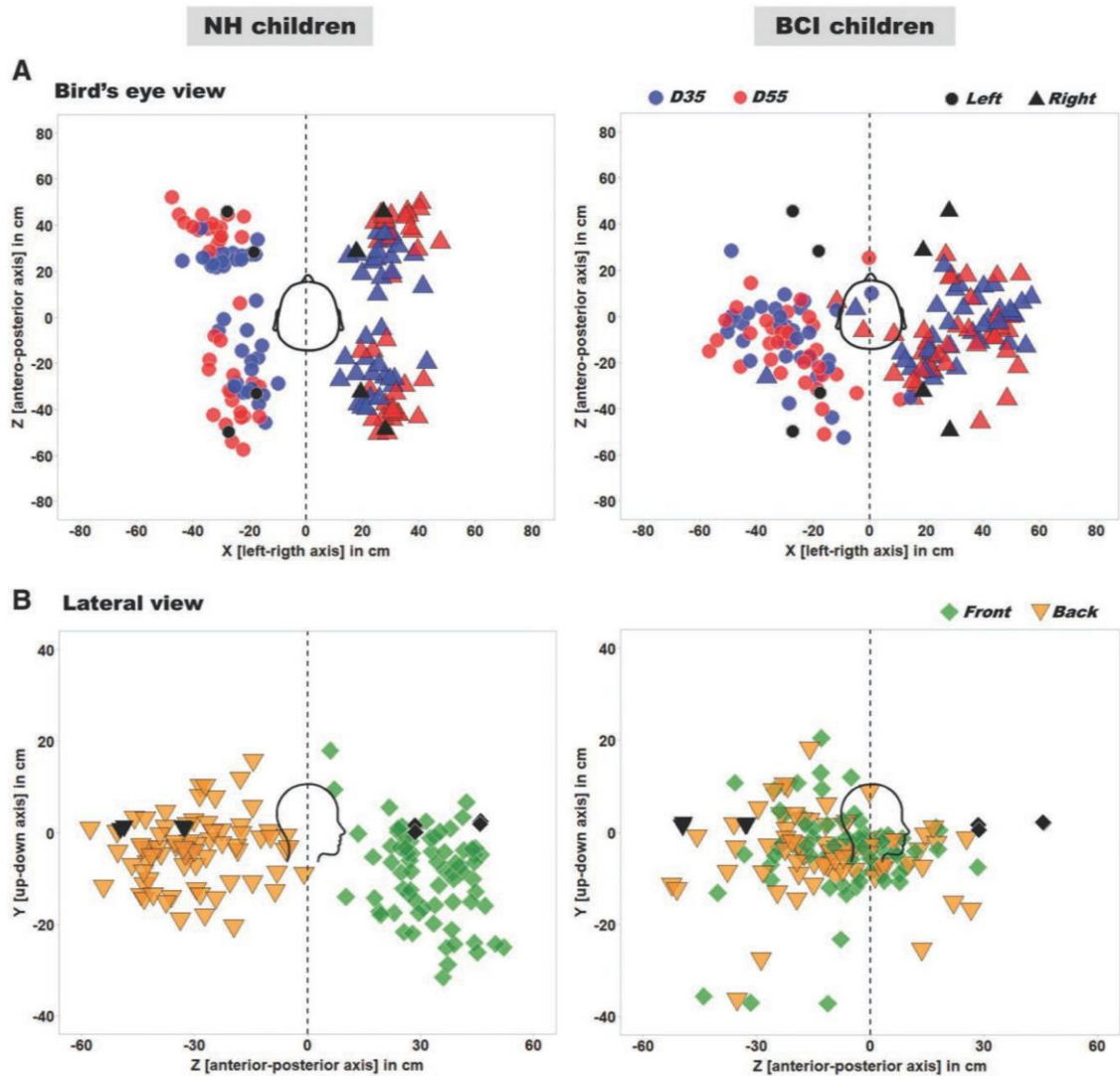


Fig. 2. Three-dimensional sound localization performance of normal-hearing (NH) and bilateral cochlear implant (BCI) children under the head immobile condition. Black symbols represent the sound sources and colored dots correspond to the mean response of each participant per target. A, Bird's eye view 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).

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 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

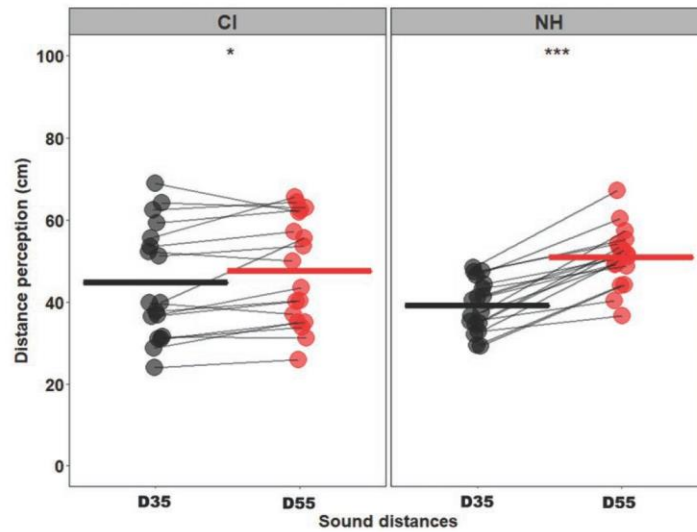


Fig. 3. Sound distance perception in normal-hearing (NH) and bilateral cochlear implant (BCI) children under the head immobile condition. Thick lines represent the mean response distances for each group for D35 (black lines: i.e., sound sources at 35 cm), and D55 (red lines: i.e., sound sources at 55 cm). Thin black lines join black and red dots for each participant. Asterisks indicate significant differences (paired t-test, $*p < 0.05$; $***p < 0.001$).

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,

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.

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.

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 back sound sources (Fig. 4A). Visualization of pointing data for NH participants are available in Supplemental Digital Content

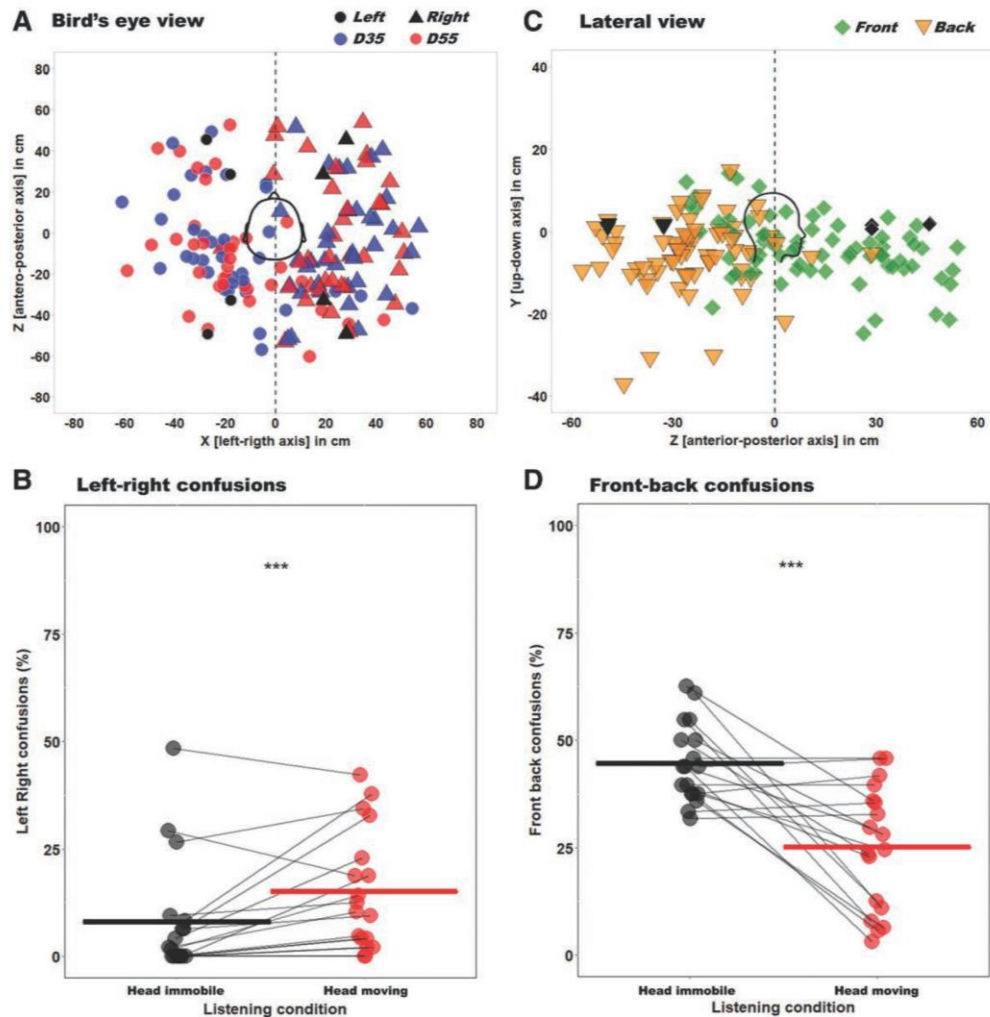


Fig. 4. Three-dimensional sound localization performance of children fitted with bilateral cochlear implant during the head moving condition. A, Bird's eye view showing hand responses as a function of stimulation side (circles for left sounds and triangles for right sounds) and distances (blue and red 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.

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 index and the percentage of trials with head movements for the BCI group

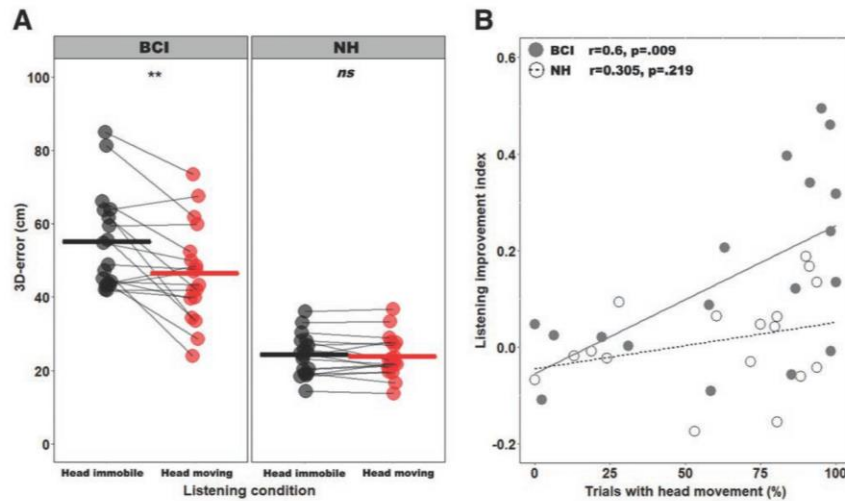


Fig. 5. Effect of head motion on spatial performance. A, Three-dimensional 3D error in both groups (BCI and NH) as a function of listening condition. Thick lines represent the mean 3D error within each group during HI listening (black) and HM listening (red) condition. Thin black lines join black and 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.

(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).

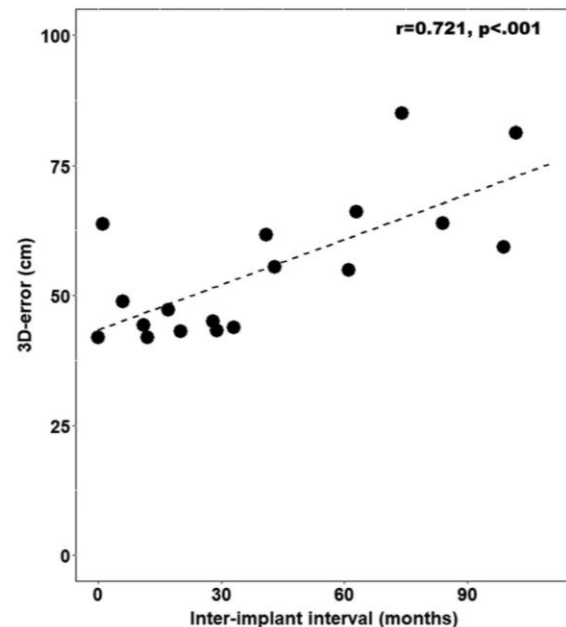


Fig. 6. Three-dimensional error (in centimeters) under the head immobile condition as a function of interimplant interval (in months).

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., Valzoghger 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

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., Loisel, 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., Rübtsamen, 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., Jiwani, 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., Rübtsamen, 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., Salemmme, 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., Salemmme, 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 Hoessel, 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., Salemmme, 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.

Etude 2

KID SSQ12

Développement et validation d'une forme courte du questionnaire pédiatrique SSQ

Soumis prochainement dans Journal of Clinical Medicine

Adaptation du SSQ pédiatrique en forme courte : le KID SSQ12

Aurélie COUDERT, Annie MOULIN, Eric TRUY, Valérie GAVEAU

Article soumis prochainement sous ce format

Introduction

L'évaluation subjective des performances auditives est un élément majeur dans la prise en charge sur le long terme d'enfants atteints de surdité. Elle permet de suivre l'évolution du ressenti et des acquis de l'enfant après mise en place d'un appareillage auditif et offre une vision complémentaire des tests objectifs (e.g. audiométrie tonale et vocale) réalisés dans le suivi médical. Peu de questionnaires de qualité de vie pédiatriques spécifiques à l'audition existent dans la littérature. Ils sont très souvent généralistes et évaluent le bien-être physique, psychique, et l'insertion sociale de l'enfant (pour revue, voir les recommandations de la SFORL 2018 ; Simon et al. 2019). Le seul à avoir été validé en français est le KINDL Cochlear Implant Module (Warner-Czyz et al. 2011) chez 138 enfants implantés cochléaires (IC) âgés de 4 à 16 ans. Cette échelle est composée de 6 items permettant un remplissage rapide (e.g. d'items : « la semaine passée, combien de fois tes amis ont-ils trouvé tes IC cool ? » ; « la semaine passée, combien de fois tes implants t'ont-ils dérangé ? ») mais néanmoins succinct des compétences réelles de l'enfant.

Afin d'évaluer plus spécifiquement un large panel des capacités auditives du quotidien, Galvin et Noble (2013) ont développé le Speech, Spatial, and Qualities of Hearing Scale (SSQ) for children, adapté du SSQ pour l'adulte par Gatehouse et Noble (2004). Le questionnaire initial est composé de 49 questions réparties en 3 sections évaluant des fonctions auditives différentes : section A – audition de la parole, section B – audition spatiale, et section C – qualités d'audition. L'adaptation pédiatrique suit le même format mais n'a conservé que 33 questions afin que l'enfant soit en mesure de le remplir directement sans aide parentale. Les auteurs conseillent toutefois de ne pas le proposer à des enfants de moins de 10 ans car l'auto-évaluation est difficile en-dessous de cet âge. Actuellement, le SSQ for children n'a été validé dans aucune autre langue et n'a aucun équivalent français.

Pour les besoins de l'étude SPHERIC (Coudert et al. 2021), le SSQ for children a été traduit en français et rempli en face à face par 18 enfants bi-implantés cochléaires (BCI) âgés de 8 à 17 ans. La passation du questionnaire a révélé quelques limites. Tout d'abord, le temps de

remplissage était en moyenne de 30 minutes, conduisant à une baisse du niveau attentionnel des plus jeunes enfants. De plus, le taux de non-réponse sur l'ensemble du questionnaire était élevé : 19 questions sur 33 (dont 12 dans la section C) présentaient au moins une fois la réponse « ne sait pas » ou « non applicable-NA ». La dernière question de la section C (#B13) avait d'ailleurs le taux de non-réponse le plus important (5 ont répondu « ne sait pas »). Enfin, la sensibilité du questionnaire n'était pas suffisante pour révéler une association avec les difficultés objectives de spatialisation auditive (SA) enregistrées lors du test en réalité virtuelle (RV).

Aussi, l'objectif principal de cette étude était de développer en français une version courte du SSQ for children à partir de questionnaires existants (Galvin & Noble 2013, Moulin et al. 2019). Le deuxième objectif était de valider cette nouvelle version dans une population d'enfants porteurs d'un ou deux IC afin d'obtenir un outil d'évaluation simple des difficultés auditives subjectives.

Matériel et méthode

Cette étude prospective a été approuvée par le Comité d'Éthique (Ile de France VIII, N°19.09.02.63029 RIPH3).

Participants

Cent enfants bi-implantés cochléaires (BCI) âgés de 7,3 à 17,8 ans (moyenne \pm écart type ; 12,5 \pm 2,9 ans) et 54 enfants mono-implantés cochléaires (UCI) âgés de 7,7 à 17,8 ans (moyenne \pm écart type ; 13, 8 \pm 2,2 ans) ont été inclus dans cette étude. Parmi les UCI, 67% portaient une prothèse controlatérale. Quatre enfants inclus dans le groupe UCI étaient d'anciens BCI ayant abandonné le port du premier IC.

Pour les groupes BCI et UCI respectivement, les trois principales étiologies étaient « origine inconnue » (42 et 39%), « génétiques » (35 et 33%), et « acquises » (9 et 11%).

Sur le plan scolaire, 21 % des enfants BCI et 39 % des UCI étaient scolarisés dans une structure spécialisée (e.g. ULIS, SEGPA, Institut National de Jeune Sourd).

Les caractéristiques cliniques et audiométriques de chaque groupe (BCI et UCI) sont synthétisées dans le Tableau VI.1.

Tableau VI.1 : Caractéristiques de la population comprenant les données cliniques et les audiométriques en fonction du groupe.

		BCI n=100	UCI n=54
Données cliniques	Sex ratio	1,4	1,2
	Age moyen au diagnostic en mois	17,6 (14,2)	22,2 (19,8)
	Age moyen à l'appareillage en mois	19,0 (13,8)	26,5 (19,4)
	Age moyen au 1 ^e IC en mois	41,0 (30,0)	62,1 (45,8)
	Délai inter-IC moyen en mois	38,8 (35,3)	NA
	Age moyen au test en ans	12,5 (2,9)	13,8 (2,2)
	Nb 1 ^e IC gauche	31	17
	Nb 1 ^e IC droit	54	37
	Nb IC bilatéral simultané	15	NA
	Nb UCI avec prothèse	NA	36
	Scolarisation en structure classique	79	33
	Scolarisation en structure adaptée	21	21
Données audiométriques	PTA OG en dB HL	25,5 (5,9)	NA
	PTA OD en dB HL	24,4 (5,7)	NA
	PTA IC seul en dB HL	NA	26,0 (7,0)
	Intelligibilité avec deux IC en %	32,7 (6,7)	NA
	Intelligibilité IC seul - monauraux	NA	36,7 (9,7)
	Intelligibilité IC + PA - bimodaux	NA	34,4 (6,6)

Les déviations standard sont notées entre parenthèses

Cent enfants étaient inclus dans le groupe BCI, 54 dans le groupe UCI dont 36 étaient porteurs d'un IC seul et 18 d'un IC avec une prothèse controlatérale.

Le PTA était calculé pour chaque oreille séparément sans port de la prothèse ou de l'IC controlatéral.

Le seuil d'intelligibilité était mesuré avec les 2 IC pour les BCI (n=100) ; avec 1 IC seul (n=18) ou IIC et une prothèse controlatérale (n=36) pour les UCI.

IC : implant cochléaire ; BCI : bi-implanté cochléaire ; UCI : mono-implanté cochléaire ;

OG : oreille gauche ; OD : oreille droite ; Nb : nombre d'enfants avec ; PA : prothèse auditive ; PTA : seuil audiométrique tonal

Construction du questionnaire

La création d'un questionnaire court à partir d'une forme longue validée (i.e. faisant office de référence) nécessite de respecter certaines règles de base :

- Suivre la même structure interne que la forme longue de référence en conservant le même nombre de sections (Widaman et al. 2011),
- Garder par section un nombre d'items compris entre 3 et 5 (Loewenthal & Lewis 2001),
- Eliminer les items dont le taux de non-réponse est élevé dans la forme longue.

Pour le développement de la forme courte du SSQ pédiatrique, que nous appellerons par la suite le KID SSQ12, nous avons gardé les 3 sections du SSQ qui font référence à différentes fonctions de l'audition (section A : audition de la parole, section B : audition spatiale, section C : qualité d'audition) et nous avons intégré 4 items pour chaque section. Pour la section A, chaque item était nommé de #A.1 à # A.4, pour la section B de #B.1 à # B.4, et pour la section C de #C.1 à # C.4.

Le choix des items a été réalisé en combinant la forme courte du questionnaire adulte, le 15iSSQ (Moulin et al. 2019), et la forme longue du questionnaire pédiatrique que nous appellerons le SSQ33 (Galvin & Noble 2013). Les questions communes à ces deux questionnaires mais présentant un taux de non-réponse élevé ont été éliminées. Le Tableau VI.2 compare les deux versions et permet de visualiser les items inclus dans le KID SSQ12.

La cotation de chaque item, 0 pour « je n'y arrive pas du tout » à 10 pour « j'y arrive parfaitement », a été préservée. Nous avons également conservé pour chaque item la réponse « non applicable » afin de laisser l'opportunité à l'enfant de ne pas répondre.

Dans son format initial adulte (Gatehouse & Noble 2004, Gatehouse & Akeroyd 2006), les 49 items du SSQ avaient été classifiés en 10 catégories faisant référence à différentes situations auditives (e.g. parole dans le bruit, localisation, ségrégation des sons, effort d'écoute). Dans l'élaboration du SSQ33, Galvin et collaborateurs (2013) avaient réduit le nombre d'items tout en conservant 90% des catégories afin qu'une majeure partie soit représentée dans le format pédiatrique. Nous avons suivi la même méthodologie en conservant 5 catégories dans le KID SSQ12 : « speech in noise », « speech in speech contexts », « localization », « distance and movement », « identification of sounds » (voir Tableau VI.3).

Trois catégories de la forme initiale n'ont pas été retenues pour les raisons suivantes :

- « Speech in quiet » : les items de cette catégorie sont majoritairement bien côtés et ne permettent pas de faire émerger les difficultés des patients (Noble et al. 2013)
- « Multiple speech stream and switching » : les items traitant de la gestion des flux de paroles ne sont pas toujours bien compris par les plus jeunes enfants et sont à risque de non-réponse (Galvin & Noble 2013)
- « Sound quality and naturalness » : les enfants n'ayant pas eu d'expérience auditive avant implantation cochléaire (e.g. sourds profonds congénitaux) ne sont pas en mesure d'effectuer une comparaison sur la clarté du son avant et après implantation (Galvin & Noble 2013)

Tableau VI.2 : comparaison des items par section dans la littérature

	<u>KID SSQ 12</u>	Galvin et al. 2013 (SSQ33)	Moulin et al. 2019 (15iSSQ)
Section A : Audition de la parole	A.1 - Tu discutes avec quelqu'un, la TV est allumée dans la même pièce. Peux-tu comprendre ce qu'il te dit sans baisser le son ?	A.1 - You are talking with your Mum or Dad and there is a TV on in the same room. Without turning the TV down, can you understand what your Mum or Dad is saying to you?	A.1 - Vous discutez avec une autre personne dans une pièce dans laquelle un téléviseur est allumé. Pouvez-vous suivre les propos de votre interlocuteur sans baisser le son du téléviseur ?
	A.2 – Tu es assis autour d'une table avec un groupe de 5 personnes environ. L'endroit est bruyant, comme une salle de classe où les élèves se déplacent et discutent. Tu peux voir toutes les personnes autour de la table. Peux-tu comprendre ce que les personnes du groupe disent ?	A.4 - You are in a group of about five people, sitting round a table. It is a noisy room, like a busy classroom where students are moving around and talking. You can see everyone in the group around your table. Can you understand what the group is talking about?	A.2 - Vous êtes assis autour d'une table avec un groupe de cinq personnes environ, dans un restaurant animé. Vous pouvez voir toutes les personnes du groupe. Pouvez-vous suivre la conversation ?
	A.3 - Tu es assis autour d'une table avec un groupe de 5 personnes environ. L'endroit est bruyant, comme une salle de classe où les élèves se déplacent et discutent. Tu ne peux pas voir toutes les personnes autour de la table. Peux-tu comprendre ce que les personnes du groupe disent ?	A.6 - You are in a group of about five people, sitting round a table. It is a noisy room, such as a busy classroom where students are moving around and talking. You cannot see everyone in the group around your table. Can you understand what the group is talking about?	A.4 - Vous êtes assis autour d'une table avec un groupe de cinq personnes environ, dans un restaurant animé. Vous NE pouvez PAS voir toutes les personnes du groupe. Pouvez-vous suivre la conversation ?
	A.4 - Tu discutes avec quelqu'un dans une pièce où beaucoup d'autres personnes parlent. Peux-tu comprendre ce qu'il te dit ?	A.8 - You are talking to one person in a room in which there are many other people talking. Can you understand what the other person is saying?	A.5 - Vous discutez avec quelqu'un dans une pièce dans laquelle beaucoup d'autres personnes parlent. Pouvez-vous suivre ce que vous dit votre interlocuteur?

Section B : Audition spatiale	<p>B.1 - Tu es à la maison et tu entends un bruit fort (une porte qui claque ou l'aspirateur). Peux-tu dire immédiatement d'où vient le bruit ?</p>	<p><i>Aucune correspondance exacte</i></p> <p>B.4 - You are at home and it is quiet. Your Mum or Dad calls you from another room. Will you know where they are without having to look?</p>	<p><i>Aucune correspondance exacte</i></p>
	<p>B.2 - Tu es dehors. Un chien aboie très fort. Peux-tu dire immédiatement où il se trouve, sans le voir ?</p>	<p>B.5 - You are outside. A dog barks loudly. Can you tell straight away where the dog is, without having to look?</p>	<p>B.2 - Vous êtes à l'extérieur. Un chien aboie bruyamment. Pouvez-vous indiquer immédiatement où il se trouve, sans regarder ?</p>
	<p>B.3 - Tu es sur le trottoir d'une rue où il y a beaucoup de bruit. Tu entends un bus ou un camion. Peux-tu dire immédiatement d'où il vient avant de le voir ?</p>	<p>B.6 - You are standing on the footpath of a busy street. You can hear a bus or truck. Can you tell right away where it is coming from before you see it?</p>	<p>B.3 - Vous êtes sur le trottoir d'une rue animée. Pouvez-vous entendre immédiatement de quelle direction un bus ou un camion arrive avant de l'avoir vu ?</p>
	<p>B.4 - Tu entends des voix ou des bruits de pas. Juste en écoutant, peux-tu dire où la personne va ? (par exemple, est-ce que la personne se rapproche de toi ou se déplace de ta gauche vers ta droite ?)</p>	<p>B.8 - You can hear voices or footsteps. Can you tell if the person is coming towards you or moving away just by the sound?</p>	<p>B.4 - Pouvez-vous indiquer dans quelle direction une personne se déplace, uniquement au son de sa voix ou de ses pas, par exemple de votre gauche à votre droite ou inversement?</p>
Section C : Qualité d'audition	<p>C.1 - Peux-tu reconnaître facilement la voix des personnes de ta famille, de tes amis, quand tu les entends arriver sans les voir ?</p>	<p>C.3 - Do you know which person in your family is talking just by the sound of their voice, even if you can't see them?</p>	<p>C.1 - Pouvez-vous reconnaître facilement les différentes personnes que vous connaissez, au son de leur voix ?</p>
	<p>C.2 - Tu entends une chanson ou une musique que tu connais bien. Arrives-tu facilement à la reconnaître ?</p>	<p>C.4 - You can hear a song you know being played. Is it easy for you to tell what song it is just by listening?</p>	<p>C.2 - Pouvez-vous reconnaître facilement les différents morceaux de musique que vous connaissez ?</p>
	<p>C.3 - Peux-tu faire la différence entre deux sons</p>	<p>C.5 - Can you tell the difference between noises</p>	<p>C.3 - Pouvez-vous différencier certains bruits,</p>

	qui se ressemblent, comme par exemple un robinet qui coule et un ventilateur ?	that are a bit the same, like a car versus a bus, OR the tap running and a fan?	par exemple une voiture par rapport à un bus ou de l'eau qui bout par rapport à la nourriture qui frit dans une poêle ?
	C.4 - Tous les jours, tu entends les mêmes bruits à la maison. Est-ce que tu peux les reconnaître facilement ? (Par exemple, le bruit de la cafetière, la sonnerie du téléphone, l'eau qui coule...)	<i>Aucune correspondance exacte</i>	<i>Aucune correspondance exacte</i> C.5 - Les bruits quotidiens que vous entendez facilement, vous semblent-ils clairs et distincts (non brouillés, non mélangés) ?

Tableau VI.3 : Items sélectionnés dans chaque catégorie (d'après Gatehouse & Akeroyd, 2006) pour les 3 formes de SSQ (SSQ33, 15iSSQ, KID SSQ12)

Section	Catégorie	Nb items SSQ33	Nb items 15iSSQ	Nb items KID SSQ12	N° items KID SSQ12
A	Speech in quiet	2	0	0	A.1 ; A.2 ; A.3 A.4
	Speech in noise	4	4	3	
	Speech in speech contexts	2	1	1	
	Multiple speech stream and switching	1	0	0	
B	Localization	5	3	2	B.1 ; B.2
	Distance and movement	7	2	2	B.3 ; B.4
C	Identification of sounds	4	3	4	C.1 ; C.2 ; C.3 ; C.4
	Segregation of sounds	2	0	0	
	Sound quality and naturalness	0	2	0	
	Listening effort	3	0	0	

Afin de faciliter la bonne compréhension du questionnaire, l'échelle numérique du 15iSSQ (Moulin et al. 2019) a été modifiée pour que les enfants puissent appréhender cette auto-évaluation (Figure VI.1). Des smileys (« non content » pour une cotation à 0 », à « content » pour une cotation à 10) et un dégradé de couleur (clair pour 0, à foncé pour 10) ont été intégrés ; la forme de l'échelle a été modifiée pour ressembler à un entonnoir mimant l'intensité croissante de la réponse. Au terme de quelques passations pilotes, une minorité d'enfants de moins de 9 ans n'arrivaient pas à se projeter avec l'échelle numérique. Une échelle

additionnelle sans chiffre a été développée dans un second temps (Figure VI.2), sur le modèle de l'échelle des visages utilisée dans l'évaluation de la douleur.

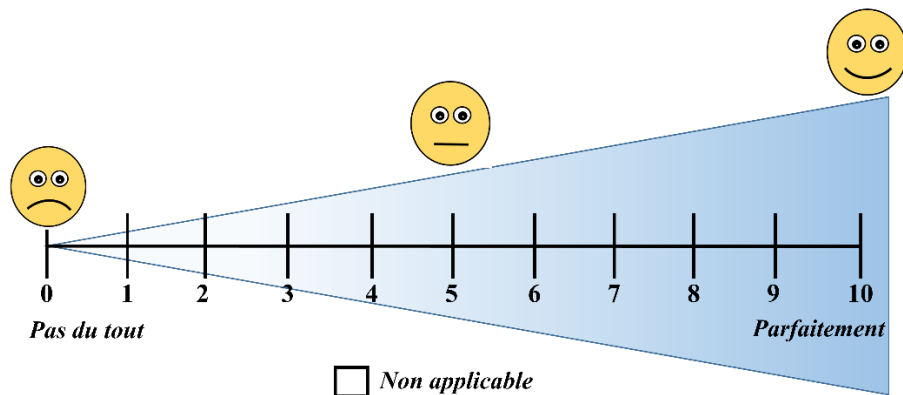


Figure VI.1 : Echelle visuelle du KID SSQ12

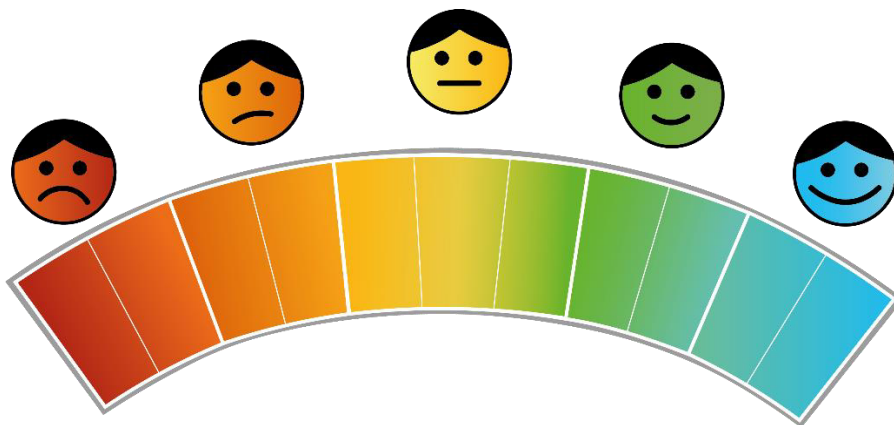


Figure VI.2 : Echelle des visages du KID SSQ12

Cette échelle est à privilégier pour les enfants âgés de 7 à 9 ans. Chaque case correspond implicitement à un chiffre (e.g. la première case à 0, la dernière à 10) permettant ainsi une cotation rapide par l'examineur

Modalités de passation

La passation du questionnaire était réalisée une seule fois en face à face avec l'enfant sous forme d'interview avec utilisation d'un format papier. Les items étaient lus à haute voix par l'expérimentateur ou l'enfant et, si besoin après lecture de l'item, l'expérimentateur donnait un exemple du quotidien pour que l'enfant se projette plus facilement dans la situation énoncée. Puis l'enfant entourait lui-même un chiffre ou un smiley sur la réglette présentée en face de

l'item (Annexe A de l'Etude 2). Il avait pour consigne de choisir un chiffre entier et s'il ne pouvait répondre à la question, il pouvait cocher la case « non applicable ».

Analyses des données

1. Structure interne

La validité interne du questionnaire était évaluée en recherchant une concordance entre la structure du SSQ33 de Galvin et celle du KID SSQ12. Nous faisons l'hypothèse que les données recueillies dans le KID SSQ12 permettaient de retrouver la structure en 3 sections du SSQ33. Ainsi, les items ayant un lien structurel fort (i.e. se rapportant au même concept théorique) doivent être regroupés dans la même section. Selon la méthode de Revelle (1978, 1979) une analyse en cluster était réalisée ; elle repose sur deux indices : le coefficient α (mesure de la cohérence interne de l'échelle) et le coefficient β (mesure de l'homogénéité factorielle). En résumé, la structure de l'échelle est construite à partir de deux groupes d'items qui sont les plus similaires l'un à l'autre ; un item est ajouté aux deux groupes d'items initiaux uniquement si cet ajout améliore α et/ou β . Les résultats sont présentés sous forme d'un diagramme arborescent de clusters qui relie des items de moins en moins similaires de gauche à droite : les items les plus similaires sont combinés en premier, et les items les plus dissemblables sont ajoutés en dernier. Les coefficients α et β sont fournis pour chaque cluster, et les coefficients de corrélation sont donnés pour chaque connexion entre les clusters et/ou les items. Un α supérieur à 0,8 et un β supérieur à 0,7, avec une différence entre ces deux coefficients inférieure à 0,1, indiquent une bonne homogénéité et une bonne cohérence interne (Cooksey & Soutar 2006).

Dans un second temps, le coefficient alpha de Cronbach était calculé pour évaluer l'homogénéité de l'ensemble du questionnaire. Il varie de 0 à 1, et plus sa valeur est proche de 1 plus grande est l'homogénéité du test. En pratique, l'homogénéité du test est satisfaisante si le coefficient est supérieur à 0,7 (Cronbach 1951).

2. Validité externe

A l'échelle de chaque groupe (BCI et UCI) ont été calculés le score moyen total au questionnaire ainsi que le score moyen pour chaque section (A, B et C).

Le seuil audiométrique tonal (PTA) en dB HL était calculé pour chaque implant en moyennant les seuils obtenus pour les fréquences 0,5 ; 1 ; 2 ; et 4 kHz.

Le seuil d'intelligibilité (i.e. le seuil en dB HL pour lequel l'enfant arrivait à répéter 50% des mots de la liste de Lafon) était calculé pour les UCI du côté de l'IC seul, et pour les BCI avec les 2 IC simultanément.

Trois facteurs étaient calculés spécifiquement dans le groupe des BCI : la différence entre le PTA de chaque implant (en dB HL), la durée d'expérience binaurale avec les deux ICs (en mois), et le délai entre les deux implantations cochléaires (en mois).

La validité externe du questionnaire était testée en évaluant la corrélation entre le score moyen total et les scores moyens par section avec l'âge au moment du test, le PTA, l'écart entre les 2 PTA, le seuil d'intelligibilité, la durée d'exposition au premier implant, la durée d'expérience binaurale, et le délai inter-implant.

3. Comparaison avec les données de l'étude SPHERIC

A titre comparatif, les scores moyens par section obtenus préalablement dans l'étude SPHERIC (Coudert et al. 2021) sur 18 enfants bi-implantés ont été analysés graphiquement avec ceux du groupe BCI de cette étude. Les groupes n'étant pas homogènes, aucune analyse statistique n'a été réalisée.

Résultats

Remplissage du questionnaire

Tous les enfants inclus ont rempli sans difficulté le questionnaire en moins de 10 minutes. Le taux de non-réponse (i.e. réponse « non applicable – NA ») était de 0,87% pour l'ensemble des participants (soit 16 réponses au total). Les deux items ayant le plus de réponses NA étaient #C.2. (5 répondants sur 154) et #B.4 (4 répondants sur 154). Ces non-réponses concernaient 12 patients (i.e. 7,8% des enfants) ; le maximum de non-réponse pour un enfant était de 2. Ces 12 enfants n'étaient pas significativement plus jeunes que le reste de la population (âge moyen de 13 ans). Le développement de l'échelle des visages sans chiffre a été utilisée par 5 enfants BCI de moins de 9 ans.

Analyse des scores du SSQ

Le score moyen total était équivalent dans les deux groupes : 6,6 ($\pm 1,3$). A l'échelle des sections, les scores moyens des groupes BCI et UCI étaient respectivement de 5,6 ($\pm 1,8$) et 5,9 ($\pm 1,7$) pour la section A, 6 ($\pm 2,1$) et 5,8 ($\pm 1,9$) pour la section B, et 8,3 ($\pm 1,4$) et 8,1 ($\pm 1,4$) pour la section C (Figure VI.3). Une ANOVA à mesures répétées a montré une différence significative entre les trois sections (Subscale $F(2, 304) = 131.86, p < .001, \eta_g^2 = .27$), la section C étant significativement plus élevée que les deux autres ($p < .0001$, correction de Bonferroni). Aucune interaction entre les groupes et les sections n'est ressortie ($F(2, 304) = 0.91, p = .405, \eta_g^2 < .01$).

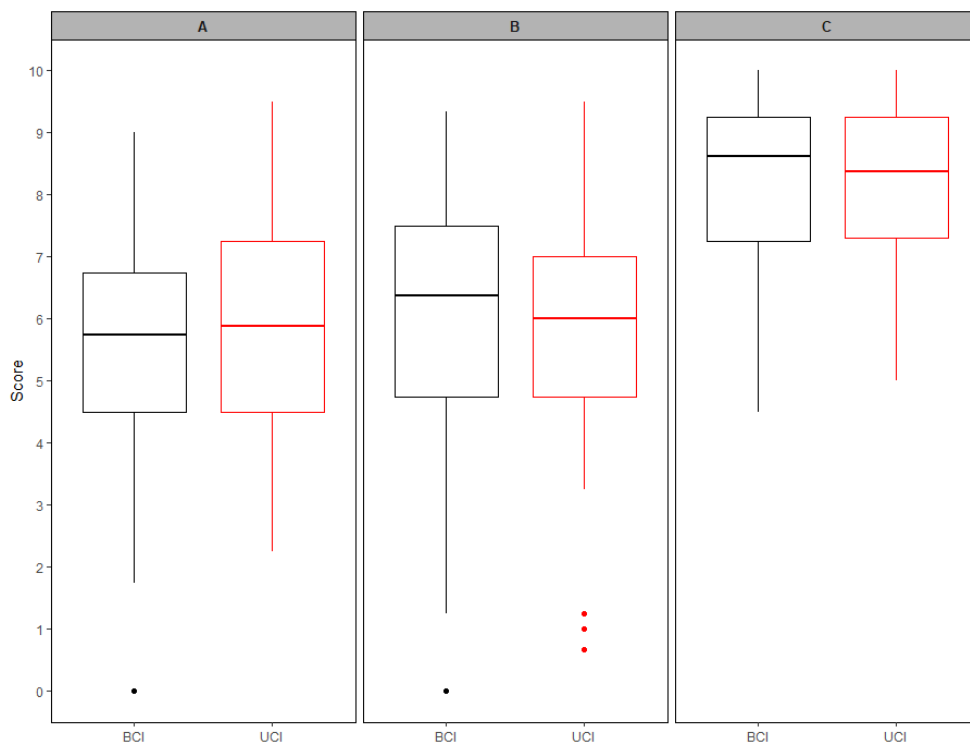


Figure VI.3 : Scores moyen par groupe (BCI avec $n=100$; UCI avec $n=54$) et section (A : audition de la parole ; B : audition spatiale ; C : qualité d'audition)

Les lignes centrales représentent les médianes, les limites des boîtes les 25^e et 75^e percentiles.

Le groupe des UCI était composé de deux catégories de patients : les UCIP appareillés en bimodal (i.e. porteurs d'une prothèse controlatérale à l'IC), et les UCIs appareillés en monaural (i.e. sans prothèse controlatérale). Les effectifs dans chaque sous-groupe étant trop faibles comparés au groupe des BCI, aucune analyse statistique n'a été réalisée. Toutefois, les scores moyens par section sont présentés en annexes B et C de ce chapitre, et montrent une tendance

émergente. Les UCIs avaient un score moyen plus faible pour les sections A et B comparés aux UCIp et aux BCI.

Structure interne

L'analyse en cluster (Figure VI.4) a permis de retrouver spontanément les 3 sections avec une bonne consistance et homogénéité interne (différence alpha et bêta < 0.10). Chaque section comprenait les 4 items appartenant à la même catégorie (cluster C10 de #A1 à #A4; cluster C7 de #B1 à #B4 et cluster C8 de #C1 à #C4). Par contre, en essayant de relier ces 3 clusters ensemble, l'homogénéité a diminué (i.e. différence alpha et bêta > 0.10 pour les clusters C9 et C11). Cela montre que l'ensemble du questionnaire explore des capacités auditives différentes qui ne peuvent pas être regroupées sous une seule catégorie.

Les coefficients alpha des 3 clusters C10, C7, et C8 étaient respectivement de 0,73 ; 0,72 ; et 0,58 attestant d'une validité interne correcte de l'échelle. Le coefficient alpha de Cronbach était de 0,78 ce qui renforce la cohérence interne du questionnaire.

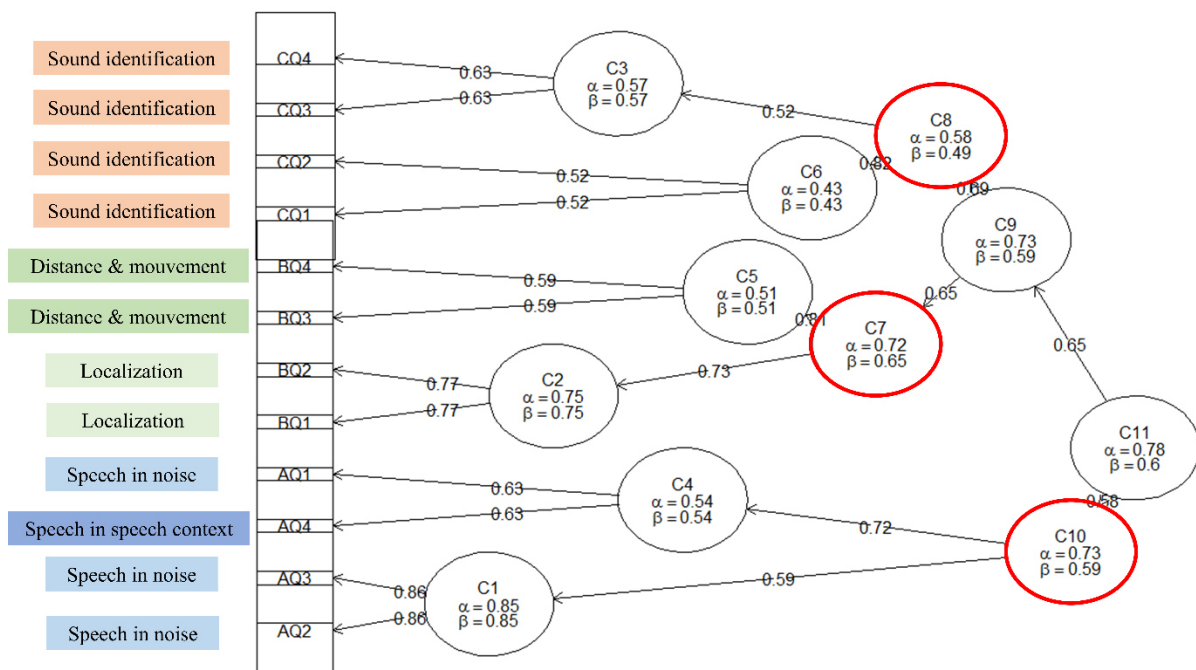


Figure VI.4 : Diagramme d'analyse en cluster pour le KID SSQ12.

Les items les plus ressemblants sont combinés ensemble à gauche. Les trois principaux clusters sont cerclés de rouge et correspondent aux 3 sections du questionnaire. Pour chaque cluster, les coefficients alpha et bêta sont notés.

Validité externe : corrélation avec des facteurs externes

Le score total au questionnaire n'était pas corrélé avec la durée d'exposition au premier implant, ni avec les seuils audiométriques tonal (PTA) et d'intelligibilité, tout groupe confondu. Une tendance émergeait entre le score moyen à la section A (i.e. audition de la parole) et le PTA ($r=0.24$, $p=0.055$). Une corrélation significative était retrouvée entre le score total et l'âge des enfants au moment du test ($r=0.23$, $p=0.02$). Ce résultat peut être attribué à une corrélation significative entre le score moyen de la section C (i.e. qualités d'audition) et l'âge ($r=0.31$, $p<0.002$, voir Figure VI.5), et particulièrement avec les items #C3 ($r=0.26$, $p=0.001$) et #C4 ($r=0.23$, $p=0.004$).

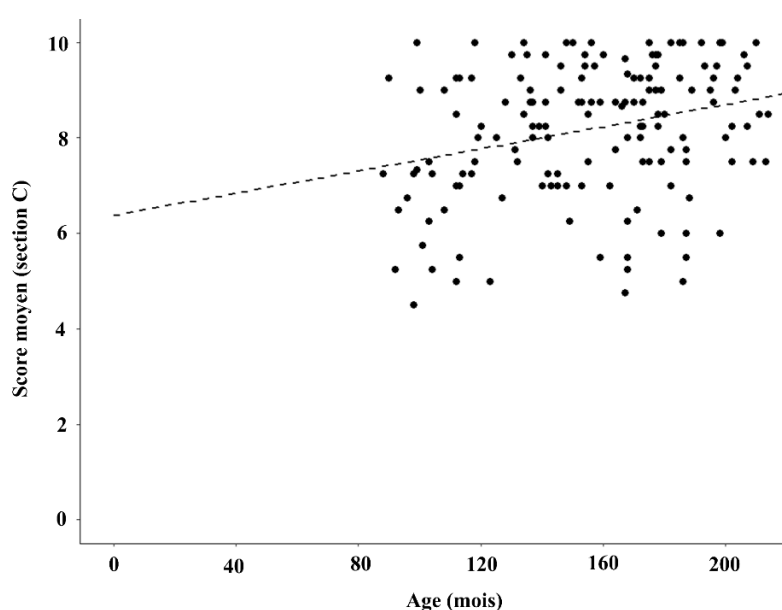


Figure VI.5 : Score moyen de la section C en fonction de l'âge des enfants implantés ($n=154$) au moment du test.

Ensuite, nous avons corrélé les données du groupe BCI avec les scores moyens à chaque section (Figure VI.6). Le score moyen de la section A était corrélé positivement avec le PTA ($r=0.294$, $p=0.003$) : plus le seuil augmentait et plus le score était élevé. Aucune corrélation n'émergeait avec le seuil d'intelligibilité et l'écart entre les deux PTA.

Une tendance émergeait entre le score moyen de la section B (i.e. localisation spatiale) et la durée d'expérience binaurale ($r=0.173$, $p=0.086$), sans corrélation avec la durée d'expérience avec le premier implant ni le délai inter implant.

Le score moyen de la section C était faiblement corrélé avec la durée d'expérience avec le premier implant ($r=0.18$, $p<0.03$), ce qui peut être attribué à la corrélation avec l'âge vue à l'échelle des 2 groupes.

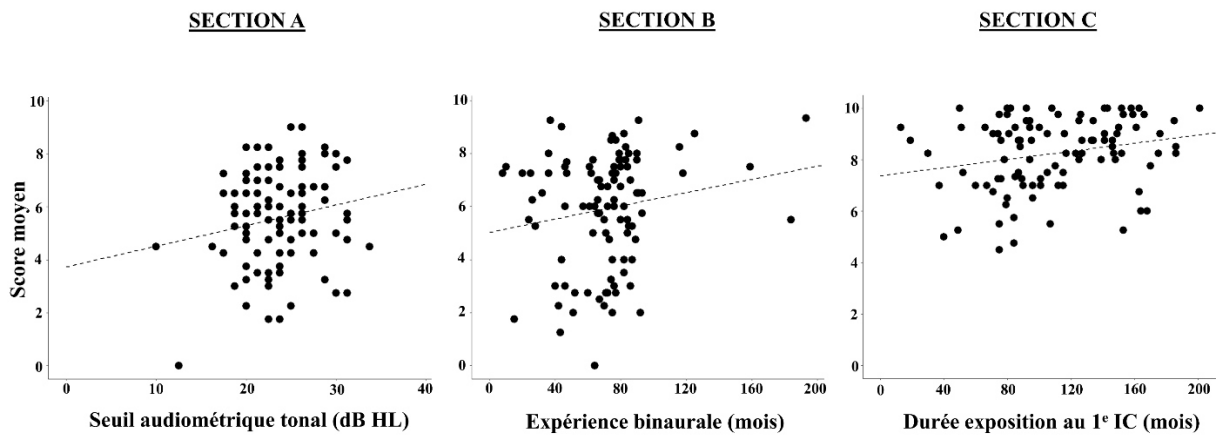


Figure VI.6 : Scores moyens à chaque section en fonction des paramètres cliniques du groupe bi-implantés (n=100).

La section A était corrélée avec le seuil audiométrique tonal, la section B avec la durée d'expérience binaurale avec les 2 implants cochléaires (IC), et la section C avec la durée d'exposition au 1^{er} IC.

Comparaison avec l'étude SPHERIC

Les scores moyens par section obtenus chez les 100 BCI ont été comparés graphiquement (Figure VI.7) avec ceux des 18 BCI de l'étude SPHERIC (Coudert et al. 2021). Le score moyen était respectivement de $5,6 \pm 1,8$ vs. $7,2 \pm 1,2$ pour la section A ; $6 \pm 2,1$ vs. $7,1 \pm 1,5$ pour la section B ; et $8,3 \pm 1,4$ vs. $7,6 \pm 1,0$ pour la section C.

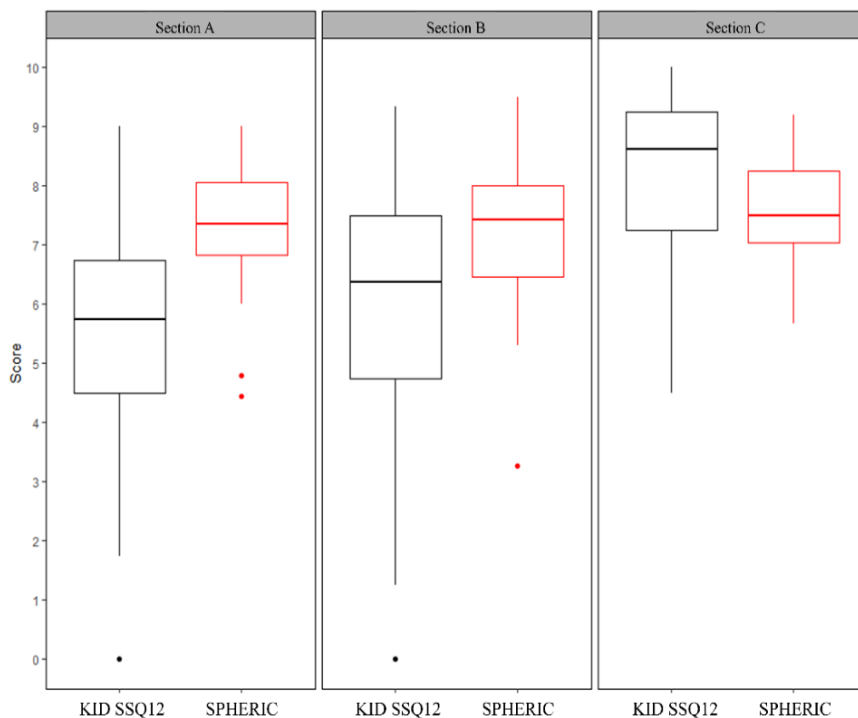


Figure VI.7 : Scores moyen par section (A : audition de la parole ; B : audition spatiale ; C : qualité d'audition) par groupe d'enfants bi-implantés cochléaires (KID SSQ12, n=100 ; SPHERIC, n=18). Les lignes centrales représentent les médianes, les limites des boîtes les 25^e et 75^e percentiles.

Discussion

L'évaluation de la qualité de vie des enfants implantés cochléaires est au cœur de la prise en charge clinique. En complément des tests objectifs, elle permet d'apprécier les bénéfices à court et long terme de l'implantation et cibler spécifiquement les difficultés auditives persistantes. Galvin et Noble (2013) ont été les premiers à développer un questionnaire pédiatrique spécifique de l'audition. Son utilisation reste limitée à des enfants de plus de 10 ans et impose un temps de passation long d'au moins 30 minutes. Dans la littérature, nous n'avons pas retrouvé d'échelles similaires en langues étrangères adaptées en format court pour une utilisation en soins courants.

Validation du premier SSQ pédiatrique de forme courte

Dans cette étude, nous avons développé un questionnaire SSQ pédiatrique composé de 12 questions (4 questions par section) puis nous l'avons validé auprès d'une large cohorte d'enfant implantés cochléaires (100 bi-implantés et 54 mono-implantés). Ce questionnaire présentait plusieurs avantages sur sa forme. Tout d'abord, il se remplissait rapidement en moins de 10 minutes. De plus, l'ensemble des questions était bien compréhensible pour des enfants âgés au minimum de 7 ans puisque moins de 1% de non-réponse ont été enregistrées sur l'ensemble des participants. L'utilisation d'une échelle visuelle numérique avec la présence de smileys était très appréciée des enfants car celle-ci leur était déjà familière des séances de réglages avec l'audioprothésiste et leur permettait de mieux apprécier la gradation de leurs difficultés.

Sur le fond, les 12 questions sélectionnées à partir du travail princeps de Galvin et Noble (2013) ont permis d'explorer une large variété de catégories auditives spécifiques : compréhension de la parole dans le bruit, localisation spatiale d'une source fixe ou en mouvement, identification des bruits du quotidien. Les items sélectionnés respectaient la structure interne de la forme longue en conservant les 3 sections (A- audition de la parole, B- localisation spatial, C- qualités d'audition).

L'aspect multidimensionnel de ce format court a révélé des profils de réponse et des corrélations différentes suivant les sections. Le score moyen de la section A augmentait avec celui du seuil audiométrique tonal (PTA). Ce résultat paraît contre-intuitif de prime abord. Tous les items de la section A dépeignent des situations en milieu bruyant. Il est possible qu'une augmentation du PTA soit également associée à une baisse du niveau de bruit perçu conduisant à une amélioration globale du rapport signal sur bruit et donc à de meilleurs scores perceptifs.

L'absence de corrélation entre le score de la section A et le seuil d'intelligibilité s'explique par les conditions de tests en cabine ne reflétant pas la réalité du quotidien (e.g. conversations en classe). En effet, le seuil est communément calculé dans le silence avec des mots isolés dissyllabiques. La section B montrait des scores d'autant plus élevés que l'expérience binaurale augmentait, confirmant toute l'importance de la binauralité dans la spatialisation auditive. Les qualités d'audition explorées dans la section C étaient significativement améliorées avec la durée d'exposition au premier IC en lien direct avec l'augmentation de l'âge de l'enfant au moment du test. Cette corrélation était portée par les items #C3 et #C4 explorant spécifiquement les capacités de séparation et reconnaissance de deux sons. Il est probable que cette capacité soit directement liée à la maturation auditive de l'enfant car celle-ci semble se poursuivre au-delà de l'âge de 5 ans (Freigang et al. 2015, Litovsky 2015).

L'ensemble des analyses réalisées suggère qu'une optimisation du questionnaire serait possible afin d'en augmenter la sensibilité. Pour cela, il faudrait sélectionner les 2 items par section ayant le plus de similitudes et avec un niveau de corrélation le plus fort avec les facteurs cliniques vus ci-dessus : items #A2, #A3, #B1, #B2, #C3, #C4.

Variabilité des réponses suivant le domaine exploré

Les données obtenues avec les 100 BCI ont également été comparées avec celles des 18 BCI évalués dans l'étude SPHERIC (voir Etude 1). Les deux populations testées n'avaient pas des âges significativement différents au moment du test (145 versus 150 mois) mais le profil de réponse était beaucoup plus hétérogène entre les 3 sections. En effet, les scores moyens du KID SSQ12 étaient bien inférieurs pour les sections A et B (respectivement 5,6 et 6 versus 7,2 et 7,1 dans l'étude SPHERIC), mais supérieur pour la section C (8,3 versus 7,6). Cette variabilité entre les sections concorde avec les données de la littérature chez l'adulte IC (Zhang et al. 2015) et l'adulte appareillé (Moulin et al. 2019). L'audition de la parole et la localisation spatiale sont souvent moins bien cotées que la qualité d'audition qui est bien restaurée avec le port de deux IC. Il est possible que l'élimination des items posant problèmes aux enfants dans la forme longue ait permis d'augmenter la sensibilité du questionnaire dans sa forme courte et révèle plus fidèlement les capacités quotidiennes des enfants. De plus, la cohorte de SPHERIC était peut-être trop faible pour faire émerger des différences entre les sections. En effet, plusieurs auteurs préconisent de valider un questionnaire sur un large effectif, entre 5 à 10 participants

par item (Kass & Tinsley 1979, Tabachnick & Fidell, 2007, Akeroyd et al. 2014), soit 70 à 120 participants dans notre cas.

La comparaison avec des données de la littérature sont difficiles à mener car le SSQ pédiatrique dans sa forme longue n'a été utilisé que récemment pour évaluer les résultats à long terme (3 ans et demi) de l'implantation cochléaire chez 11 enfants souffrant de surdité profonde unilatérale (Rauch et al. 2021). Seuls 5 enfants avaient réussi à compléter le questionnaire ; hormis un enfant âgé de 17 ans, tous avaient entre 5 et 9 ans, âge jugé comme étant trop faible par Galvin pour comprendre l'ensemble des questions du SSQ. Après 3 ans d'implantation, les scores moyens étaient plus élevés que les nôtres (section A : 8,6 ; section B : 7,3 ; section C : 8,2). Cette différence s'explique certainement par un profil de surdité moins sévère dans leur étude puisque l'oreille controlatérale à l'IC avait une audition normale.

De manière anecdotique, une autre étude récente a repris les travaux de Galvin en sélectionnant 21 items du questionnaire sur 33 mais ceux-ci étaient complétés directement par les parents (Kronenberger et al. 2021), laissant apparaître des scores corrélés à la qualité de l'audition plutôt qu'aux mesures objectives d'audiométrie.

Alternatives de la forme longue du SSQ pédiatrique

Galvin et collaborateurs ont développé une alternative au SSQ for children pour permettre une hétéro-évaluation facile des performances de l'enfant : le SSQ for parents (Galvin et al. 2007, 2013). Ce questionnaire a été adapté de la forme longue du SSQ adulte (Gatehouse & Noble 2004) et comprend 23 questions. Ses modalités de passation diffèrent du SSQ for children car les auteurs préconisent que les parents observent pendant une semaine (pour chaque section) leur enfant dans le quotidien afin d'augmenter la sensibilité des réponses. Les principales études ayant utilisé le SSQ for parents avaient pour objectifs d'évaluer les bénéfices de la bilatéralisation (Beijen et al. 2007, Lovett et al. 2010, Sparreboom et al. 2012, Galvin & Mok 2016). Ainsi, les parents remplissaient le questionnaire avant l'intervention (i.e. en condition mono-implantée) et jusqu'à 24 mois après (en condition bi-implantée). Les scores moyens aux 3 sections augmentaient significativement après 2 ans d'expérience binaurale mais le bénéfice était le plus important pour les items de localisation spatiale. Du fait de l'hétéro-évaluation, ce questionnaire permet d'évaluer de plus jeunes enfants, de moins de 5 ans, voire moins de 2 ans (Lovett et al. 2010). Mais les scores obtenus doivent être interprétés avec prudence car sur cette tranche d'âge inférieure, certains items peuvent être difficilement cotables par les parents (e.g. « conversations de groupe en milieu bruyant » ne s'appliquent pas

au quotidien d'un enfant de 3 ans). De plus, les bénéfices attribuables à la deuxième implantation cochléaire doivent être mis en balance avec la maturation auditive naturelle à cet âge.

Comparaison avec les SSQ de forme courte pour adulte

La forme longue du SSQ adulte de Gatehouse et Noble (2004) a déjà été adaptée en formes courtes dans plusieurs pays. Les versions allemande (Kiessling et al. 2011), anglaise (Noble et al. 2013) et française (Moulin et al. 2019) ont choisi de sélectionner entre 12 et 15 items afin de conserver la structure initiale en 3 sections. L'étude de Noble se démarque par le nombre conséquent de sujets inclus (n=1220) et la diversité des surdités (386 patients non appareillés, 627 porteurs d'une prothèse unilatérale, 207 appareillés bilatéralement). Elle se rapproche de près de la nôtre dans le choix des items mais aussi son objectif d'utilisation. Ainsi, après avoir sondé au préalable 3 centres cliniques indépendants, les auteurs ont sélectionné les items liés à la compréhension de la parole en milieu bruyant, la localisation spatiale avec et sans mouvement de la source et la ségrégation des sons. Les items traitant de la compréhension de la parole dans le silence étaient supprimés du fait d'une surestimation des scores.

Une version beaucoup plus courte de 5 questions a été développée par Demeester et collaborateurs (2012). L'utilisation de ce questionnaire était différente des précédents car celui-ci servait d'outil de dépistage rapide d'une perte auditive chez des sujets normo-entendant âgés de plus de 55 ans.

Il est notable de constater qu'aucune des études sus-citées n'a validé son questionnaire chez des patients implantés cochléaires. La majorité de ces études s'accorde sur l'intérêt des formes courtes comme outils de suivi clinique et rééducatif (Noble et al. 2013, Moulin et al. 2019).

Conclusion

Le KID SSQ12 est un nouvel outil pédiatrique utilisable en pratique clinique pour suivre l'évolution dans le temps des bénéfices de l'implantation cochléaire, à l'image du questionnaire SSQ destiné aux parents. Il a été validé dans une large population d'enfants bi-implantés et mono-implantés. A l'avenir, il serait intéressant de l'utiliser à grande échelle chez des enfants appareillés. De plus, une optimisation de son contenu serait envisageable, notamment pour les plus jeunes, sous un format de 6 questions.

REFERENCES

- Akeroyd, M. A., Guy, F. H., Harrison, D. L., et al. (2014). A factor analysis of the SSQ (Speech, Spatial, and Qualities of Hearing Scale). *Int J Audiol*, 53, 101–114.
- Beijen J.-W., Mylanus E.A.M., Snik A.F.M. (2007). Sound localization ability of young children with bilateral cochlear implants. *Otology and Neurotology*, 28(4): 479–485.
- Bullinger, M., Brütt, A. L., Erhart, M., Ravens-Sieberer, U., & BELLA Study Group. (2008). Psychometric properties of the KINDL-R questionnaire: Results of the BELLA study. *European Child & Adolescent Psychiatry*, 17 Suppl 1, 125-132.
- Cooksey, R. W., & Soutar, G. N. (2006). Coefficient beta and hierarchical item clustering an analytical procedure for establishing and displaying the dimensionality and homogeneity of summated scales. *Organ Res Methods*, 9, 78–98.
- Coudert, A., Gaveau, V., Gatel, J., Verdelet, G., Salemme, R., Farnè, A., Pavani, F., Truy, E. (2021). Spatial hearing difficulties in reaching space in bilateral cochlear implant children improve with head movements. *Ear Hear*. Jul 1. doi: 10.1097/AUD.0000000000001090. Online ahead of print.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297-334.
- Demeester, K., Topsakal, V., Hendrickx, J. J., et al. (2012). Hearing disability measured by the speech, spatial, and qualities of hearing scale in clinically normal-hearing and hearing-impaired middle-aged persons, and disability screening by means of a reduced SSQ (the SSQ5). *Ear Hear*, 33, 615–616.
- 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., Mok M., Dowell R.C. (2007). Perceptual benefit and functional outcomes for children using sequential bilateral cochlear implants. *Ear and Hearing*, 28(4): 470–482.
- 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(1):93-102.
- Gatehouse S, Noble W. (2004). The Speech, Spatial and Qualities of Hearing Scale (SSQ). *Int J Audiol*, 43(2) : 85–99

Gatehouse S., Akeroyd M. (2006). Two-eared listening in dynamic situations. *International Journal of Audiology*, 45(Suppl 1): 120–124.

Kass R.A. & Tinsley H.E.A. (1979). Factor analysis. *Journal of Leisure Research*, 11, 120–138.

Kiessling, J., Grugel, L., Meister, I. G., et al. (2011). Übertragung der Fragebögen SADL, ECHO und SSQ ins Deutsche und deren Evaluation. German translations of questionnaires SADL, ECHO and SSQ and their evaluation. *Z Audiol*, 6–16.

Kronenberger WG, Bozell H, Henning SC, Montgomery CJ, Ditmars AM & Pisoni DB. (2020): Functional hearing quality in prelingually deaf school-age children and adolescents with cochlear implants, *International Journal of Audiology*, DOI: 10.1080/14992027.2020.1826586

Litovsky, R. (2015). Development of the auditory system. *Handb Clin Neurol*, 129, 55–72.

Loewenthal, K. M., & Lewis, C. A. (2001). *An Introduction to Psychological Tests and Scales*. Psychology Press.

Lovett R.E., Kitterick P.T., Hewitt C.E., Summerfield A.Q. (2010). Bilateral or unilateral cochlear implantation for deaf children: an observational study. *Archives of Disease in Childhood*, 95(2): 107–112.

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 and Hearing*. 40(4):938-950.

Noble, W., Jensen, N. S., 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.

Rauch AK, Arndt S, Aschendorff A, Beck R, Speck I, Ketterer MC, Jakob TF, Hassepas F. (2021). Long-term results of cochlear implantation in children with congenital single-sided deafness. *European Archives of Oto-Rhino-Laryngology*, 278:3245–3255.

Recommandation pour la pratique clinique. (2018). Indications de l'implant cochléaire chez l'adulte et chez l'enfant. pp 32, 34.

Revelle, W. (1978). ICLUST: A cluster analytic approach to exploratory and confirmatory scale construction. *Behav Res Methods Instrum*, 10, 739–742.

Revelle, W. (1979). Hierarchical cluster analysis and the internal structure of tests. *Multivariate Behav Res*, 14, 57–74.

Sparreboom M., Snik A.F.M., Mylanus E.A.M. (2012). Sequential bilateral cochlear implantation in children: quality of life. *Archives of Otolaryngology – Head & Neck Surgery*, 138(2): 134–141.

Tabachnick B.G. & Fidell L.S. (2007). *Using Multivariate Statistics*. Boston: Pearson.

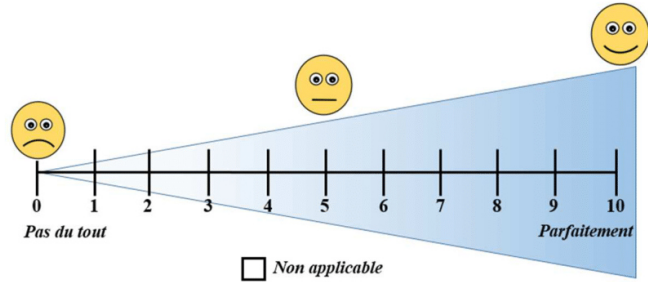
Warner-Czyz, A. D., Loy, B., Tobey, E. A., Nakonezny, P., & Roland, P. S. (2011). Health-related quality of life in children and adolescents who use cochlear implants. *International Journal of Pediatric Otorhinolaryngology*, 75(1), 95-105.

Widaman, K. F., Little, T. D., Preacher, K. J., & Sawalani, G. M. (2011). On creating and using short forms of scales in secondary research. In *Secondary data analysis : An introduction for psychologists* (p. 39-61). American Psychological Association.

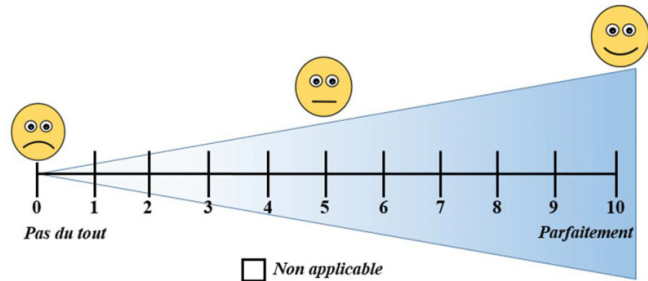
Zhang, J., Tyler, R., Ji, H., Dunn, C., Wang, N., Hansen, M., & Gantz, B. (2015). Speech, Spatial and Qualities of Hearing Scale (SSQ) and Spatial Hearing Questionnaire (SHQ) Changes Over Time in Adults With Simultaneous Cochlear Implants. *American Journal of Audiology*, 24(3), 384-397.

Partie 1 : Audition de la parole

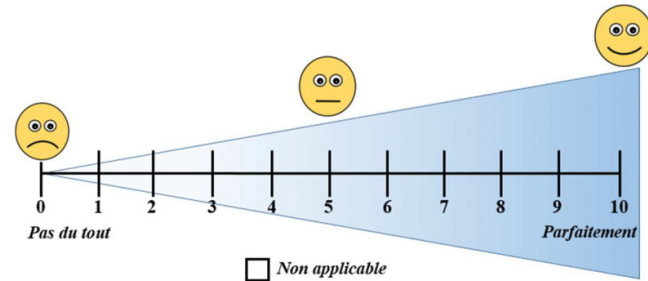
1. Tu discutes avec quelqu'un, la TV est allumée dans la même pièce.
Peux-tu comprendre ce qu'il te dit sans baisser le son ?



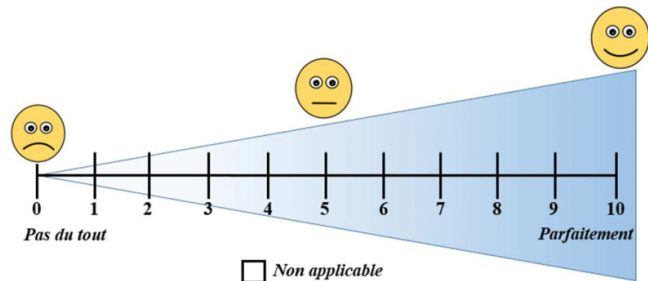
2. Tu es assis autour d'une table avec un groupe de 5 personnes environ. L'endroit est bruyant, comme une salle de classe où les élèves se déplacent et discutent. Tu **peux voir** toutes les personnes autour de la table.
Peux-tu comprendre ce que les personnes du groupe disent ?



3. Tu es assis autour d'une table avec un groupe de 5 personnes environ. L'endroit est bruyant, comme une salle de classe où les élèves se déplacent et discutent. Tu **ne peux pas voir** toutes les personnes autour de la table.
Peux-tu comprendre ce que les personnes du groupe disent ?

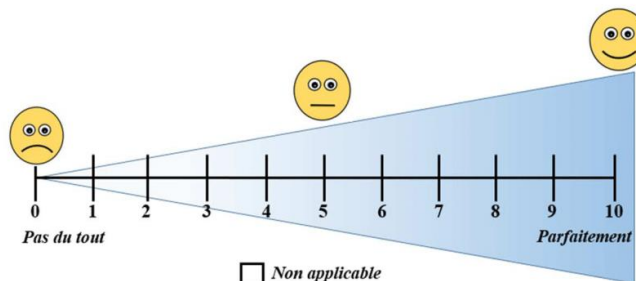


4. Tu discutes avec quelqu'un dans une pièce où beaucoup d'autres personnes parlent.
Peux-tu comprendre ce qu'il te dit ?

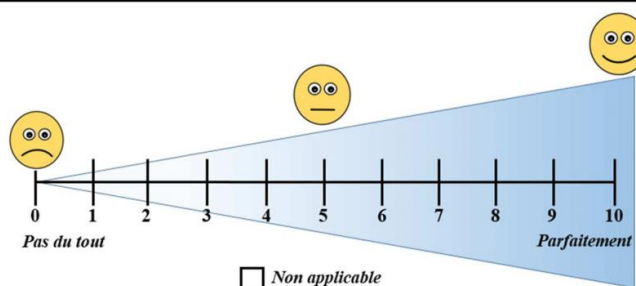


Partie 2 : Audition spatiale

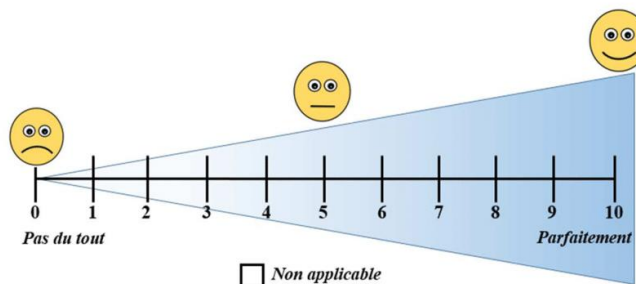
1. Tu es à la maison et tu entends un bruit fort (une porte qui claque ou l'aspirateur).
Peux-tu dire immédiatement d'où vient le bruit ?



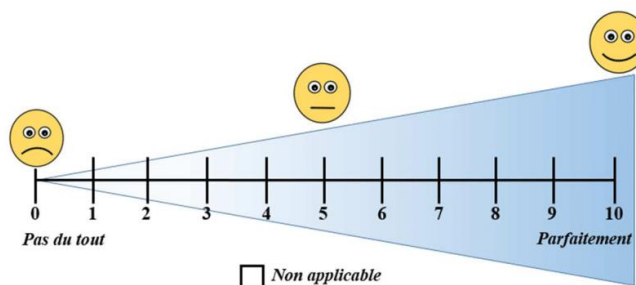
2. Tu es dehors. Un chien aboie très fort.
Peux-tu dire immédiatement où il se trouve, sans le voir ?



3. Tu es sur le trottoir d'une rue où il y a beaucoup de bruit. Tu entends un bus ou un camion.
Peux-tu dire immédiatement d'où il vient avant de le voir ?

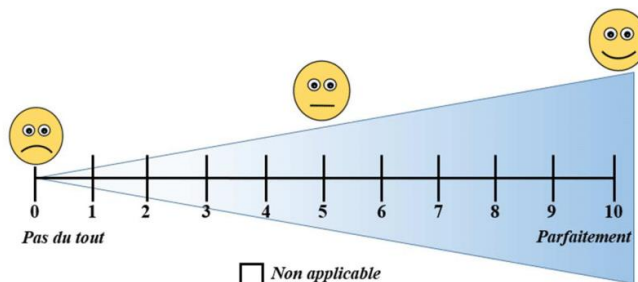


4. Tu entends des voix ou des bruits de pas.
Juste en écoutant, peux-tu dire où la personne va ? (par exemple, est-ce que la personne se rapproche de toi ou se déplace de ta gauche vers ta droite ?)

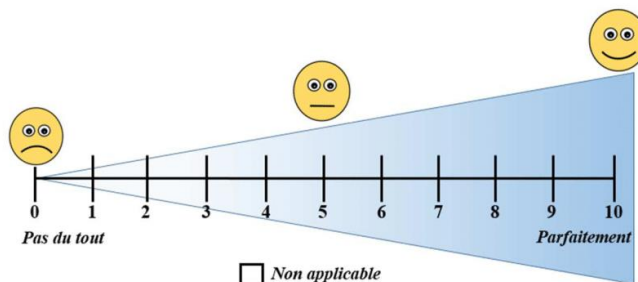


Partie 3 : Qualité d'audition

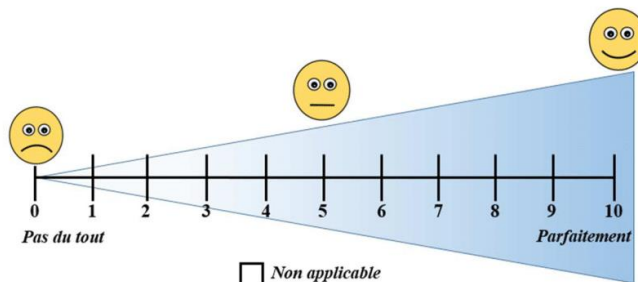
1. **Peux-tu reconnaître facilement la voix des personnes de ta famille, de tes amis, quand tu les entends arriver sans les voir ?**



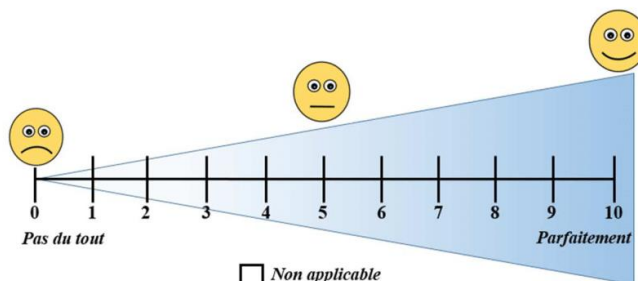
2. **Tu entends une chanson ou une musique que tu connais bien. Arrives-tu facilement à la reconnaître ?**



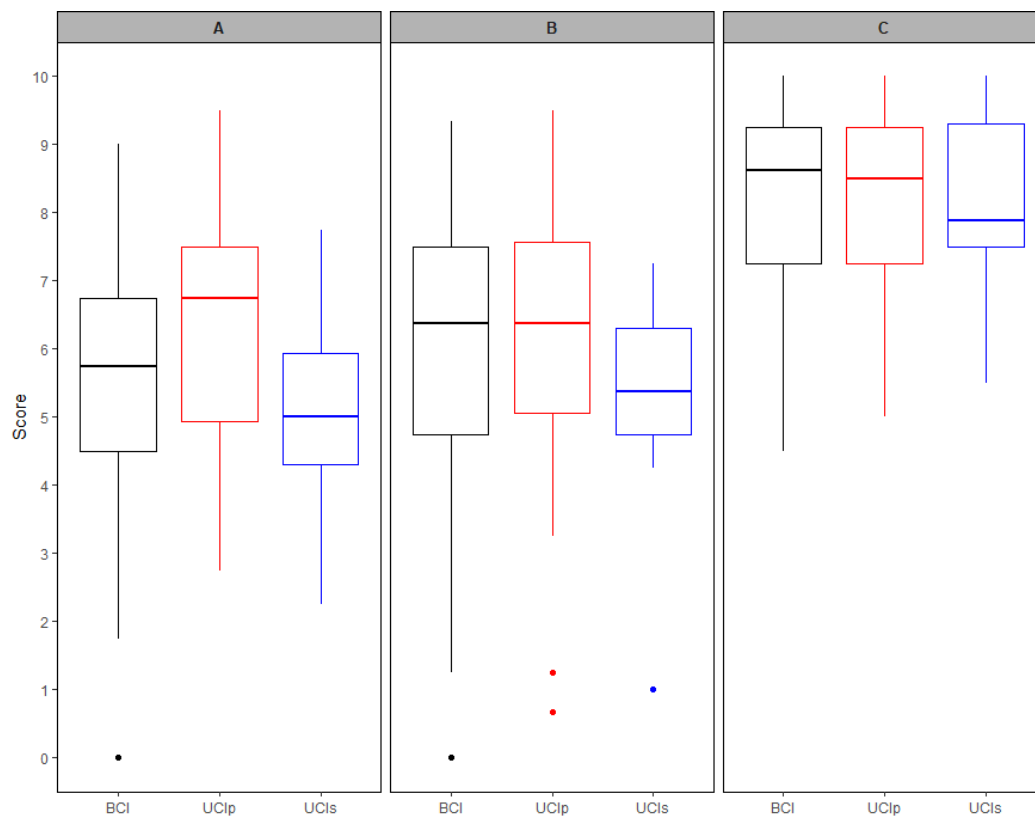
3. **Peux-tu faire la différence entre deux sons qui se ressemblent, comme par exemple un robinet qui coule et un ventilateur ?**



4. **Tous les jours, tu entends les mêmes bruits à la maison. Est-ce que tu peux les reconnaître facilement ? (par exemple, le bruit de la cafetière, la sonnerie du téléphone, l'eau qui coule...)**



Annexe B : Scores moyen par sous-groupe (BCI avec n=100, UCIP avec n=36, UCIs avec n=18) et section (A : audition de la parole ; B : audition spatiale ; C : qualité d'audition)



Les lignes centrales représentent les médianes, les limites des boîtes les 25^e et 75^e percentiles.

Annexe C : Scores moyen par sous-groupe et section.

Sous-groupe	Section A	Section B	Section C
BCI	5,6 (1,8)	6 (2,1)	8,3 (1,4)
UCIP	6,3 (1,7)	6,1 (2,1)	8,1 (1,5)
UCIs	5,2 (1,5)	5,5 (1,5)	8,2 (1,4)

L'écart type est noté entre parenthèses.

Etude 3

RECOVER

Rééducation de l'espace auditif 3D chez des adultes bi-implantés cochléaires

En révision dans Ear and Hearing

Partie 1 : Résumé de l'étude

1. Questions soulevées au regard du contexte scientifique

- Les patients adultes bi-implantés cochléaires (BCI) peuvent-ils adhérer à un protocole d'entraînement spatial intensif dans un environnement contrôlé ?
- Les patients BCI peuvent-ils apprendre à extraire suffisamment d'indices acoustiques enrichis par les mouvements de tête pour construire de nouvelles cartes de spatialisation auditive (SA) ?
- Un entraînement de la SA peut-il impacter d'autres capacités auditives non entraînées ?
- Un feedback multisensoriel est-il plus efficace qu'un feedback unisensoriel pour promouvoir l'amélioration des performances de SA ?

Afin de répondre à ces interrogations, nous proposons dans cette troisième étude de tester la faisabilité d'un protocole d'entraînement multisensoriel de la SA en réalité virtuelle (RV) à destination d'adultes BCI. En objectifs secondaires, nous évaluerons l'impact de cet entraînement sur différentes capacités auditives, spatiales et non spatiales. Puis nous comparerons les effets induits par le feedback multisensoriel versus le feedback unisensoriel.

2. Justification du design de l'étude

Douze patients BCI ont été inclus dans l'étude RECOVER. Ils devaient avoir au moins 1 an d'expérience binaurale avec leurs implants afin de limiter le risque d'interaction avec deux facteurs confondants : les variations importantes des réglages et l'amélioration spontanée des performances de SA la première année post-implantation (Zhang et al. 2015).

Les séances de rééducation se sont étalées sur 1 mois dans le service hospitalier d'implantation cochléaire de Lyon, à raison de 2 séances par semaine de 45 minutes chacune. Ce choix a été fait sur la base du nombre minimal de séances nécessaires pour induire une modification de performances chez des patients (Tyler et al. 2010, Kuk et al. 2014, Firszt et al. 2015) tout en essayant de ne pas saturer leur niveau attentionnel du fait d'une fatigabilité importante induites par les entraînements auditifs. Ces séances étaient encadrées par 5 périodes d'évaluation :

- 2 sessions de tests avant de débiter le protocole d'entraînement pour évaluer les capacités de base et éliminer un éventuel effet intrinsèque des tests sur les performances,
- 1 session à mi-parcours après 4 séances d'entraînement (3 ciblées sur l'espace auditif avant, 1 sur l'espace arrière),

- 2 sessions après la fin du protocole – l’une ayant lieu quelques jours après la fin de la rééducation, pour évaluer les effets induits à court terme, et la dernière 1 mois après pour évaluer le maintien des effets à moyen terme.

Au cours de ces sessions d’évaluation, un test objectif de localisation spatiale en 3D était réalisé avec : (1) l’utilisation de positions de sources sonores différentes de celles entraînées afin d’apprécier la généralisation de la rééducation à l’ensemble de l’espace auditif, et (2) une tâche de réponse motrice différente de celle sollicitée pendant la rééducation afin de tester l’effet sensoriel et non moteur. Nous avons également évalué l’extension des effets de l’entraînement à d’autres capacités auditives *via* deux tests : un test de compréhension de la parole dans le bruit (Framatrix avec 2 sources sonores spatialisées), et un questionnaire de qualité de vie (SSQ dans sa forme courte de 15 questions) incluant notamment des items sur la SA dans la section B.

La structure d’une séance d’entraînement était la même pour l’ensemble du protocole : une phase de recherche dynamique pendant l’émission du son suivie d’un feedback sensoriel. Les intégrations multisensorielles étaient optimisées par la présence constante d’informations visuelles de la phase de recherche au feedback. Ainsi, à chaque essai, les patients étaient immergés en RV dans une pièce structurée reproduisant à l’identique l’ensemble des objets présents dans la pièce réelle, à l’exception du haut-parleur (HP) émettant le son. Une liste de mots (matériel vocal le plus saillant pour le cerveau en rééducation, Verschuur et al., 2005; Grantham et al., 2007) suffisamment longue (15 secondes) était émise afin de laisser le temps au patient d’explorer la scène auditive avec les mouvements de tête. Sa tâche était d’orienter sa tête en direction de la source sonore le plus précisément possible. Une fois la position de la source sonore validée par le sujet, un feedback visuo-auditif ou visuel apparaissait en RV. Nous avons scindé la population en deux groupes afin de tester l’hypothèse suivante : un feedback multisensoriel induit des effets plus importants qu’un feedback unisensoriel. Pour le groupe visuo-auditif (VA), les participants visualisaient l’emplacement exacte du HP tandis que celui-ci continuait d’émettre le son ; pour le groupe visuel (V), seule la position exacte du HP était affichée. A la présentation du feedback, les participants devaient réorienter leur tête sur le HP et tenir cette position corrigée pendant 5 secondes afin d’encoder la position correcte du HP (et du son délivré pour le groupe VA).

3. Résumé de l'analyse des données

Pour le test de localisation 3D, nous avons mesuré le pourcentage moyen de confusions entre les positions des sources sonores avants et arrières (appelées confusions avant-arrière), ainsi que le pourcentage moyen de confusions entre les positions des sources situées au-dessus du niveau des oreilles et celles situées en-dessous (appelées confusions haut-bas), et cela pour chaque participant.

Pour le questionnaire SSQ, nous avons calculé le score moyen pour chaque section et chaque participant ; nous ne présenterons dans ce résumé que le score obtenu à la section B (items relatifs à la spatialisation auditive).

Pour le Framatrix, le score moyen SRT, correspondant au niveau sonore pour lequel les participants arrivaient à répéter 50% du contenu des phrases à niveau de bruit constant, était calculé pour chaque participant.

Les effets de la rééducation ont été mesurés en comparant les performances obtenues à 3 instants différents par rapport aux performances de base (mesurées avant rééducation). Ces 3 moments d'évaluation étaient définis à mi-parcours (après 4 séances d'entraînement), en fin de rééducation (après 8 séances) et à distance de la fin de la prise en charge rééducative (1 mois après la fin de la rééducation).

Enfin, les performances de chaque test ont été comparées en fin de rééducation entre les deux groupes de patients (groupe avec feedback VA et feedback V).

Pour plus de détails sur la méthodologie, voir l'article soumis en partie 3.

4. Principaux résultats

Nous souhaitons tout d'abord tester la faisabilité de notre protocole d'entraînement spatial chez des adultes BCI. Tous les participants sont allés au bout des séances du protocole (i.e. au total 8 entraînements sur 1 mois) et ont montré une bonne adhésion au dispositif de RV. Dans un deuxième temps, nous voulions évaluer les effets de cet entraînement sur différentes capacités auditives spatiales (modifications du pourcentage de confusions avant-arrière et haut-bas au test de localisation 3D et du score moyen à la section B), et non spatiales (changement du score moyen au Framatrix). A mi-parcours, aucun changement notable n'était visible quel que soit le test ; une amélioration significative des performances a été observée en fin

d'entraînement pour l'ensemble des tests d'évaluation. Ainsi, les confusions avant-arrière sont passées de 32% à 14,8% ($p=0,017$) tandis que les confusions haut-bas n'ont pas été modifiées et sont restées proches de la chance (de 49% à 42%, $p=0,088$) à l'échelle de la population. Il faut toutefois noter que 2 patients ont réussi à faire moins de 30% d'erreurs en fin de rééducation. Ce résultat est un exemple de l'importante variabilité existante entre les capacités de spatialisation auditive des patients implantés.

Ces améliorations objectives ont également été perçues par les patients puisque le score moyen de la section B du SSQ a augmenté de 5,3 à 6,6/10 ($p=0,015$). Les bénéfices de l'entraînement spatial se sont étendus vers d'autres compétences de la sphère auditive. Le score moyen SRT a baissé de 1,5 dB à -0,5 dB ($p=0,029$) signant une meilleure compréhension des phrases dans le bruit. L'ensemble des effets bénéfiques observés se sont maintenus 1 mois après la fin de la rééducation : pourcentage de confusions avant-arrière à 14,1% ($p=0,023$), score à la section B à 6,7 ($p=0,032$), et score SRT à -0,7 dB ($p=0,018$).

L'analyse séparée des deux groupes d'entraînement (i.e. feedback VA versus feedback V) n'a mis en évidence aucune différence significative de performances en fin de rééducation quel que soit le test d'évaluation.

En analyse complémentaire, *non présentée dans l'article soumis*, nous avons exploré le degré d'amélioration des confusions avant-arrière en fonction des performances initiales (i.e. (% de confusions avant rééducation - % de confusions après rééducation) / % de confusions avant rééducation) (Figure VII.1). Aucune corrélation graphique n'émerge de cette analyse.

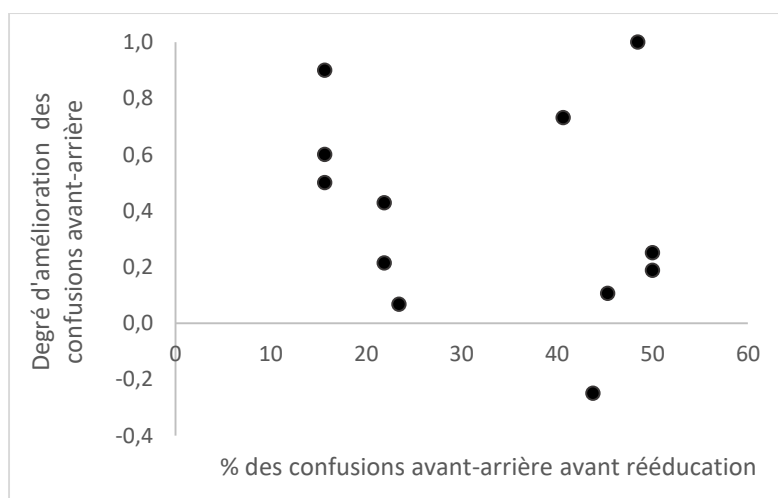


Figure VII.1 : Degré d'amélioration des confusions avant-arrière en fonction des performances individuelles de départ.

Partie 2 : Discussion générale de l'étude RECOVER

1. Bénéfice de l'entraînement centré sur la discrimination en azimut

Malgré l'importante distorsion des indices acoustiques générée par l'implant (e.g. Seeber & Fastl 2008, Dorman et al. 2014), les patients ont amélioré significativement leurs capacités de discrimination avant-arrière. Les sons venant de l'espace arrière étaient initialement perçus en majorité à l'avant, comme dans une étude chez des patients appareillés (Kuk et al. 2014). Dans notre étude, 8 séances de rééducation ont été nécessaires pour qu'une meilleure ségrégation de l'espace sonore apparaisse. Cela pourrait s'expliquer en partie par le fait que l'espace arrière commençait à être travaillé à la séance 4, donc juste avant la session d'évaluation intermédiaire. Notons que l'espace arrière est une zone peu rééduquée en post-opératoire au cours des séances orthophoniques. De plus, afin de faciliter la communication orale avec des locuteurs en face à face, et notamment lors de rééducation, les réglages de l'implant cochléaire sont optimisés (e.g. réduction de tout bruit de fond situé dans l'espace arrière - donc appauvrissement de l'environnement acoustique arrière, microphones directionnels se focalisant sur la parole dans l'espace avant du patient). Ainsi, tous ces éléments suggèrent que l'apprentissage perceptuel de l'espace arrière reste délaissé mais conserve un potentiel qui nécessite certainement plusieurs séances de rééducation avant de rétablir un équilibre spatial sur l'ensemble de l'espace azimutal.

Malgré le fait que plus de la moitié des séances étaient dédiées à l'entraînement de l'élévation, seuls 2 patients sur 11 ont amélioré leurs capacités de discrimination haut-bas. Ce manque de bénéfice avait déjà été constaté par Zahorik (2006) au cours d'un entraînement pour des sujets normo-entendant recevant des indices spectraux dégradés. Ainsi, les sujets diminuaient significativement leurs pourcentages de confusions avant-arrière tandis que les bénéfiques en élévation étaient restreints. L'auteur faisait l'hypothèse que le poids des indices spectraux étant bien plus important dans la perception de l'élévation que de l'azimut (e.g. Bloom, 1977 ; Macpherson & Middlebrooks, 2003), et la recalibration spatiale de cet espace nécessitait un apprentissage beaucoup plus long. De manière analogue, chez les patients IC, les indices spectraux étant majoritairement altérés, il n'est pas surprenant d'observer un effet plus limité de notre rééducation. Il serait donc intéressant dans la suite de cette étude de poursuivre sur une période de temps plus longue la rééducation spatiale (au minimum 2 à 3 mois) afin de permettre à une majorité de patients de mettre en place des mécanismes de compensation

centrale pour compléter leurs cartes spatiales qui, de base, sont probablement très rudimentaires en élévation.

2. Généralisation des effets de l'entraînement spatial

A notre connaissance, RECOVER est la première étude à avoir mis en évidence la généralisation des effets d'un entraînement auditif spatial chez des patients IC. En effet, les modalités de rééducation et d'évaluation étaient différentes : (i) la rééducation était riche en stimulations visuelles (scène visuelle) et auditives saillantes (voix) et complétée par des retours sensoriels (feedback), tandis que les évaluations ne s'appuyaient que sur une stimulation auditive simple (bruit blanc) ; (ii) les positions des sources sonores utilisées lors de la rééducation étaient différentes de celles testées lors des évaluations ; (iii) enfin, pour éliminer un éventuel biais moteur induit par l'entraînement, le type de réponse motrice était aussi différent entre les phases de rééducation et d'évaluation : en rééducation les sujets répondaient en orientant la tête en direction de la cible tandis qu'en évaluation, les sujets devaient indiquer la position de la cible dans leur espace proche à l'aide de leur main.

Nous pouvons donc affirmer qu'il n'y a pas eu un simple apprentissage sensorimoteur de la tâche de spatialisation auditive en rééducation mais un apprentissage d'une nouvelle carte sensorielle auditive qui s'est généralisée à d'autres points de l'espace auditif et qui peut être utilisée indépendamment du type de réponse motrice.

Les résultats présentés pour cette étude ne concernent que les données obtenues lors des tests d'évaluation. A l'échelle individuelle, il serait pertinent d'analyser les comportements d'exploration céphalique au cours des séances de rééducation (e.g. orientation du premier mouvement de tête, temps de latence, etc.) afin de mieux comprendre les mécanismes comportementaux mis en place par les patients IC pour optimiser l'extraction des indices acoustiques. De façon subjective, nous avons constaté que la plupart des mouvements de tête étaient au départ très anarchiques et aléatoires puis au fil des séances, des stratégies se sont développées. Certains patients effectuaient un tilt de la tête vers le haut ou vers le bas afin de mieux percevoir les sons en élévation. Cela revenait à transposer les coordonnées auditives spatiales dans la dimension azimutale et à analyser les indices selon une modalité binaurale et non plus monaurale. Ces constatations sont semblables à celles réalisées par Brimijoin et collaborateurs (2010) chez des patients déficients auditifs. Lors d'une tâche de localisation, les auteurs avaient remarqué que la trajectoire des mouvements de tête des patients (versus sujets

normo-entendant) était beaucoup plus complexe (Figure VII.2) avec des tilts de la tête plus prononcés.

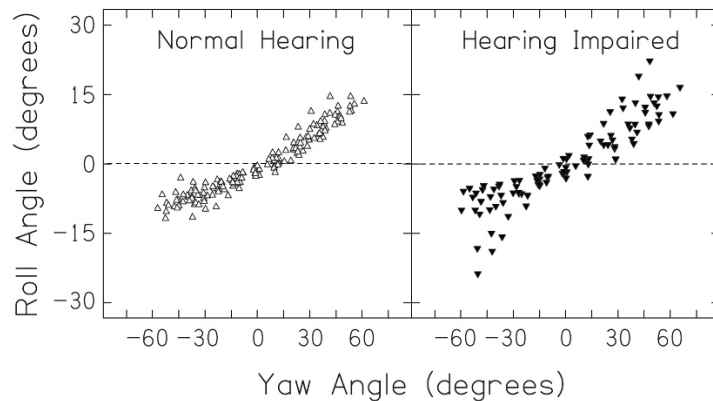


Figure VII.2 : Caractéristiques des mouvements de tête au cours d'une tâche de localisation chez des sujets normo-entendant et des sujets porteurs d'un déficit auditif (d'après Brimijoin et al. 2010). La rotation de la tête (roll) a été analysée en fonction du tilt (yaw) de celle-ci, toute cible confondue.

Dans une prochaine étude, l'analyse des mouvements de tête au cours des séances de rééducation serait peut-être plus sensible pour constater l'évolution des comportements et l'amélioration des performances des sujets.

3. Variabilité des performances

La variabilité inter-individuelle constatée avant et à l'issue du protocole d'entraînement est un point commun à de nombreuses études réalisées chez des patients IC. Les causes de cette variabilité sont multiples (e.g. âge au moment du test, âge d'apparition de la surdité, type de processeur et partie interne, directionnalité des microphones, réducteurs de bruits...) et n'ont pas été un critère de sélection dans l'étude. En effet, nous souhaitons évaluer la faisabilité d'un programme d'entraînement de la SA à destination d'une majorité d'adultes BCI. Il est difficile de dire si certains des critères cliniques sus-cités seraient prédictifs d'un éventuel bénéfice de la rééducation spatiale. Pour l'instant, seule une étude a corrélé le degré d'erreurs en azimut avant rééducation avec le pourcentage d'amélioration de patients porteurs d'une surdité unilatérale profonde non appareillée : plus l'erreur était élevée et plus le patient avait une marge de progression importante (Firstz et al. 2015). Sur le même modèle, nous n'avons retrouvé aucune corrélation entre le degré d'amélioration des confusions avant-arrière et les

performances initiales. Il est possible que le test utilisé ne soit pas suffisamment sensible pour faire émerger les effets de la rééducation sur la SA. Aussi, il serait pertinent de se rapprocher des conditions de la vie réelle en utilisant un test de localisation 3D dans le bruit. Cela permettrait également d'évaluer dans une étude ultérieure les bénéfices de la directionnalité des microphones en comparant un groupe de patients réglés avec un microphone directionnel (i.e. se focalisant sur la source de la parole dans l'espace frontal), versus un groupe de patients contrôles réglés avec un microphone omnidirectionnel (i.e. captant l'ensemble des sons de l'environnement).

4. Poids des interactions multisensorielles dans la rééducation

Lors du développement du protocole de rééducation, nous avons fait le choix de mettre la multisensorialité au cœur des séances d'entraînement car la SA est par essence un processus multisensoriel (e.g. Burr & Alais 2006, Recanzone 2009). Le groupe VA recevait des informations visuelles et auditives congruentes, de la phase de recherche du son jusqu'à la phase de feedback sensoriel. Ainsi, pendant l'émission du son, les patients recherchaient l'origine de la source en bougeant la tête ce qui leur permettait d'exploiter de manière concomitante les informations provenant de leur environnement visuel (i.e. scène de réalité virtuelle) et les informations acoustiques dynamiques induites par les mouvements de tête. Puis lors de la phase de feedback, ils visualisaient la position exacte du HP tandis que le son continuait d'être émis. A l'opposé, pour le groupe V, les interactions multisensorielles se limitaient à la phase de recherche du son puisque le feedback permettait de visualiser la position du HP sans information auditive associée.

Nous avons fait l'hypothèse que le feedback multisensoriel conduirait à une amélioration plus importante des performances que le feedback unisensoriel visuel. Notre hypothèse ne s'est pas vérifiée, aucune différence de performance n'a émergé entre les deux groupes de rééducation. Il est possible que la seule recherche de la source sonore dans un environnement visuel structuré puisse favoriser la SA et que le feedback final n'ait qu'un poids mineur dans ce contexte rééducatif. En effet, plusieurs auteurs ont montré que la moindre information visuelle disponible aidait le cerveau à calibrer les indices auditifs pour établir une carte spatiale de l'environnement (Isaiah et al. 2014, Ahrens et al. 2019, Valzolgher et al. 2020b) et qu'en cas de conflit visuo-auditif les informations visuelles aidaient au décodage et au maintien à long terme des informations auditives (Weinberger 2004, Schorr et al. 2005). De plus, l'équipe de Cappelloni (2019) a montré que lors d'une tâche de discrimination spatiale auditive, les

performances des sujets étaient améliorées en situation visuo-auditive, alors même que les indices visuels n'étaient pas pertinents pour la tâche. Dans le cas présent, l'environnement visuel de la pièce en RV est resté constant tout au long des séances ce qui a permis à tous les patients d'avoir un ancrage visuel permanent. Il serait intéressant dans la suite de cette étude de rajouter un groupe contrôle qui participerait au programme d'entraînement sans recevoir de feedback final. Cela permettrait d'évaluer si l'interaction multisensorielle qui a lieu pendant la phase de recherche se suffit à elle-même ou est potentialisée par l'ajout d'un feedback final.

La présence d'un autre groupe contrôle n'effectuant pas l'entraînement mais uniquement les sessions d'évaluation permettrait d'affirmer que les effets obtenus ne sont pas imputables à un apprentissage des tests au fil des sessions. Pour autant, nous faisons l'hypothèse que si cet effet existe il reste négligeable comparé à celui induit par l'entraînement. Plusieurs études ont montré la stabilité des performances de SA en l'absence d'entraînement chez des sujets normo-entendant (Steadman et al. 2019) et surtout chez des patients appareillés (Kuk et al. 2014) et implantés (Tyler et al. 2010).

5. Conclusion et perspectives

Cette étude a montré en premier lieu la faisabilité d'un programme d'entraînement de la spatialisation auditive chez des adultes implantés cochléaires sur une période prolongée. Malgré un accès très limité aux indices acoustiques, l'intégration d'informations multisensorielles a certainement joué un rôle clé dans la modification des performances des patients. Aussi, devant les retours positifs des patients et les résultats encourageants de cette étude pilote, nous nous interrogeons sur l'opportunité de proposer ce type de protocole plus précocement dans le parcours de soins des implantés, et de complexifier la rééducation (e.g. localisation auditive dans un environnement bruyant). Plusieurs auteurs ont ainsi suggéré que l'entraînement précoce permettait de potentialiser les effets naturels induits par l'expérience auditive du quotidien et aiderait à l'adaptation du décalage inter-aural induit par les deux IC. Il ne serait peut-être donc pas nécessaire d'attendre la stabilisation des performances de localisation après plus d'un an, mais commencer l'entraînement spatial en parallèle de la rééducation langagière dans le bruit. Cela faciliterait la ségrégation des différentes sources sonores et *in fine* leur reconnaissance.

Partie 3 : Article relatif à l'étude RECOVER (soumis dans *Ear and Hearing* le 01/08/2021, en cours de révision)

Intensive training of spatial hearing promotes auditory abilities of bilateral cochlear implant adults: a pilot study

Aurélie COUDERT^{1,3,4,5}, Grégoire VERDELET^{1,6}, Karen T. REILLY^{2,5}, Eric TRUY^{1,3,4,5}, Valérie GAVEAU^{1,5}

1. Integrative Multisensory Perception Action & Cognition Team — ImpAct, Lyon Neuroscience Research Center, INSERM U1028, CNRS U5292, Lyon, France
2. Trajectoires Team — Lyon Neuroscience Research Center, INSERM U1028, CNRS U5292, Lyon, France
3. Department of Pediatric Otolaryngology – Head & Neck Surgery, Femme Mere Enfant Hospital - Hospices Civils de Lyon, Lyon, France
4. Department of Otolaryngology – Head & Neck Surgery, Edouard Herriot Hospital - Hospices Civils de Lyon, Lyon, France
5. University of Lyon 1, Lyon, France
6. Hospices Civils de Lyon, Neuro-immersion Platform, Lyon, France

All correspondence should be addressed to:

Aurélie Coudert, Department of Pediatric Otolaryngology – Head & Neck Surgery, Femme Mere Enfant Hospital - Hospices Civils de Lyon, 59 boulevard Pinel, 69677 Bron France.

Email: aurelie.coudert2@chu-lyon.fr

ABSTRACT

Objective: The aim of this study was to evaluate the feasibility of virtual-reality-based spatial hearing training 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 of 8 virtual reality training sessions (2 per week) interspersed with 5 evaluation sessions (2 weeks before training started, after 4 and 8 training sessions, and 1 month after the end of training). During each 45-minute training session patients had to localize a sound source whose position varied in azimuth and/or in elevation (depending on session difficulty). At the start of each trial CI users received no information about sound location, but after validating their answer feedback was given to enable error correction. Patients were divided into two groups: a multisensory feedback group (audiovisual spatial cue) and an unisensory one (spatial cue) who received only feedback in a wholly intact sensory modality. Training benefits were measured at each evaluation using 3 auditory 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 by the patients and all patients 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% ($p=.017$); SRT-score from 1.5 dB to -0.7 dB ($p=.029$) and 8 CI users successfully achieved a negative signal to noise ratio. Quality of life also improved for sound localization (from 5.3 to 6.7, $p=.015$), and for speech understanding (from 5.2 to 5.9, $p=.048$) but with a one month delay. All benefits were maintained 1 month after the end of training.

Conclusions: This pilot study shows the feasibility and potential clinical interest 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: spatial hearing, rehabilitation, cochlear implant, 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

INTRODUCTION

Accurate analysis of the auditory scene is essential for orienting attention in space and necessary for daily behaviors (Kerber and Seeber 2012; Shin-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 3D, i.e. 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 auditory cues reaching each ear: binaural and monaural cues. 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 pinna, head and to a lesser extent by the shoulders depending on the incidental angle and the sound frequency (Angell 1901a, 1901b). 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 1999a; 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, Brimijoin et al. 2012). In 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 audio-visual association is essential in complex multisensory environments (Da Silva, 1985; Loomis et al., 1998; Bologni et al. 2005), and many studies have highlighted its benefits in terms of sound localization (Strelnikov et al. 2011) and speech discrimination (e.g. Middleweerd 1987, Schwartz et al. 2004, Macleod et al. 2009).

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 comfort 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 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 2003, Grantham 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 and 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 two years of binaural hearing experience (Zhang 2015). 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. 2020b, Coudert et al. 2021). Furthermore, since the position of the CI microphone 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 et al. 2015). These types of training protocols are most commonly proposed to hearing aid users (Kuk et al. 2014, Firszt et al. 2015), but have also been tested with 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 focusses 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 (Isaiah et al. 2015). There are, however, several methodological challenges that must be overcome in order to sufficiently control the proposed training sessions. First, since spatial hearing is three-dimensional, it is essential to work in the whole space surrounding the patient, and not only in front space in 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.

In order 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; Valzolgher et al. 2020b; 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 with 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 in order 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 and therefore visual information was accurately processed. The second group received audiovisual feedback. This group received the same spatial cue as in the first group plus audio feedback. Based on evidence suggesting that multisensory training is better than unisensory (Strelnikov et al. 2011), we hypothesized that the improvement would be greater in the group that received audiovisual feedback.

MATERIAL AND METHOD

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 patients gave informed consent for inclusion before participating in the study.

Patients

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 binaural experience (to avoid large setting variations of cochlear implants), and no areflexia. Binaural 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 patients we did not target a specific patient profile. As such, no inclusion criteria regarding CI processor brand or settings were applied. The bilateral CI patients were all daily users of their cochlear implants and had excellent monosyllabic word recognition performance (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, patients wore their own cochlear implant processors and no parameter adjustments were made before testing. Since neither the training nor evaluation conditions

included any background noise those patients with directional microphones used omnidirectional settings.

Experimental Protocol

The sound localization rehabilitation protocol was conducted across 10 weeks and consisted of 8 training sessions interspersed with evaluation sessions (Figure 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 4 (E2) and 8 training sessions (E3), and one month after the 8th training session (E4).

All sessions (training and evaluation) took place in a reverberant room (3.6m x 3.9m x 2.7m, reverberation time RT60: 0.32s) within a hospital. All patients 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 patients 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 × 82.7 × 30.8 mm, Output Power 3.0 W; frequency response: 180 Hz – 20 kHz) that was moved in space by the experimenter. For each sound localization trial, the experimenter silently moved the loudspeaker to the desired sound position. Further details of the experimental set-up 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; Valzolgher et al. 2020b; Coudert et al. 2021). Our VR system consisted of a head-mounted display-HMD (HTC VIVE® System, resolution: 1080 x 1200 px, Field Of View (FOV): 110°, Refresh rate: 90 Hz) with integrated eye-tracking technology (SensoriMotoric Instruments, Berlin, Germany; www.smivision.com; 60 Hz frequency and 0.5 degrees spatial precision). The main advantage of using the virtual reality apparatus for training sessions was (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 hand-held VIVE® controller that patients 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). Tested positions were predetermined in advance 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 patients responded by orienting their head in the direction of the sound; During evaluation, stimuli were 3-second white noise bursts and patients indicated sound sources using a hand-held 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 (Figure 1). During training, patients wore a 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 meter from the head). 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 patients did not habituate to the words. During training the patient'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 90cm, and depending on the training session its position varied in azimuth and/or elevation (see Figure 2).

At the beginning of each session patients were informed of the portion of space that would be trained. Tested stimulus locations and session difficulty were varied across training sessions in order to maintain motivation and to challenge patients throughout the training protocol. The level of difficulty was increased across the 8 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 patients 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 were 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 patients 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. To familiarize themselves with the apparatus, patients underwent a 5-minute training session during which they had to localize sounds from 10 different possible loudspeaker positions that could be anywhere around them - similar to the T7/T8 training sessions (see Figure 2, panel E).

Each training trial was divided into 2 phases: a sound localization phase and a feedback phase. During the sound localization phase the experimenter positioned the loudspeaker in a predetermined position (unknown to the patient) and a sound was emitted for 15 seconds. The patient's task was to identify the direction from which the sound was emitted by orienting their head to the perceived sound source location. They then validated their head position (indicated by a white cross in the HMD - see videos in supplementary material) by clicking a button on a hand-held VIVE controller. This initiated the feedback phase during which they received immediate feedback on their performance.

If the patient did not respond during the 15-second trial, feedback appeared automatically at the end of the trial. Half of the patients (group A) received both visual and auditory feedback, with an image of the loudspeaker appearing in virtual reality in a position spatially congruent with its actual position in space while the sound stimulus continued to be played 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). We chose to investigate the added value of providing of multisensory information 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), and because vision was intact in all patients.

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 size of the arrow varied as a function of their error (larger 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 supplementary material.

Patients 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 one 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

The 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. We chose this questionnaire because (1) it has been validated in French (Moulin et al. 2018) and (2) it contains specific questions on sound localization and speech comprehension in noise. The short form of the SSQ has 15 items, divided into 3 domains: subscale A for speech perception; subscale B for spatial hearing; and subscale C for other qualities of hearing. Each question is scored from 0 (not at all) to 10 (perfectly) and higher scores represent greater comfort 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, patients wore the HMD but received no visual information (the screen was grey). Patients were told that sounds would be played from various positions around their body and that their task was to place a hand-held pointer at the exact position at which they perceived the sound. They were also told they could reach each sound source without leaving their chair. Unbeknownst to them, 8 pre-determined sound positions at a constant distance of 55cm from

the center of the head (within reaching space) were used. The loudspeaker could be located at +30°, +70°, +120°, +160°, -30°, -70°, -120°, and -160° in azimuth with respect to the patient's straight ahead (see Figure 3, positive values indicating right space and negative values left space). Two elevations were evaluated; +25° and -25°. Each position was repeated eight times, for a total of 64 trials. It is important to note that all of the tested 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, modulated in amplitude at 2.5 Hz frequency (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 (4.1m x 2.8m x 2.7m, reverberation time RT60: 0.39s). Two loudspeakers were placed at ear level, at -45° and +45° azimuth angle, and at 1 meter from the patient'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 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 patient's 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 patient we analyzed performance during each of the 4 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. Based on previous studies on sound localization, performance at chance level was defined as a score of 41% or more confusions (Mueller et al. 2013; Fischer et al. 2020b). 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 domain (A, B and C), as well as the mean total SSQ-score across the 3 domains.

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 data were not normally distributed we used non-parametric 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 patients attended all training and evaluation 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 Figure 1). Three separate Wilcoxon signed-rank tests failed to identify significant performance differences between E0 and E1 on any of the tests: sound localization test ($V=29.5$, $p=.121$), matrix test ($V=56$, $p=.051$); SSQ questionnaire ($V=40$, $p=.563$). These results suggest that the content of the evaluation sessions did not induce any learning. For all further 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 patients, for all 8 sound sources separately for each evaluation session. Each colored symbol represents the average hand pointing position for a given patient 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 inter-aural axis (around their cochlear implants) and in front space. This resulted in a large number of front-back confusions (median of 32.0%, E1, Figure 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, Figure 5B). Figure 6 shows median SRT scores on the matrix test. Before training the median SRT was 1.5 dB and between-patient variability was very high (range -4.7 to 6.8). The median for the total SSQ score was 5.9/10.

Auditory performance before 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 four points was compared to baseline performance obtained during the pre-training 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, patients' performance for back sound sources progressively improved, with the median percentage of front-back confusion decreasing from 25.8% during E2 to 14.8% during E3 and stabilizing at 14.1% during E4 (Figure 5A). The group-level improvement at E3 was driven by 8/12 bilateral CI users all of whom had fewer than 25% of confusions, although it is important to note that at E4 9/12 patients 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 ($\chi^2(3)=10.9$, $p=.012$, effect size $W=0.31$). Pairwise comparisons (Wilcoxon test with Bonferroni adjustment) showed that this improvement was significant between E1 and E3 ($p=.017$) and between E1 and E4 ($p=.023$) but not between E1 and E2 ($p=.169$).

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 remained 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 confusion across the four evaluation sessions (from 46.1% during E2 to 42.2% during E4). Interestingly, even if the median of up-down confusions at E4 was similar to that at E3, several patients continued to improve between E3 and E4 and the improvements in 2 patients (A01 and B03) put them below chance level (see

Table 2). A Friedman test revealed no significant differences between the four evaluation sessions ($\chi^2(3)=6.54$, $p=.088$).

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 above 40% (below chance level) at all evaluation sessions.

One patient (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 4 evaluation sessions for 11/12 patients. The median SRT score decreased from 1.5 dB at E1 to 0.2 dB at E2 and -0.5 dB at E3 then stabilized at -0.7 dB 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-patient variability across sessions. SRT scores ranged from 6.8 to -4.7 at E1 and from 1.3 to -4.8 at E4. SRT scores improved between E1 and E2 for 8/11 patients and between E2 and E3 for 10/11 patients. Some patients continued to improve between E3 and E4, although these improvements were generally small and were observed in only 5/11 patients (see Table 3 for individual data). A Friedman test revealed a significant decrease in SRT scores across all 4 evaluation sessions ($\chi^2(3)=12.6$, $p=.006$, 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 ($p=.194$). It is important to note that at the end of training eight bilateral CI users were able to discriminate sentences with a *negative* signal to noise ratio, meaning that they were able to correctly repeat sentences with a higher noise than 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 4 evaluation sessions ($\chi^2(3)=22.3$, $p<.001$, effect size $W=0.62$). Pairwise comparisons (Wilcoxon test with Bonferroni adjustment) showed that this improvement was significant between E1 and E3 ($p=.003$), and between E1 and E4 ($p=.015$) but not between E1 and E2 ($p=.205$). To investigate whether this improvement concerned all three domains of the SSQ-questionnaire we compared the scores on each subscale across all 4 evaluation sessions (Figure 7). The median scores of subscales A and B at E1 were similar (5.2 and 5.3) whereas the score for subscale C was higher (7.9). Subscale A 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, $\chi^2(3)=8.78$, $p=.032$, effect size $W=0.24$), but only between E1 and E4 ($p=.048$). Subscale B scores steadily increased over the 4 evaluations; 5.8, 6.6, and 6.7 from E2 to E4. This improvement was significant (Friedman test, $\chi^2(3)=14.7$, $p=.002$, effect size $W=0.41$) between E1 and E3 ($p=.015$), and E1 and E4 ($p=.032$). Subscale C remained constant around 8 throughout the evaluations ($\chi^2(3)=6.85$, $p=.077$). It is interesting to note that the individual patient data in Table 4 reveal substantial variability in SSQ scores. Some patients reported no difficulties in their daily lives in any of the domains (e.g. B03, B05) whereas others still reported sound localization problems after the training (e.g. A02, A06, B01).

Overall the questionnaire revealed that quality of life improved substantially with training, that this improvement was due to changes in 2 of the 3 subdomains, and was still present 4 weeks after the end of training.

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 all patients, regardless of the type of feedback they received. Despite the small number of patients in each group ($n=6$), we were also interested to know whether there was an added benefit of training for those patients 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 ($V=13$, $p=.466$), the matrix test ($V=22$, $p=.247$) or the SSQ questionnaire ($V=17$, $p=.937$). We then compared performance in the two groups at E3 at the end of the 8 training sessions. Similar to E1, no significant differences on any of the evaluation tests emerged between groups at E3 (Mann–Whitney tests): the sound localization test ($V=21$, $p=.699$), the matrix test ($V=20$, $p=.429$) and the SSQ questionnaire ($V=15.5$, $p=.748$). Although the sample size in each group is small, these data suggest that compared with visual feedback alone, there is no added benefit to 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 the 8 biweekly training sessions. Development of the training sessions

was based on three observations for which there is increasing clinical and scientific support: i) spatial hearing is a multisensory process, ii) exploration of the auditory environment with head movements improves sound localization, and iii) 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 and we evaluated performance on two auditory tests and 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 all three evaluations improved after eight training sessions spread across 4 weeks. It is important to note that all patients benefited from the training, regardless of their clinical profile, CI device brand or setting, or duration of binaural experience. Interestingly, the patients with the greatest improvement were not necessarily those with the best performance before training, suggesting that this protocol can be offered to all patients regardless of their pre-inclusion performance.

Until now, most spatial hearing protocols have been developed with normally hearing (NH) adults with little homogeneity between protocols and poor quality control: 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 and colleagues were one of the first groups to propose training sessions spread over several days, and this was done with NH adults wearing a monaural plug (Strelnikov et al. 2011). Studies conducted with hearing impaired people are rare, with only 3 published studies of note. These included either patients with unilateral hearing loss (Firstz 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

Spatial hearing improvement was most noticeable for front-back confusions, which decreased by more than 50% (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.

Despite this improvement, front-back discrimination remained difficult for CI users partly due to the lack of salient 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), patients in our sample fitted with this technology did not show the most improvement in sound localization performance. Bilateral CI users also experience difficulties extracting ILD cues because the magnitude of ILDs are 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 mostly explain the presence of large sound localization errors in azimuth in bilateral CI users (e.g. Kerber & Seeber 2012). 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. 2013, Pastore et al. 2018, Fischer et al. 2020b, Coudert et al. 2021) and in the present study these ILD variations could have been large enough to be usable by the bilateral CI users. This reinforces the importance of focusing auditory rehabilitation on the dynamic interaction between the two CIs that occurs naturally in everyday behavior in which the head moves freely.

There is little data available on sound localization in elevation in CI users, but one study by Majdak and colleagues (2011) reported results similar to our pre-training 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 and 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 placed behind the ear 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 patients

improved their up-down discrimination after training, while the other 10 remained at chance level. The improvement shown by these two patients 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 and colleagues that sound localization in elevation is possible from low spectral cues when sound sources are positioned laterally (Algazi et al. 2001).

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 after 8 training sessions, meaning that after training most patients 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 clinically significant. Some patients even managed to reach the mean score of NH peers, i.e. -6 dB (Jansen et al. 2012). All patients reported tiredness after the 50 sentences and 1 patient 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 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, Lorenzi et al. 2016). Improvement on the Matrix test after spatial localization training could be due to the fact that patients learned to better exploit spatial hearing cues necessary for segregating speech from noise which in turn facilitated their speech understanding. Indeed, the large inter-individual 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 patients' daily life using a validated quality of life questionnaire, the short form of the SSQ

(Moulin et al. 2018). 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 to a lesser extent 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 two years after surgery (Hasspass et al. 2013, Zhang 2015), and no data exist at longer delays, when performance and device settings are stable. Interestingly, when compared to the results of a large cohort of patients suffering from moderate hearing loss, i.e. a maximum loss of 55dB HL (Moulin et al. 2018), our mean scores per domain were similar (subscale A: 6.7, subscale B: 6.5, and subscale C: 8.2). This suggests that our rehabilitation protocol allowed bilateral CI users to reach a similar quality of life to that of 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 continued to improve between the end of training and the 1-month follow-up evaluation. Previous studies of spatial hearing rehabilitation did not include a follow-up evaluation of the benefits in hearing impaired patients (Kuk et al. 2014, Firstz 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 patients maintain their performance because they indirectly continue to train using the multi-sensory 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 multi-sensory) influenced training improvement. The multi-sensory 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 et al. 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. 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 observation 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 to 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 the results of an imaging study, it is possible that the absence of a difference between the two groups is due to the predominance of the visual system in hearing impaired patients. Indeed, 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. Third, the small size of each group may have led to a lack of statistical power to highlight any differences between groups. 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 binaural 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.

Perspectives

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 binaural 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. Considering that hearing impaired patients routinely face challenging situations when attempting to understand speech (e.g. interfering background noise, competing speakers, 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 5 evaluation sessions across the 10-week study duration. The ease of use of the virtual reality system regardless of the patients' 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.

Author contributions: 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.

Funding: 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). We thank the IMPACT team administrative staff for their administrative and technical support.

Acknowledgments: The authors are grateful to all patients who took part in this study.

Conflicts 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.

REFERENCES

- Algazi, V. R., Avendano, C., & Duda, R. O. (2001). Elevation localization and head-related transfer function analysis at low frequencies. *The Journal of the Acoustical Society of America*, 109(3), 1110-1122. <https://doi.org/10.1121/1.1349185>
- Anderson, S., & Kraus, N. (2010). Sensory-cognitive interaction in the neural encoding of speech in noise: A review. *Journal of the American Academy of Audiology*, 21(9), 575-585. <https://doi.org/10.3766/jaaa.21.9.3>
- Angell JR, Fite W. (1901a). The monaural localization of sound. *Psychol Rev*, 8:225-46, <http://dx.doi.org/10.1037/h0073690>.
- Angell JR, Fite W. (1901b). Further observations on the monaural localization of sound. *Psychol Rev*, 8:449-58, <http://dx.doi.org/10.1037/h0075258>.
- 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. *The Journal of the Acoustical Society of America*, 127(3), EL87-92. <https://doi.org/10.1121/1.3298451>
- 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. *Frontiers in Neurology*, 11, 915. <https://doi.org/10.3389/fneur.2020.00915>
- Bolognini, N., Frassinetti, F., Serino, A., & Làdavvas, E. (2005). « Acoustical vision » of below threshold stimuli : Interaction among spatially converging audiovisual inputs. *Experimental Brain Research*, 160(3), 273-282. <https://doi.org/10.1007/s00221-004-2005-z>
- 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>
- 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>
- Brungart, D. S. (1999). Auditory localization of nearby sources. III. Stimulus effects. *The Journal of the Acoustical Society of America*, 106(6), 3589-3602. <https://doi.org/10.1121/1.428212>

Bulkin, D. A., & Groh, J. M. (2006). Seeing sounds : Visual and auditory interactions in the brain. *Current Opinion in Neurobiology*, 16(4), 415-419. <https://doi.org/10.1016/j.conb.2006.06.008>

Coudert, A., Gaveau, V., Gatel, J., Verdelet, G., Salemme, R., Farnè, A., Pavani, F., Truy, E. (2021). Spatial hearing difficulties in reaching space in bilateral cochlear implant children improve with head movements. *Ear Hear.* Jul 1. doi: 10.1097/AUD.0000000000001090. Online ahead of print.

Da Silva, J. A. (1985). Scales for perceived egocentric distance in a large open field: Comparison of three psychophysical methods. *The American Journal of Psychology*, 98(1), 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 and Hearing*, 39(4), 679-686. <https://doi.org/10.1097/AUD.0000000000000525>

Dorman, M. F., Loisel, 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 and Hearing*, 35(6), 633-640. <https://doi.org/10.1097/AUD.0000000000000057>

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. *Hearing research*, 319, 48-55. <https://doi.org/10.1016/j.heares.2014.11.005>

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(1), 214-222. <https://doi.org/10.1097/AUD.0000000000000912>

Gatehouse, S., & Noble, W. (2004). The Speech, Spatial and Qualities of Hearing Scale (SSQ). *International journal of audiology*, 43(2), 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(3), 657-663. [https://doi.org/10.1016/s0896-6273\(01\)00318-x](https://doi.org/10.1016/s0896-6273(01)00318-x)

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

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. *Otology & Neurotology: Official Publication of the American Otological Society, American Neurotology Society [and] European Academy of Otology and Neurotology*, 34(7), 1278-1283. <https://doi.org/10.1097/MAO.0b013e3182937ad4>

Isaiah, A., & Hartley, D. E. H. (2015). Can training extend current guidelines for cochlear implant candidacy? *Neural Regeneration Research*, 10(5), 718-720. <https://doi.org/10.4103/1673-5374.156964>

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>

Kerber, S., & Seeber, B. U. (2012). Sound localization in noise by normal-hearing listeners and cochlear implant users. *Ear and Hearing*, 33(4), 445-457. <https://doi.org/10.1097/AUD.0b013e318257607b>

Kuk, F., Keenan, D. M., Lau, C., Crose, B., & Schumacher, J. (2014). Evaluation of a Localization Training Program for Hearing Impaired Listeners: *Ear and Hearing*, 35(6), 652-666. <https://doi.org/10.1097/AUD.0000000000000067>

Laback, B., Egger, K., & Majdak, P. (2015). Perception and coding of interaural time differences with bilateral cochlear implants. *Hearing Research*, 322, 138-150. <https://doi.org/10.1016/j.heares.2014.10.004>

Lafon, J.C. (1964). *Le Test Phonétique Et La Mesure De L'Audition*, Editions 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. *Perception & Psychophysics*, 60(6), 966-980. <https://doi.org/10.3758/bf03211932>

MacLeod, A., & Summerfield, Q. (1987). Quantifying the contribution of vision to speech perception in noise. *British Journal of Audiology*, 21(2), 131-141. <https://doi.org/10.3109/03005368709077786>

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

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>

Middelweerd, M. J., & Plomp, R. (1987). The effect of speechreading on the speech-reception threshold of sentences in noise. *The Journal of the Acoustical Society of America*, 82(6), 2145-2147. <https://doi.org/10.1121/1.395659>

Middlebrooks, J. C. (2015). Sound localization. *Handbook of Clinical Neurology*, 129, 99-116. <https://doi.org/10.1016/B978-0-444-62630-1.00006-8>

Mo, B., Lindbaek, M., & Harris, S. (2005). Cochlear implants and quality of life : A prospective study. *Ear and Hearing*, 26(2), 186-194. <https://doi.org/10.1097/00003446-200504000-00006>

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. *Journal of the Association for Research in Otolaryngology: JARO*, 9(4), 399-406. <https://doi.org/10.1007/s10162-008-0143-x>

Moulin, A., Vergne, J., Gallego, S., & Micheyl, C. (2018). A New Speech, Spatial, and Qualities of Hearing Scale Short-Form : Factor, Cluster, and Comparative Analyses. *Ear and Hearing*. <https://doi.org/10.1097/AUD.0000000000000675>

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 International*, 15(1), 36-42.

Musicant, A. D., & Butler, R. A. (1984). The influence of pinnae-based spectral cues on sound localization. *The Journal of the Acoustical Society of America*, 75(4), 1195-1200.

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. *The Journal of the Acoustical Society of America*, 146(1), 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 and Hearing*, Publish Ahead of Print. <https://doi.org/10.1097/AUD.0000000000000581>

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>

Perrett, S., & Noble, W. (1997). The effect of head rotations on vertical plane sound localization. *The Journal of the Acoustical Society of America*, 102(4), 2325-2332.

Rouger, J., Lagleyre, S., Fraysse, B., Deneve, S., Deguine, O., & Barone, P. (2007). Evidence that cochlear-implanted deaf patients are better multisensory integrators. *Proceedings of the National Academy of Sciences of the United States of America*, 104(17), 7295-7300. <https://doi.org/10.1073/pnas.0609419104>

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 and Hearing*, 36(5), 582-590. <https://doi.org/10.1097/AUD.0000000000000170>

Schwartz, J.-L., Berthommier, F., & Savariaux, C. (2004). Seeing to hear better : Evidence for early audio-visual interactions in speech identification. *Cognition*, 93(2), B69-78. <https://doi.org/10.1016/j.cognition.2004.01.006>

Seeber, B. U., Baumann, U., & Fastl, H. (2004). Localization ability with bimodal hearing aids and bilateral cochlear implants. *The Journal of the Acoustical Society of America*, 116(3), 1698-1709. <https://doi.org/10.1121/1.1776192>

Shams, L., Wozny, D. R., Kim, R., & Seitz, A. (2011). Influences of Multisensory Experience on Subsequent Unisensory Processing. *Frontiers in Psychology*, 2. <https://doi.org/10.3389/fpsyg.2011.00264>

Shinn-Cunningham, B., Best, V., & Lee, A. K. (2017). Auditory object formation and selection. *The auditory system at the cocktail party*, New York: Springer, 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. M., Briaire, 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 Otolaryngology-- Head & Neck Surgery*, 142(3), 249-256. <https://doi.org/10.1001/jamaoto.2015.3305>

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. *Scientific Reports*, 9(1), 18284. <https://doi.org/10.1038/s41598-019-54811-w>

Strelnikov, K., Rosito, M., & Barone, P. (2011). Effect of Audiovisual Training on Monaural Spatial Hearing in Horizontal Plane. *PLoS ONE*, 6(3). <https://doi.org/10.1371/journal.pone.0018344>

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. *Journal of the American Academy of Audiology*, 21(6), 390-403. <https://doi.org/10.3766/jaaa.21.6.4>

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. <https://doi.org/10.1016/j.neuropsychologia.2020.107665>

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. *Current Research in Behavioral Sciences* 1, 100003. <https://doi.org/10.1016/j.crbeha.2020.100003>

van Hoesel, R. J. M. (2004). Exploring the benefits of bilateral cochlear implants. *Audiology & Neuro-Otology*, 9(4), 234-246. <https://doi.org/10.1159/000078393>

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>

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. The Role of Head Movements and Vestibular and Visual Cues in Sound Localization. (1940). *Journal of Experimental Psychology*, 27 (4): 339-68. <https://doi.org/10.1037/h0054629>

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.

Witten, I. B., & Knudsen, E. I. (2005). Why seeing is believing : Merging auditory and visual worlds. *Neuron*, 48(3), 489-496. <https://doi.org/10.1016/j.neuron.2005.10.020>

Zaltz, Y., Buganim, Y., Zechoval, D., Kishon-Rabin, L., & Perez, R. (2020). Listening in Noise Remains a Significant Challenge for Cochlear Implant Users : Evidence from Early Deafened and Those with Progressive Hearing Loss Compared to Peers with Normal Hearing. *Journal of Clinical Medicine*, 9(5). <https://doi.org/10.3390/jcm9051381>

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. *American Journal of Audiology*, 24(3), 384-397. https://doi.org/10.1044/2015_AJA-14-0074

FIGURES

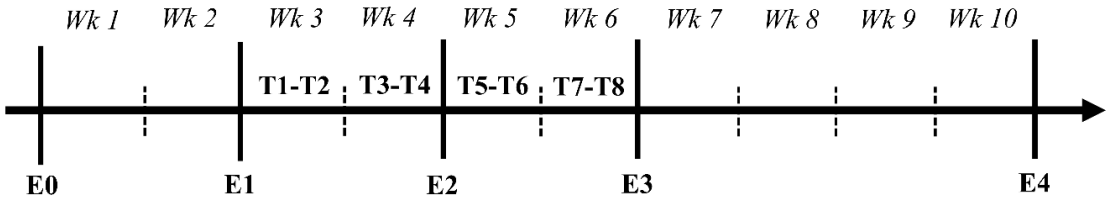


Figure 1. Flowchart of the rehabilitation training protocol conducted over 10 weeks (Wk 1 to Wk 10), with 5 evaluation sessions (E0 to E4) distributed across the 10 weeks and 8 training sessions (T1 to T8) distributed across 4 weeks (weeks 3 to 6).

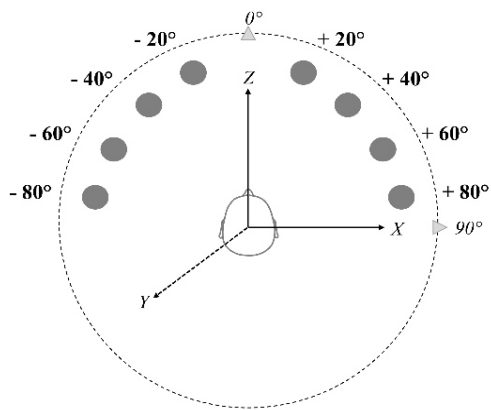
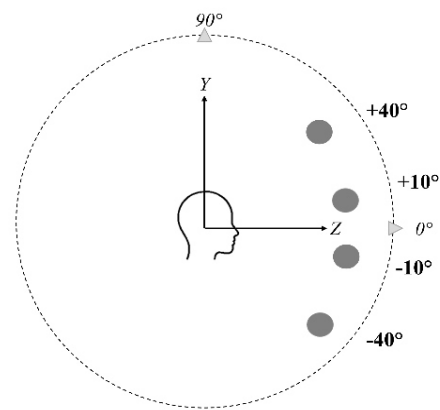
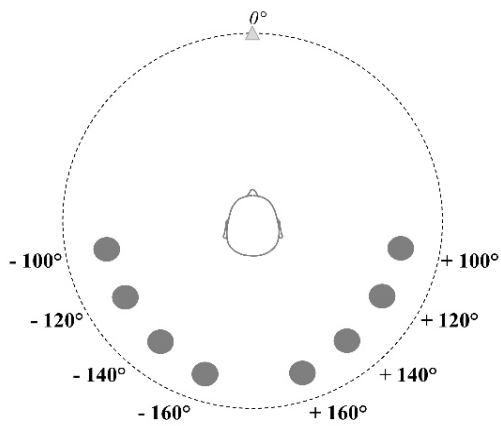
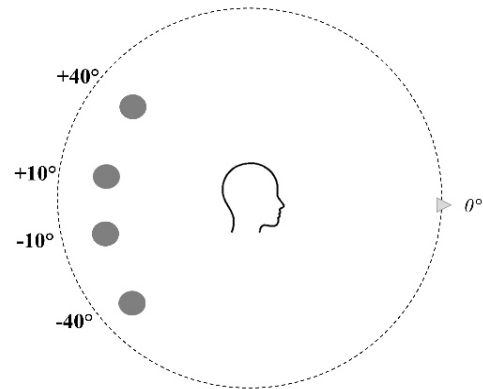
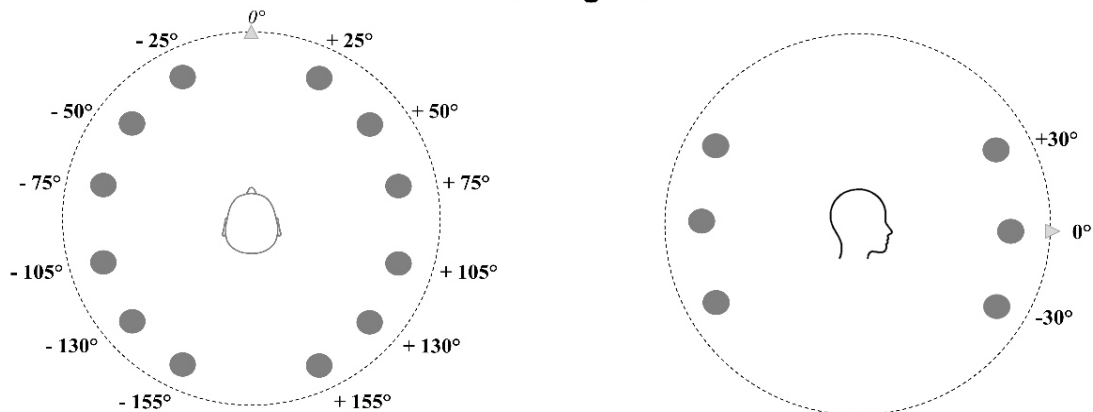
A. Training 1**B. Training 2 and 3****C. Training 4****D. Training 5 and 6****E. Training 7 and 8**

Figure 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. patient 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 $+135^\circ$.

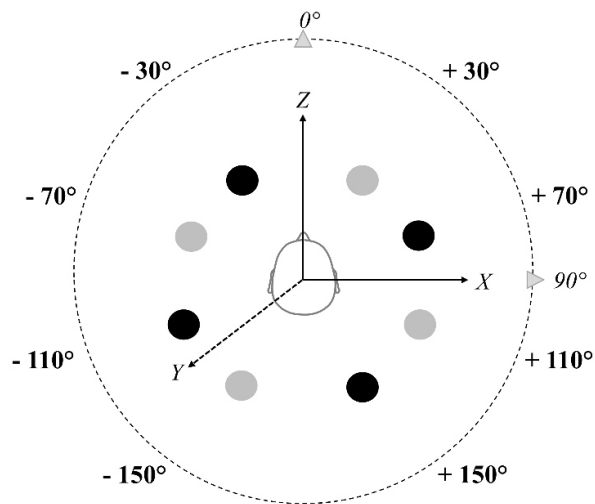


Figure 3. Set up during sound localization evaluation sessions. A total of 8 positions were tested. Black and grey circles indicate two target elevations in the near-field, at $+25^\circ$ and -25° respectively. Distance was fixed at 55cm. Three axes were defined according to the reference-frame (i.e. patient head-centered): X, azimuth; Y, elevation; and Z, distance.

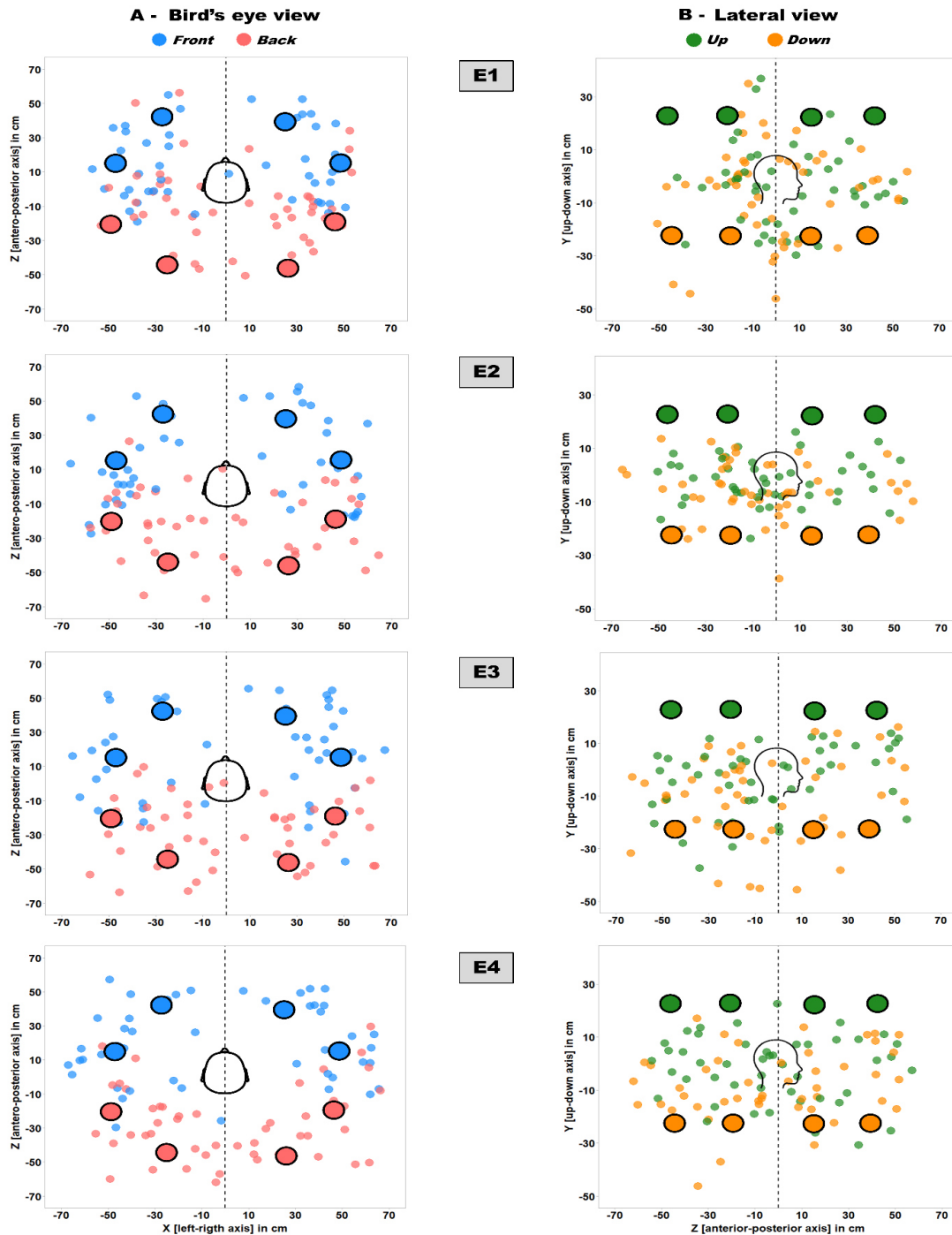


Figure 4. Sound localization performance of all 12 bilateral cochlear implant users during each evaluation session: before training (E1), after 4 (E2) and 8 (E3) training sessions, and 1 month after the end of the training sessions (E4). Black circled symbols represent the sound sources and colored symbols correspond to the mean response for each patient per target. (A) Bird's eye view showing sound localization indicated by the patient's 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).

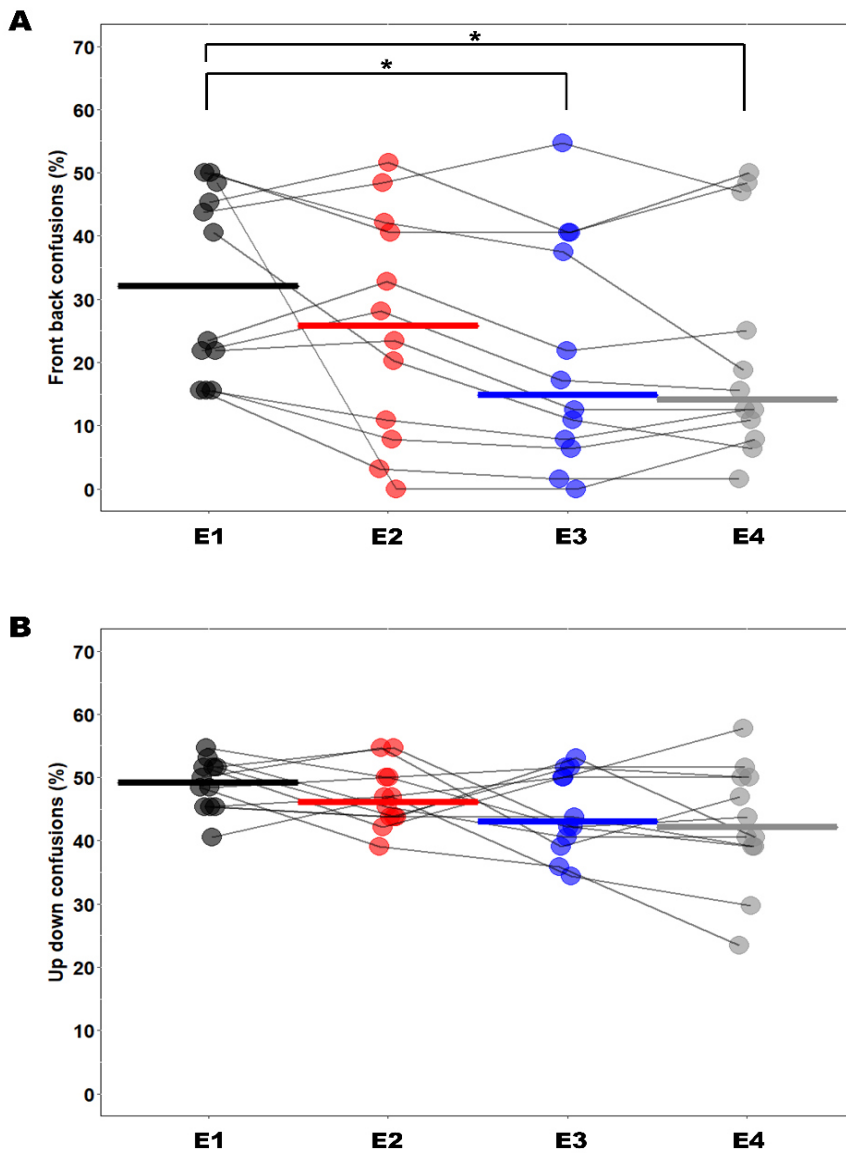


Figure 5. Sound localization performance during evaluation sessions E1 to E4. Thick lines represent the median percentage of confusions for all 12 participants: A) Front back confusions in azimuth, B) Up-down confusions in elevation. Colored dots correspond to the percentage of confusions for each patient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

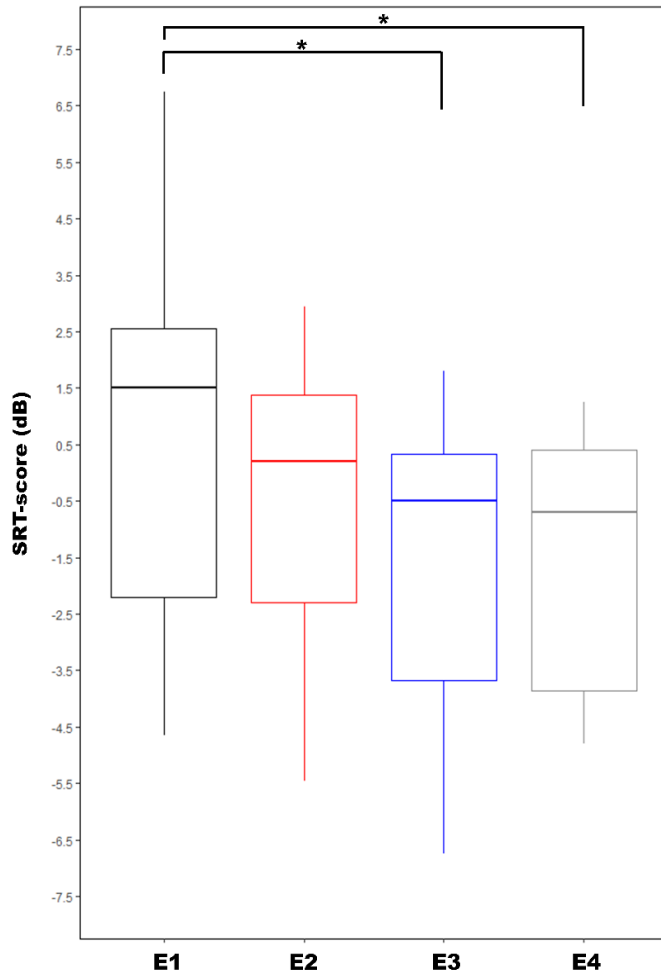


Figure 6. SRT-scores during evaluation sessions E1 to E4. Lines are medians, box limits 25th–75th percentiles and error bars 95% confidence limits (n=11 for each session). Asterisks indicate significant differences (Wilcoxon test, * p<0.05).

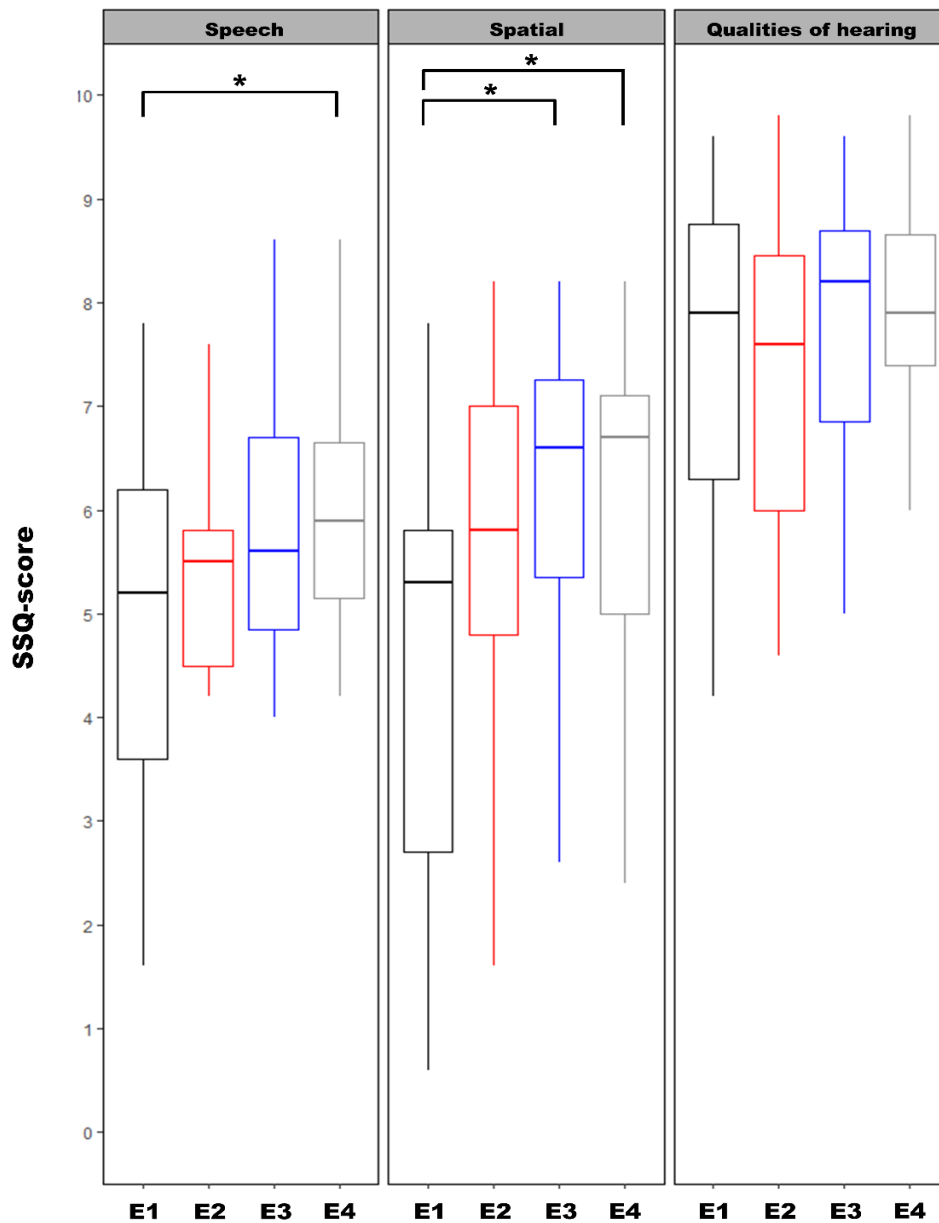


Figure 7. SSQ-score during evaluation sessions E1 to E4 shown separately for each domain (Speech, Spatial, and Qualities of hearing). Lines are medians, box limits 25th–75th percentiles and error bars 95% confidence limits (n=12 for each session). Asterisks indicate significant differences (Wilcoxon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

Supplemental Digital Content 1.mp4: example of an audiovisual training session

Supplemental Digital Content 2.mp4: example of a visual training sessi

Tables

Table 1. Demographics and device information for the 12 bilateral cochlear implant users included in the training protocol. Values in bold represent group averages with standard deviations.

Group	Id	Sex	First Cochlear Implant			Second Cochlear Implant					Interimplant interval (y)	Duration of binaural experience (y)								
			Age at testing (y)	Etiology	Age at diagnosis (y)	Age at hearing aids (y)	Age at CI1 (y)	Internal Part, Processor, Ear side	Strategy	Micro			PTA threshold (dB HL)	Speech reception threshold (%)	Age at CI2 (y)	Internal Part, Processor, Ear side	Strategy	Micro	PTA threshold (dB)	Speech reception threshold (%)
A	A01	M	18	genetic	6	6	12	Nucleus CI422, CP1000, R	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	ACE	Omni	25	86	45	Nucleus CI522, CP1000, R	ACE	Omni	30	90	3	1
	A03	F	45	meningitidis	2	5	39	Concerto, RONDO2, R	Fs4p	Direc	35	88	44	Synchrony, RONDO2, L	Fs4p	Direc	35	96	5	1
	A04	F	23	Ototoxicity	9	9	15	Concerto, SONNET, L	Fs4p	Omni	25	96	17	Concerto, R	OPUS2, Fs4p	Omni	25	97	2	6
	A05	F	49	unknown	6	6	42	CONCERTO, SONNET, R	Fs4p	Omni	25	70	46	Synchrony, SONNET, L	Fs4p	Omni	25	70	4	3
	A06	M	31	meningitidis	2	2	2	Nucleus, CP910, R	SPEAK	Omni	25	72	28	Nucleus CI522, CP910, L	ACE	Omni	30	100	26	3
				35.3 ± 13.2		4.7 ± 2.8	5.1 ± 2.5	25.3 ± 17.7			26.7 ± 4.1	84.7 ± 11.4	32.8 ± 13.9			28.3 ± 4.1	91.8 ± 11.2	7.5 ± 9.1	2.5 ± 2.0	
B	B01	F	69	unknown	55	58	64	Digisonic SP EVO, SAPHYR SP, R	Crystalis XDP	Omni	35	92	67	Digisonic Zti EVO, NEURO2, L	Crystalis CAP	Omni	35	84	3	2
	B02	M	40	Ototoxicity	2	6	34	Digisonic SP EVO, SAPHYR SP, R	Crystalis XDP	Omni	30	82	35	Digisonic SP EVO, SAPHYR SP, L	Crystalis SPD	Omni	35	96	1	5
	B03	F	46	genetic	10	12	42	Concerto, SONNET, R	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	Fs4	Direc	30	87	39	Sonata, RONDO2, L	Fs4	Direc	30	90	1	9
	B05	M	55	genetic	1	1	50	HiRes90K, NAIDA, R	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 2, R	Fs4	Omni	30	87	20	Synchrony, SONNET, L	Fs4	Omni	30	99	3	4
			47.0 ± 15.0		14.0 ± 20.7	18.3 ± 22.1	40.8 ± 15.8			28.3 ± 5.2	90.3 ± 6.3	42.5 ± 15.9			29.2 ± 5.8	91.7 ± 6.7	1.8 ± 1.0	4.5 ± 2.4		

Group A: audiovisual feedback; Group B: visual feedback; Id: patient identification; CI: cochlear implant; y: years; L: left; R: right; Omni: omnidirectional; Direc: directional

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
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

Group A: audiovisual feedback; Group B: visual feedback; Id: patient identification

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	-0.7
	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

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

Group	Id	Domain	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

PERSPECTIVES GENERALES

Grâce aux avancées technologiques de ces dernières décennies, les implants cochléaires (IC) ont permis de changer le quotidien de milliers de patients sourds congénitaux ou devenus sourds tardifs. La restauration d'une bonne compréhension de la parole a longtemps été au cœur des travaux scientifiques qui ont nourri ces avancées. Aujourd'hui, le thème de la spatialisation auditive (SA) représente un second défi d'envergure dans le développement prothétique, notamment en termes d'évaluation pour mieux caractériser les déficits des patients implantés et ainsi adapter les techniques de prises en charge au niveau clinique et technologique.

Nos travaux ont permis de mieux cerner les limites perceptives des enfants bi-implantés (BCI) en 3D mais également à l'échelle individuelle de l'espace azimutal et de la profondeur grâce à un test de localisation 3D et au développement d'un court questionnaire de qualité de vie. Ces déficits de SA étaient contrebalancés par les capacités d'exploitation d'indices acoustiques dynamiques générés par les mouvements de tête. Les enfants BCI sembleraient même plus performants dans l'extraction et l'interprétation de ces indices acoustiques que les adultes BCI. Pour autant, les adultes BCI ont su tirer profit d'un programme d'entraînement de la SA basé sur la recherche active du son et l'intégration multisensorielle.

Au terme de nos recherches, plusieurs perspectives se sont dégagées pour avancer dans la compréhension de la SA chez le patient implanté. Ces perspectives ont été développées individuellement dans la partie discussion de nos 3 études. Nous proposons ici trois pistes futures, synthèse de l'ensemble de nos travaux personnels.

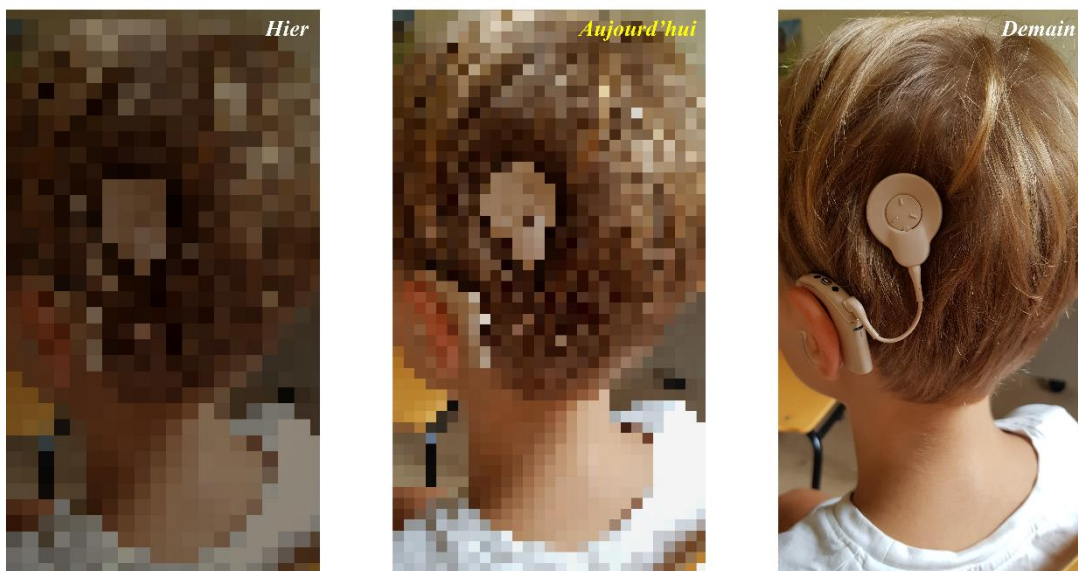
Tout d'abord, l'évolution naturelle des capacités de localisation après la pose d'un puis deux IC chez l'enfant a très peu été explorée. A notre connaissance, elle reste limitée à une seule étude (Galvin et al. 2007) sur une courte période de temps (maximum 8 mois après bilatéralisation). Il serait intéressant de mettre en place un suivi longitudinal composé de mesures itératives des performances de spatialisation, tous les 3 mois la première année post-implantation (ce qui correspond à chaque visite audioprothétique où les réglages sont modifiés), et avec un recul d'au moins 24 mois après la bilatéralisation. Ces évaluations comporteraient systématiquement un test objectif de localisation en 3D adapté à l'âge de l'enfant et un test subjectif, la forme courte pédiatrique du questionnaire SSQ remplie par l'enfant ou les parents en fonction du niveau de développement du langage de l'enfant.

Dans un 2^e temps, les enfants BCI étant en mesure de tirer profit des mouvements de tête au cours de la recherche d'un son, il nous semblerait pertinent de leur proposer un programme d'entraînement de la SA basé sur l'exploitation d'indices acoustiques dynamiques et l'interaction avec d'autres modalités sensorielles, particulièrement visuelles, puisque des

phénomènes de réorganisation corticale au profit de la modalité visuelle se mettent en place précocement chez l'enfant implanté (Campbell et al. 2016).

Enfin, l'ensemble de nos travaux chez l'adulte et l'enfant BCI s'est focalisé sur l'aspect comportemental des capacités de SA à plusieurs instants T. Il paraîtrait intéressant de coupler nos tests de localisation à des mesures électrophysiologiques afin d'évaluer l'impact d'un protocole d'entraînement sur les structures auditives centrales et de mieux comprendre les subtilités des changements cérébraux induits. Plusieurs études ont montré la faisabilité de telles mesures chez les patients IC, par le biais de la tomographie par émission de positons (Giraud et al. 2001, Strelnikov et al. 2015) ou l'électro-encéphalographie (Petersen et al. 2015, Schäfer et al. 2021, Wang et al. 2021).

La voie de la spatialisation auditive est encore longue, mais les perspectives sont prometteuses...



**La science ne peut pas résoudre le mystère ultime de la nature (*Max Planck*).
Mais elle nous fait entrevoir des horizons plus élargis (*Berthelot*).**

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), e0214603.
- 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 of the United States of America*, *98*(21), 12301-12306.
- Amitay, S., Halliday, L., Taylor, J., Sohoglu, E., & Moore, D. R. (2010). Motivation and intelligence drive auditory perceptual learning. *PloS One*, *5*(3), e9816.
- Angell, J. R. (1901). Monaural localization of sound. *Science (New York, N.Y.)*, *13*(333), 775-777.
- Angell, J.R., & Fite, W. (1901) Further observations on the monaural localization of sound. *Science (New York, N.Y.)*, *13*(333), 849-858
- Antunes, F. M., & Malmierca, M. S. (2011). Effect of auditory cortex deactivation on stimulus-specific adaptation in the medial geniculate body. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *31*(47), 17306-17316.
- 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. *The Journal of the Acoustical Society of America*, *145*(3), 1389.
- Arlinger, S. D. (2001). How to assess outcome of hearing aid fitting in children. *Scandinavian Audiology. Supplementum*, *53*, 68-72.
- 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.
- 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. *Frontiers in Neurology*, *11*, 915.
- Ausili, S. A., Backus, B., Agterberg, M. J. H., van Opstal, A. J., & van Wanrooij, M. M. (2019). Sound Localization in Real-Time Vcoded Cochlear-Implant Simulations With Normal-Hearing Listeners. *Trends in Hearing*, *23*, 2331216519847332.
- Barkat, T. R., Polley, D. B., & Hensch, T. K. (2011). A critical period for auditory thalamocortical connectivity. *Nature Neuroscience*, *14*(9), 1189-1194.
- Batteau, D. W. (1967). The role of the pinna in human localization. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, *168*(1011), 158-180.

- Beijen, J., Snik, A. F. M., Straatman, L. V., Mylanus, E. a. M., & Mens, L. H. M. (2010). Sound localization and binaural hearing in children with a hearing aid and a cochlear implant. *Audiology & Neuro-Otology*, *15*(1), 36-43.
- Best, V., van Schaik, A., & Carlile, S. (2004). Separation of concurrent broadband sound sources by human listeners. *The Journal of the Acoustical Society of America*, *115*(1), 324-336.
- Bishop, C. W., London, S., & Miller, L. M. (2012). Neural time course of visually enhanced echo suppression. *Journal of Neurophysiology*, *108*(7), 1869-1883.
- Blauert, J. (1997). Spatial hearing: the psychophysics of human sound localization. *Cambridge, Massachusetts/Cambridge, MA: MIT Press*.
- Bloom, P. J. (1977). Determination of monaural sensitivity changes due to the pinna by use of minimum-audible-field measurements in the lateral vertical plane. *The Journal of the Acoustical Society of America*, *61*(3), 820-828.
- Bolognini, N. , Leor, F. , Passamonti, C. , Stein, B.E. , Làdavas, E. (2007). Multisensory-mediated auditory localization. *Perception*, *36* (10), 1477–1485.
- Bregman, A.S. (1990). Auditory Scene Analysis: The Perceptual Organization of Sound. *MIT Press, Cambridge, MA*.
- 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.
- Brimijoin, W. O., & Akeroyd, M. A. (2014). Head-movement compensation results in a slightly moving auditory world. Paper presented at the Association for Research in Otolaryngology (ARO), San Diego, CA.
- Brimijoin, W. O., & Akeroyd, M. A. (2016). The Effects of Hearing Impairment, Age, and Hearing Aids on the Use of Self-Motion for Determining Front/Back Location. *Journal of the American Academy of Audiology*, *27*(7), 588-600.
- 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.
- Brungart, D. S., Durlach, N. I., & Rabinowitz, W. M. (1999a). Auditory localization of nearby sources. II. Localization of a broadband source. *The Journal of the Acoustical Society of America*, *106*(4 Pt 1), 1956-1968.
- Brungart, D. S. (1999b). Auditory localization of nearby sources. III. Stimulus effects. *The Journal of the Acoustical Society of America*, *106*(6), 3589-3602.

- Bronkhorst, A. W. (2000). "The cocktail party phenomenon: A review of research on speech intelligibility in multiple-talker conditions." *Acust. Acta Acust.* 86, 117–128.
- Bulkin, D. A., & Groh, J. M. (2006). Seeing sounds : Visual and auditory interactions in the brain. *Current Opinion in Neurobiology*, 16(4), 415-419.
- Burr, D., & Alais, D. (2006). Combining visual and auditory information. *Progress in Brain Research*, 155, 243-258.
- Butler, R. A. (1986). The bandwidth effect on monaural and binaural localization. *Hearing Research*, 21(1), 67-73.
- Butler, R. A., & Humanski, R. A. (1992). Localization of sound in the vertical plane with and without high-frequency spectral cues. *Perception & Psychophysics*, 51(2), 182-186.
- Campbell, J., & Sharma, A. (2014). Cross-modal re-organization in adults with early stage hearing loss. *PloS One*, 9(2), e90594.
- Campbell, J., & Sharma, A. (2016). Visual Cross-Modal Re-Organization in Children with Cochlear Implants. *PloS One*, 11(1), e0147793. <https://doi.org/10.1371/journal.pone.0147793>
- Campus, C., Sandini, G., Concetta Morrone, M., & Gori, M. (2017). Spatial localization of sound elicits early responses from occipital visual cortex in humans. *Scientific Reports*, 7(1), 10415.
- Cappelloni, M. S., Shivkumar, S., Haefner, R. M., & Maddox, R. K. (2019). Task-uninformative visual stimuli improve auditory spatial discrimination in humans but not the ideal observer. *PloS One*, 14(9), e0215417.
- Carlile, S., Leong, P., & Hyams, S. (1997). The nature and distribution of errors in sound localization by human listeners. *Hearing Research*, 114(1-2), 179-196.
- Carlile, S. & Leung, J. (2016). The perception of auditory motion. *Trends in Hearing*, 20, 1–19.
- Chang, S.-A., Tyler, R. S., Dunn, C. C., Ji, H., Witt, S. A., Gantz, B., & Hansen, M. (2010). Performance over time on adults with simultaneous bilateral cochlear implants. *Journal of the American Academy of Audiology*, 21(1), 35-43.
- 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 and Hearing*, 38(4), 426-440.
- Chun, S., Bayazitov, I. T., Blundon, J. A., & Zakharenko, S. S. (2013). Thalamocortical long-term potentiation becomes gated after the early critical period in the auditory cortex. *The*

- Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 33(17), 7345-7357.
- Clavagnier, S., Falchier, A., & Kennedy, H. (2004). Long-distance feedback projections to area V1 : Implications for multisensory integration, spatial awareness, and visual consciousness. *Cognitive, Affective & Behavioral Neuroscience*, 4(2), 117-126.
- Cooper, J., Carlile, S., & Alais, D. (2008). Distortions of auditory space during rapid head turns. *Experimental Brain Research*, 191(2), 209-219.
- Crair, M. C., & Malenka, R. C. (1995). A critical period for long-term potentiation at thalamocortical synapses. *Nature*, 375(6529), 325-328.
- 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.
- Demeester, K., Topsakal, V., Hendrickx, J.-J., Fransen, E., van Laer, L., Van Camp, G., Van de Heyning, P., & van Wieringen, A. (2012). Hearing disability measured by the speech, spatial, and qualities of hearing scale in clinically normal-hearing and hearing-impaired middle-aged persons, and disability screening by means of a reduced SSQ (the SSQ5). *Ear and Hearing*, 33(5), 615-616.
- 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.
- Dillon, H. (2001). Hearing Aids. *Boomerang press*.
- Dorman, M. F., Loiselle, L. H., Cook, S. J., Yost, W. A., & Gifford, R. H. (2016). Sound Source Localization by Normal-Hearing Listeners, Hearing-Impaired Listeners and Cochlear Implant Listeners. *Audiology & Neuro-Otology*, 21(3), 127-131.
- 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 and Hearing*, 35(6), 633-640.
- Dorman, M. F., Natale, S. C., Butts, A. M., Zeitler, D. M., & Carlson, M. L. (2017). The Sound Quality of Cochlear Implants : Studies With Single-sided Deaf Patients. *Otology & Neurotology: Official Publication of the American Otological Society, American Neurotology Society [and] European Academy of Otology and Neurotology*, 38(8), e268-e273.
- Dorman, M. F., Natale, S., & Loiselle, L. (2018). Speech Understanding and Sound Source Localization by Cochlear Implant Listeners Using a Pinna-Effect Imitating Microphone and an Adaptive Beamformer. *Journal of the American Academy of Audiology*, 29(3), 197-205.

- 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. *Journal of the American Academy of Audiology*, 29(9), 835-846.
- Edmonds, B. A., & Culling, J. F. (2005). The spatial unmasking of speech : Evidence for within-channel processing of interaural time delay. *The Journal of the Acoustical Society of America*, 117(5), 3069-3078.
- Egger, K., Majdak, P., & Laback, B. (2016). Channel Interaction and Current Level Affect Across-Electrode Integration of Interaural Time Differences in Bilateral Cochlear-Implant Listeners. *Journal of the Association for Research in Otolaryngology: JARO*, 17(1), 55-67.
- Ehlers, E., Goupell, M. J., Zheng, Y., Godar, S. P., & Litovsky, R. Y. (2017). Binaural sensitivity in children who use bilateral cochlear implants. *The Journal of the Acoustical Society of America*, 141(6), 4264.
- Eklöf, M., & Tideholm, B. (2018). The choice of stimulation strategy affects the ability to detect pure tone inter-aural time differences in children with early bilateral cochlear implantation. *Acta Oto-Laryngologica*, 138(6), 554-561.
- Falchier, A., Clavagnier, S., Barone, P., & Kennedy, H. (2002). Anatomical evidence of multimodal integration in primate striate cortex. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 22(13), 5749-5759.
- Faller, C., & Merimaa, J. (2004). Source localization in complex listening situations : Selection of binaural cues based on interaural coherence. *The Journal of the Acoustical Society of America*, 116(5), 3075-3089.
- Fandakova, Y., & Hartley, C. A. (2020). Mechanisms of learning and plasticity in childhood and adolescence. *Developmental Cognitive Neuroscience*, 42, 100764.
- Feeny, D., Furlong, W., Boyle, M., & Torrance, G. W. (1995). Multi-attribute health status classification systems. Health Utilities Index. *Pharmacoeconomics*, 7(6), 490-502.
- 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. *Hearing Research*, 319, 48-55.
- Fischer, T., Schmid, C., Kompis, M., Mantokoudis, G., Caversaccio, M., & Wimmer, W. (2020a). Pinna-Imitating Microphone Directionality Improves Sound Localization and Discrimination in Bilateral Cochlear Implant Users. *Ear and Hearing*, 42(1), 214-222.
- Fischer, T., Caversaccio, M., & Wimmer, W. (2020b). A Front-Back Confusion Metric in Horizontal Sound Localization: The FBC Score. *In ACM Symposium on Applied Perception 2020 (SAP '20), September 12–13, 2020, Virtual Event, USA. ACM, New York, NY, USA 5*

- Fitzgerald, M. B., Kan, A., & Goupell, M. J. (2015). Bilateral Loudness Balancing and Distorted Spatial Perception in Recipients of Bilateral Cochlear Implants. *Ear and Hearing*, 36(5), e225-236.
- Freeman, T. C. A., Leung, J., Wufong, E., Orchard-Mills, E., Carlile, S., & Alais, D. (2014). Discrimination contours for moving sounds reveal duration and distance cues dominate auditory speed perception. *PloS One*, 9(7), e102864.
- Freigang, C., Richter, N., Rübsamen, R., & Ludwig, A. A. (2015). Age-related changes in sound localisation ability. *Cell and Tissue Research*, 361(1), 371-386.
- Fu, Q.-J., & Nogaki, G. (2005). Noise susceptibility of cochlear implant users : The role of spectral resolution and smearing. *Journal of the Association for Research in Otolaryngology: JARO*, 6(1), 19-27.
- Fuller, J. H. (1992). Head movement propensity. *Exp. Brain Res.* 92, 152–164.
- Gage, F. H. (2004). Structural plasticity of the adult brain. *Dialogues in Clinical Neuroscience*, 6(2), 135-141.
- Galván, A. (2010). Neural plasticity of development and learning. *Human Brain Mapping*, 31(6), 879-890.
- 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 and Hearing*, 37(1), 93-102.
- Galvin, K. L., Mok, M., & Dowell, R. C. (2007). Perceptual benefit and functional outcomes for children using sequential bilateral cochlear implants. *Ear and Hearing*, 28(4), 470-482.
- 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.
- Garas, J. Adaptive 3D sound systems. (2000). *Boston, MA: Springer US*.
- Gatehouse, S., & Noble, W. (2004). The Speech, Spatial and Qualities of Hearing Scale (SSQ). *International journal of audiology*, 43(2), 85-99.
- Giraud, A. L., Price, C. J., Graham, J. M., & Frackowiak, R. S. (2001). Functional plasticity of language-related brain areas after cochlear implantation. *Brain: A Journal of Neurology*, 124(Pt 7), 1307-1316.
- Glick, H., & Sharma, A. (2017). Cross-modal plasticity in developmental and age-related hearing loss : Clinical implications. *Hearing Research*, 343, 191-201.
- Goossens, H. H., & van Opstal, A. J. (1999). Influence of head position on the spatial representation of acoustic targets. *Journal of Neurophysiology*, 81(6), 2720-2736.

- 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. *Current Opinion in Otolaryngology & Head and Neck Surgery*, 25(6), 514-519.
- Gordon, K. A., Deighton, M. R., Abbasalipour, P., & Papsin, B. C. (2014). Perception of binaural cues develops in children who are deaf through bilateral cochlear implantation. *PloS One*, 9(12), e114841.
- Gordon, K. A., Jiwani, S., & Papsin, B. C. (2011). What is the optimal timing for bilateral cochlear implantation in children? *Cochlear Implants International*, 12 Suppl 2, S8-14.
- Gordon, K. A., Jiwani, S., & Papsin, B. C. (2013). Benefits and detriments of unilateral cochlear implant use on bilateral auditory development in children who are deaf. *Frontiers in Psychology*, 4, 719.
- Gordon, K. A., Papsin, B. C., & Harrison, R. V. (2003). Activity-dependent developmental plasticity of the auditory brain stem in children who use cochlear implants. *Ear and Hearing*, 24(6), 485-500.
- Gordon, K. A., Salloum, C., Toor, G. S., van Hoesel, R., & Papsin, B. C. (2012). Binaural interactions develop in the auditory brainstem of children who are deaf : Effects of place and level of bilateral electrical stimulation. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 32(12), 4212-4223.
- 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. *Otology & Neurotology: Official Publication of the American Otological Society, American Neurotology Society [and] European Academy of Otology and Neurotology*, 29(2), 193-198.
- Gordon, K. A., Wong, D. D. E., & Papsin, B. C. (2010). Cortical function in children receiving bilateral cochlear implants simultaneously or after a period of interimplant delay. *Otology & Neurotology: Official Publication of the American Otological Society, American Neurotology Society [and] European Academy of Otology and Neurotology*, 31(8), 1293-1299.
- 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: A Journal of Neurology*, 136(Pt 5), 1609-1625.
- Gordon, K., Henkin, Y., & Kral, A. (2015). Asymmetric Hearing During Development : The Aural Preference Syndrome and Treatment Options. *Pediatrics*, 136(1), 141-153.
- Gori, M., Del Viva, M., Sandini, G., & Burr, D. C. (2008). Young children do not integrate visual and haptic form information. *Current Biology: CB*, 18(9), 694-698.

- Gori, M., Sandini, G., & Burr, D. (2012). Development of visuo-auditory integration in space and time. *Frontiers in Integrative Neuroscience*, 6, 77.
- Grant, K. W., & Seitz, P. F. (2000). The use of visible speech cues for improving auditory detection of spoken sentences. *The Journal of the Acoustical Society of America*, 108(3 Pt 1), 1197-1208.
- Grantham, D. W., Ashmead, D. H., Ricketts, T. A., Labadie, R. F., & Haynes, D. S. (2007). Horizontal-plane localization of noise and speech signals by postlingually deafened adults fitted with bilateral cochlear implants. *Ear and Hearing*, 28(4), 524-541.
- Greenough, W. T., Black, J. E., & Wallace, C. S. (1987). Experience and brain development. *Child Development*, 58(3), 539-559.
- Grothe, B. (2000). The evolution of temporal processing in the medial superior olive, an auditory brainstem structure. *Progress in Neurobiology*, 61(6), 581-610.
- Grothe, B., Pecka, M., & McAlpine, D. (2010). Mechanisms of sound localization in mammals. *Physiological Reviews*, 90(3), 983-1012. <https://doi.org/10.1152/physrev.00026.2009>
- Hartmann, W. M. (1999). How we localize sound. *Phys Today*, 52, 24-9.
- Hawley, M. L., Litovsky, R. Y., & Colburn, H. S. (1999). Speech intelligibility and localization in a multi-source environment. *The Journal of the Acoustical Society of America*, 105(6), 3436-3448.
- Heffner, R. S. (2004). Primate hearing from a mammalian perspective. *The Anatomical Record. Part A, Discoveries in Molecular, Cellular, and Evolutionary Biology*, 281(1), 1111-1122.
- Heffner, R. S., & Heffner, H. E. (1992). Visual factors in sound localization in mammals. *The Journal of Comparative Neurology*, 317(3), 219-232.
- Heffner, H.E., Heffner, R.S., (2014). *The Behavioral Study of Mammalian Hearing*. Springer, New York, NY, pp. 269-285 .
- Henshaw, H., McCormack, A., & Ferguson, M. A. (2015). Intrinsic and extrinsic motivation is associated with computer-based auditory training uptake, engagement, and adherence for people with hearing loss. *Frontiers in Psychology*, 6, 1067.
- Hermann, R., Lescanne, E., Loundon, N., 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, M. B., Roman, S., & Truy, E. (2019). French Society of ENT (SFORL) guidelines. Indications for cochlear implantation in adults. *European Annals of Otorhinolaryngology, Head and Neck Diseases*, 136(3)
- Hofman, P. M., Van Riswick, J. G., & Van Opstal, A. J. (1998). Relearning sound localization with new ears. *Nature Neuroscience*, 1(5), 417-421.

- 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.
- Illg, A., Giourgas, A., Kral, A., Büchner, A., Lesinski-Schiedat, A., & Lenarz, T. (2013). Speech comprehension in children and adolescents after sequential bilateral cochlear implantation with long interimplant interval. *Otology & Neurotology: Official Publication of the American Otological Society, American Neurotology Society [and] European Academy of Otology and Neurotology*, 34(4), 682-689.
- Irving, S., & Moore, D. R. (2011). Training sound localization in normal hearing listeners with and without a unilateral ear plug. *Hearing Research*, 280(1-2), 100-108.
- Isaiah, A., Vongpaisal, T., King, A. J., & Hartley, D. E. H. (2014). Multisensory training improves auditory spatial processing following bilateral cochlear implantation. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 34(33), 11119-11130.
- Iwaya, Y., Suzuki, Y., & Kimura, D. (2003). Effects of head movement on front-back error in sound localization. *Acoustical Science and Technology*, 24, 322-324.
- Jeffress, L. A. (1948). A place theory of sound localization. *Journal of Comparative and Physiological Psychology*, 41(1), 35-39.
- 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 and Hearing*, 37(5), e341-345.
- Kan, A., Goupell, M. J., & Litovsky, R. Y. (2019). Effect of channel separation and interaural mismatch on fusion and lateralization in normal-hearing and cochlear-implant listeners. *The Journal of the Acoustical Society of America*, 146(2), 1448.
- Kan, A., & Litovsky, R. Y. (2015a). Binaural hearing with electrical stimulation. *Hearing Research*, 322, 127-137.
- Kan, A., & Litovsky, R. Y. (2015b). Binaural hearing with electrical stimulation. *Hearing Research*, 322, 127-137.
- Kan, A., Litovsky, R. Y., & Goupell, M. J. (2015). Effects of interaural pitch matching and auditory image centering on binaural sensitivity in cochlear implant users. *Ear and Hearing*, 36(3), e62-68.
- Kan, A., Stoelb, C., Litovsky, R. Y., & Goupell, M. J. (2013). Effect of mismatched place-of-stimulation on binaural fusion and lateralization in bilateral cochlear-implant users. *The Journal of the Acoustical Society of America*, 134(4), 2923-2936.

- Keating, P., & King, A. J. (2015). Sound localization in a changing world. *Current Opinion in Neurobiology*, 35, 35-43.
- Kerber, S., & Seeber, B. U. (2012). Sound localization in noise by normal-hearing listeners and cochlear implant users. *Ear and Hearing*, 33(4), 445-457.
- Kessler A.R., Giolas T.G., Maxon A.B. (1990). The hearing performance inventory for children (HPIC): reliability and validity, *Poster presented at the American Speech-Language-Hearing Association Convention, November, Seattle.*
- 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(11), e82263.
- Kidd, G., Arbogast, T. L., Mason, C. R., & Gallun, F. J. (2005). The advantage of knowing where to listen. *The Journal of the Acoustical Society of America*, 118(6), 3804-3815.
- 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 International*, 19(5), 284-291.
- 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.
- 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. *Proc Meet Acoust, Acoust Soc Am*, 19, 050142.
- King, A. J. (2009). Visual influences on auditory spatial learning. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 364(1515), 331-339.
- Klumpp, R.G., & Eady, H.R. (1956). Some measurements of interaural time difference thresholds. *J Acoust Soc Am*, 28, 859–860.
- Knudsen, E. I., & Knudsen, P. F. (1985). Vision guides the adjustment of auditory localization in young barn owls. *Science (New York, N.Y.)*, 230(4725), 545-548.
- Kral, A., Dorman, M. F., & Wilson, B. S. (2019). Neuronal Development of Hearing and Language : Cochlear Implants and Critical Periods. *Annual Review of Neuroscience*, 42, 47-65.
- Kral, A., Heid, S., Hubka, P., & Tillein, J. (2013). Unilateral hearing during development : Hemispheric specificity in plastic reorganizations. *Frontiers in Systems Neuroscience*, 7, 93.
- Kral, A., Tillein, J., Heid, S., Hartmann, R., & Klinke, R. (2005). Postnatal cortical development in congenital auditory deprivation. *Cerebral Cortex (New York, N.Y.: 1991)*, 15(5), 552-562.

- Kral, A., Tillein, J., Hubka, P., Schiemann, D., Heid, S., Hartmann, R., & Engel, A. K. (2009). Spatiotemporal Patterns of Cortical Activity with Bilateral Cochlear Implants in Congenital Deafness. *The Journal of Neuroscience*, 29(3), 811-827.
- 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. *Audiology & Neuro-Otology*, 18(1), 48-62.
- Kuk, F., Keenan, D. M., Lau, C., Crose, B., & Schumacher, J. (2014). Evaluation of a localization training program for hearing impaired listeners. *Ear and Hearing*, 35(6), 652-666.
- 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: The Official Journal of the Society for Neuroscience*, 30(14), 4883-4894.
- Laback, B., Egger, K., & Majdak, P. (2015). Perception and coding of interaural time differences with bilateral cochlear implants. *Hearing Research*, 322, 138-150.
- Lambert, R. M. (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>
- Langendijk, E. H., Kistler, D. J., & Wightman, F. L. (2001). Sound localization in the presence of one or two distracters. *The Journal of the Acoustical Society of America*, 109(5 Pt 1), 2123-2134.
- Leung, J., Alais, D., & Carlile, S. (2008). Compression of auditory space during rapid head turns. *Proceedings of the National Academy of Sciences of the United States of America*, 105(17), 6492-6497.
- Lewald, J., Dörrscheidt, G. J., & Ehrenstein, W. H. (2000). Sound localization with eccentric head position. *Behavioural Brain Research*, 108(2), 105-125.
- Lewkowicz, D.J. (2000). Perceptual development: visual, auditory, and speech perception in infancy. *American Journal of Psychology*, 113 (3), 488–500.
- Lewkowicz, D. J. (2002). Heterogeneity and heterochrony in the development of intersensory perception. *Brain Research. Cognitive Brain Research*, 14(1), 41-63.
- Lewkowicz, D. J. & Lickliter, R. (1994). The Development of Intersensory Perception: Comparative Perspectives (*Lawrence Erlbaum Associates*).
- Litovsky, R.Y. (2011). Springer Handbook of Auditory Research. New York: Springer-Verlag; Development of binaural and spatial hearing; p. 163-195.
- Litovsky, R.Y. (2015). Development of the auditory system. *Handbook of clinical neurology*, 129, 55-72.

- Litovsky, R. Y., & Godar, S. P. (2010). Difference in precedence effect between children and adults signifies development of sound localization abilities in complex listening tasks. *The Journal of the Acoustical Society of America*, 128(4), 1979-1991.
- Litovsky, R. Y., & Gordon, K. (2016). Bilateral cochlear implants in children : Effects of auditory experience and deprivation on auditory perception. *Hearing Research*, 338, 76-87.
- Litovsky, R. Y., Jones, G. L., Agrawal, S., & van Hoesel, R. (2010). Effect of age at onset of deafness on binaural sensitivity in electric hearing in humans. *The Journal of the Acoustical Society of America*, 127(1), 400-414.
- Litovsky, R. Y., Parkinson, A., & Arcaroli, J. (2009). Spatial hearing and speech intelligibility in bilateral cochlear implant users. *Ear and Hearing*, 30(4), 419-431.
- Litovsky, R. Y., Parkinson, A., Arcaroli, J., Peters, R., Lake, J., Johnstone, P., & Yu, G. (2004). Bilateral cochlear implants in adults and children. *Archives of Otolaryngology--Head & Neck Surgery*, 130(5), 648-655.
- Lord Rayleigh. XII. (1907). On our perception of sound direction. *Philos Mag Ser*, 13,214–32.
- Lorens, A., Zgoda, M., Obrycka, A., & Skarzynski, H. (2010). Fine Structure Processing improves speech perception as well as objective and subjective benefits in pediatric MED-EL COMBI 40+ users. *International Journal of Pediatric Otorhinolaryngology*, 74(12), 1372-1378.
- Lövdén, M., Wenger, E., Mårtensson, J., Lindenberger, U., & Bäckman, L. (2013). Structural brain plasticity in adult learning and development. *Neuroscience and Biobehavioral Reviews*, 37(9 Pt B), 2296-2310.
- Lovett, R. E. S., Kitterick, P. T., Hewitt, C. E., & Summerfield, A. Q. (2010). Bilateral or unilateral cochlear implantation for deaf children : An observational study. *Archives of Disease in Childhood*, 95(2), 107-112.
- Lovett, R. E. S., Kitterick, P. T., Huang, S., & Summerfield, A. Q. (2012). The developmental trajectory of spatial listening skills in normal-hearing children. *Journal of Speech, Language, and Hearing Research: JSLHR*, 55(3), 865-878.
- Macpherson, E. A., & Middlebrooks, J. C. (2003). Vertical-plane sound localization probed with ripple-spectrum noise. *The Journal of the Acoustical Society of America*, 114(1), 430-445.
- 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.

- Majdak, P., Walder, T., & Laback, B. (2013). Effect of long-term training on sound localization performance with spectrally warped and band-limited head-related transfer functions. *The Journal of the Acoustical Society of America*, *134*(3), 2148-2159.
- Makous, J. C., & Middlebrooks, J. C. (1990). Two-dimensional sound localization by human listeners. *The Journal of the Acoustical Society of America*, *87*(5), 2188-2200.
- Marrone, N., Mason, C. R., & Kidd, G. (2008). Tuning in the spatial dimension : Evidence from a masked speech identification task. *The Journal of the Acoustical Society of America*, *124*(2), 1146-1158.
- McConkey Robbins, A., Koch, D. B., Osberger, M. J., Zimmerman-Phillips, S., & Kishon-Rabin, L. (2004). Effect of age at cochlear implantation on auditory skill development in infants and toddlers. *Archives of Otolaryngology--Head & Neck Surgery*, *130*(5), 570-574.
- Mendonça, C. (2014). A review on auditory space adaptations to altered head-related cues. *Frontiers in Neuroscience*, *8*, 219.
- Mendonça, C., Campos, G., Dias, P., & Santos, J. A. (2013). Learning auditory space : Generalization and long-term effects. *PloS One*, *8*(10), e77900.
- Middlebrooks, J. C., & Green, D. M. (1991). Sound localization by human listeners. *Annual Review of Psychology*, *42*, 135-159.
- Mills, A.W. (1958). On the minimum audible angle. *J Acoust Soc Am*, *30*,237–46.
- Mok, M., Galvin, K. L., Dowell, R. C., & McKay, C. M. (2010). Speech perception benefit for children with a cochlear implant and a hearing aid in opposite ears and children with bilateral cochlear implants. *Audiology & Neuro-Otology*, *15*(1), 44-56.
- Morrongiello, B. A., & Rocca, P. T. (1987). Infants' localization of sounds in the horizontal plane : Effects of auditory and visual cues. *Child Development*, *58*(4), 918-927.
- Mosnier, I., Sterkers, O., Bebear, J.-P., Godey, B., Robier, A., Deguine, O., Fraysse, B., Bordure, P., Mondain, M., Bouccara, D., Bozorg-Grayeli, A., Borel, S., Ambert-Dahan, E., & Ferrary, E. (2009). Speech performance and sound localization in a complex noisy environment in bilaterally implanted adult patients. *Audiology & Neuro-Otology*, *14*(2), 106-114.
- Moulin, A., Pauzie, 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. *International Journal of Audiology*, *54*(12), 889-898.
- 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 and Hearing*, *40*(4), 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 International*, *15*(1), 36-42. <https://doi.org/10.1179/1754762813Y.0000000040>
- Muir, D., & Field, J. (1979). Newborn infants orient to sounds. *Child Development*, *50*(2), 431-436.
- Munhall, K. G., Jones, J. A., Callan, D. E., Kuratate, T., & Vatikiotis-Bateson, E. (2004). Visual prosody and speech intelligibility : Head movement improves auditory speech perception. *Psychological Science*, *15*(2), 133-137. <https://doi.org/10.1111/j.0963-7214.2004.01502010>.
- Musicant, A. D., & Butler, R. A. (1984). The influence of pinnae-based spectral cues on sound localization. *The Journal of the Acoustical Society of America*, *75*(4), 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. *Otology & Neurotology: Official Publication of the American Otological Society, American Neurotology Society [and] European Academy of Otology and Neurotology*, *35*(2), 271-276.
- Neil, P. A., Chee-Ruiter, C., Scheier, C., Lewkowicz, D. J., & Shimojo, S. (2006). Development of multisensory spatial integration and perception in humans. *Developmental Science*, *9*(5), 454-464.
- Neuman, A. C., Haravon, A., Sislian, N., & Waltzman, S. B. (2007). Sound-direction identification with bilateral cochlear implants. *Ear and Hearing*, *28*(1), 73-82.
- Nicholas, J. G., & Geers, A. E. (2013). Spoken language benefits of extending cochlear implant candidacy below 12 months of age. *Otology & Neurotology: Official Publication of the American Otological Society, American Neurotology Society [and] European Academy of Otology and Neurotology*, *34*(3), 532-538.
- Niparko, J. K., Tobey, E. A., Thal, D. J., Eisenberg, L. S., Wang, N.-Y., Quittner, A. L., Fink, N. E., & CDaCI Investigative Team. (2010). Spoken language development in children following cochlear implantation. *JAMA*, *303*(15), 1498-1506.
- Noble, W., Jensen, N. S., Naylor, G., Bhullar, N., & Akeroyd, M. A. (2013). A short form of the Speech, Spatial and Qualities of Hearing scale suitable for clinical use : The SSQ12. *International Journal of Audiology*, *52*(6), 409-412.
- Nopp, P., Schleich, P., & D'Haese, P. (2004). Sound localization in bilateral users of MED-EL COMBI 40/40+ cochlear implants. *Ear and Hearing*, *25*(3), 205-214.

- Opoku-Baah, C., Schoenhaut, A. M., Vassall, S. G., Tovar, D. A., Ramachandran, R., & Wallace, M. T. (2021). Visual Influences on Auditory Behavioral, Neural, and Perceptual Processes : A Review. *Journal of the Association for Research in Otolaryngology: JARO*, 22(4), 365-386.
- Opstal, J.V. (2016). The auditory system and human sound localization behavior. *Boston,MA: Elsevier*.
- Otte, R. J., Agterberg, M. J. H., Van Wanrooij, M. M., Snik, A. F. M., & Van Opstal, A. J. (2013). Age-related hearing loss and ear morphology affect vertical but not horizontal sound-localization performance. *Journal of the Association for Research in Otolaryngology: JARO*, 14(2), 261-273.
- Ou, H., Perreau, A., & Tyler, R. (2017). Development of a Shortened Version of the Spatial Hearing Questionnaire (SHQ-S) for Screening Spatial-Hearing Ability. *American Journal of Audiology*, 26, 1.
- Pascual-Leone, A., Amedi, A., Fregni, F., & Merabet, L. B. (2005). The plastic human brain cortex. *Annual Review of Neuroscience*, 28, 377-401.
- 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 and Hearing*, 39(6), 1224-1231.
- 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.
- Perrault, T.J., Rowland, B.A., Stein, B.E., Murray, M.M., & Wallace, M.T. (2012). Chapter 15, The Organization and Plasticity of Multisensory Integration in the Midbrain. In: *The Neural Bases of Multisensory Processes*. Boca Raton (FL): CRC Press/Taylor & Francis.
- Perrett, S., & Noble, W. (1997a). The contribution of head motion cues to localization of low-pass noise. *Perception & Psychophysics*, 59(7), 1018-1026.
- 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.
- Perrott, D. R., Ambarsoom, H., & Tucker, J. (1987). Changes in head position as a measure of auditory localization performance : Auditory psychomotor coordination under monaural and binaural listening conditions. *The Journal of the Acoustical Society of America*, 82(5), 1637-1645.

- Petersen, B., Weed, E., Sandmann, P., Brattico, E., Hansen, M., Sørensen, S. D., & Vuust, P. (2015). Brain responses to musical feature changes in adolescent cochlear implant users. *Frontiers in Human Neuroscience, 9*, 7.
- Polonenko, M. J., Papsin, B. C., & Gordon, K. A. (2015). The effects of asymmetric hearing on bilateral brainstem function: Findings in children with bimodal (electric and acoustic) hearing. *Audiology & Neuro-Otology, 20 Suppl 1*, 13-20.
- Ponton, C. W., Eggermont, J. J., Kwong, B., & Don, M. (2000). Maturation of human central auditory system activity: Evidence from multi-channel evoked potentials. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology, 111*(2), 220-236.
- Populin, L. C. (2008). Human sound localization: Measurements in untrained, head-unrestrained subjects using gaze as a pointer. *Experimental Brain Research, 190*(1), 11-30.
- Punte, A.K., De Bodt, M., & Van De Heyning, P. (2014). Long-term improvement of speech perception with the fine structure processing coding strategy in cochlear implants. *ORL, 76*, 36-43.
- Purves, A., Fitzpatrick, H.L., & McNamara, W. (2005). Le système auditif. *Neurosciences. ed. De Boeck*, p. 303-309.
- Querleu, D., Renard, X., Versyp, F., Paris-Delrue, L., & Crèpin, G. (1988). Fetal hearing. *European Journal of Obstetrics, Gynecology, and Reproductive Biology, 28*(3), 191-212.
- Rakerd, B., & Hartmann, W. M. (1986). Localization of sound in rooms, III: Onset and duration effects. *The Journal of the Acoustical Society of America, 80*(6), 1695-1706.
- Ramakers, G. G. J., Smulders, Y. E., van Zon, A., Kraaijenga, V. J. C., Stegeman, I., Van Zanten, G. A., & Grolman, W. (2016). Agreement between health utility instruments in cochlear implantation. *Clinical Otolaryngology: Official Journal of ENT-UK; Official Journal of Netherlands Society for Oto-Rhino-Laryngology & Cervico-Facial Surgery, 41*(6), 737-743.
- Recanzone, G. H. (2009). Interactions of auditory and visual stimuli in space and time. *Hearing Research, 258*(1-2), 89-99.
- Recanzone, G. H., Makhamra, S. D., & Guard, D. C. (1998). Comparison of relative and absolute sound localization ability in humans. *The Journal of the Acoustical Society of America, 103*(2), 1085-1097.
- Richard, C., Pautz, A., Tingralli, S., & Moulin, A. (2014). Évaluation d'une version française de 2 questionnaires d'audition spatiale (Le SSQ et le SHQ). *Annales françaises d'Oto-rhino-laryngologie et de Pathologie Cervico-faciale, 131*(4, Supplement), A12-A13.

- Risoud, M., Hanson, J.-N., Gauvrit, F., Renard, C., Lemesre, P.-E., Bonne, N.-X., & Vincent, C. (2018). Sound source localization. *European Annals of Otorhinolaryngology, Head and Neck Diseases*, *135*(4), 259-264.
- Roffler, S. K., & Butler, R. A. (1968). Factors that influence the localization of sound in the vertical plane. *The Journal of the Acoustical Society of America*, *43*(6), 1255-1259.
- Rouger, J., Lagleyre, S., Fraysse, B., Deneve, S., Deguine, O., & Barone, P. (2007). Evidence that cochlear-implanted deaf patients are better multisensory integrators. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(17), 7295-7300.
- Salamy, A. (1984). Maturation of the auditory brainstem response from birth through early childhood. *Journal of Clinical Neurophysiology: Official Publication of the American Electroencephalographic Society*, *1*(3), 293-329.
- Salloum, C. A. M., Valero, J., Wong, D. D. E., Papsin, B. C., van Hoesel, R., & Gordon, K. A. (2010). Lateralization of interimplant timing and level differences in children who use bilateral cochlear implants. *Ear and Hearing*, *31*(4), 441-456.
- Schäfer, E., Vedoveli, A.-E., Righetti, G., Gamedinger, P., Knipper, M., Tropitzsch, A., Karnath, H.-O., Braun, C., & Li Hegner, Y. (2021). Activities of the Right Temporo-Parieto-Occipital Junction Reflect Spatial Hearing Ability in Cochlear Implant Users. *Frontiers in Neuroscience*, *15*, 613101.
- Schorr, E. A., Fox, N. A., van Wassenhove, V., & Knudsen, E. I. (2005). Auditory-visual fusion in speech perception in children with cochlear implants. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(51), 18748-18750.
- Seeber, B. U., & Fastl, H. (2008). Localization cues with bilateral cochlear implants. *The Journal of the Acoustical Society of America*, *123*(2), 1030-1042.
- SFORL - Société Française d'Oto-Rhino-Laryngologie et de Chirurgie de la Face et du Cou. (2018). Indications de l'implant cochléaire chez l'adulte et chez l'enfant. *Recommandation pour la pratique clinique*.
- Sharma, A., Dorman, M. F., & Spahr, A. J. (2002). A sensitive period for the development of the central auditory system in children with cochlear implants: Implications for age of implantation. *Ear and Hearing*, *23*(6), 532-539.
- Shelton, B. R., & Searle, C. L. (1980). The influence of vision on the absolute identification of sound-source position. *Perception & Psychophysics*, *28*(6), 589-596.
- 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. *European Annals of Otorhinolaryngology, Head and Neck Diseases*, 136(5), 385-391.
- Sparreboom, M., Snik, A. F. M., & Mylanus, E. A. M. (2012). Sequential bilateral cochlear implantation in children: Quality of life. *Archives of Otolaryngology--Head & Neck Surgery*, 138(2), 134-141.
- Stacey, P. C., & Summerfield, A. Q. (2007). Effectiveness of computer-based auditory training in improving the perception of noise-vocoded speech. *The Journal of the Acoustical Society of America*, 121(5 Pt1), 2923-2935.
- Stacey, P. C., & Summerfield, A. Q. (2008). Comparison of word-, sentence-, and phoneme-based training strategies in improving the perception of spectrally distorted speech. *Journal of Speech, Language, and Hearing Research: JSLHR*, 51(2), 526-538.
- 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. *Scientific Reports*, 9(1), 18284.
- Steel, M. M., Papsin, B. C., & Gordon, K. A. (2015). Binaural fusion and listening effort in children who use bilateral cochlear implants: A psychoacoustic and pupillometric study. *PloS One*, 10(2), e0117611.
- Stein, B. E., Stanford, T. R., & Rowland, B. A. (2014). Development of multisensory integration from the perspective of the individual neuron. *Nature Reviews. Neuroscience*, 15(8), 520-535.
- Stevens, S.S., & Newman, E.B. (1936). The localization of actual sources of sound. *Am JPsychol*, 48,297.
- 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 and Hearing*, 20(2), 104-116.
- Strelnikov, K., Marx, M., Lagleyre, S., Fraysse, B., Deguine, O., & Barone, P. (2015). PET-imaging of brain plasticity after cochlear implantation. *Hearing Research*, 322, 180-187.
- Strelnikov, K., Rosito, M., & Barone, P. (2011). Effect of audiovisual training on monaural spatial hearing in horizontal plane. *PloS One*, 6(3), e18344.
- Stropahl, M., & Debener, S. (2017). Auditory cross-modal reorganization in cochlear implant users indicates audio-visual integration. *NeuroImage. Clinical*, 16, 514-523.
- Study protocol for the World Health Organization project to develop a Quality of Life assessment instrument (WHOQOL). (1993). *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation*, 2(2), 153-159.

- Svirsky, M. A., Teoh, S.-W., & Neuburger, H. (2004). Development of language and speech perception in congenitally, profoundly deaf children as a function of age at cochlear implantation. *Audiology & Neuro-Otology*, *9*(4), 224-233.
- 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*, 932.
- 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.
- Tillein, J., Hubka, P., Syed, E., Hartmann, R., Engel, A. K., & Kral, A. (2010). Cortical representation of interaural time difference in congenital deafness. *Cerebral Cortex (New York, N.Y.: 1991)*, *20*(2), 492-506.
- Todd, A. E., Goupell, M. J., & Litovsky, R. Y. (2017). The Relationship Between Intensity Coding and Binaural Sensitivity in Adults With Cochlear Implants. *Ear and Hearing*, *38*(2), e128-e141.
- Tonelli, A., Brayda, L., & Gori, M. (2015). Task-dependent calibration of auditory spatial perception through environmental visual observation. *Frontiers in Systems Neuroscience*, *9*, 84.
- Trumpp, N. M., & Kiefer, M. (2018). Functional reorganization of the conceptual brain system after deafness in early childhood. *PloS One*, *13*(7), e0198894.
- Truy, E., Lina-Granade, G., Apruzesse, H., Herrmann, R., Coudert, A., Gallégo, S., Flamens-Sevain, B., Ionescu, E., Ayari-Khalfallah, S., & Thai-Van, H. (2017). Implantation cochléaire de l'adulte et de l'enfant. *EMC - Oto-rhino-laryngologie*, *12*(2), 1-13.
- Tyler, R. S., Perreau, A. E., & Ji, H. (2009). The Validation of the Spatial Hearing Questionnaire. *Ear and hearing*, *30*(4), 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. *Journal of the American Academy of Audiology*, *21*(6), 390-403.
- 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. *Current Research in Behavioral Sciences 1*, 100003.

- Van Barneveld, D. C. P. B. M., Van Grootel, T. J., Alberts, B., & Van Opstal, A. J. (2011). The effect of head roll on perceived auditory zenith. *Experimental Brain Research*, 213(2-3), 235-243.
- 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.
- 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.
- 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. *Audiology & Neuro-Otology*, 15(1), 7-17.
- Van Deun, L., van Wieringen, A., Van den Bogaert, T., Scherf, F., Offeciers, F. E., Van de Heyning, P. H., Desloovere, C., Dhooge, I. J., Deggouj, N., De Raeve, L., & Wouters, J. (2009). Sound localization, sound lateralization, and binaural masking level differences in young children with normal hearing. *Ear and Hearing*, 30(2), 178-190.
- Van Wanrooij, M. M., & Van Opstal, A. J. (2005). Relearning sound localization with a new ear. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 25(22), 5413-5424.
- 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.
- van Zon, A., Smulders, Y. E., Stegeman, I., Ramakers, G. G. J., Kraaijenga, V. J. C., Koenraads, S. P. C., Zanten, G. A. V., Rinia, A. B., Stokroos, R. J., Free, R. H., Frijns, J. H. M., Huinck, W. J., Mylanus, E. A. M., Tange, R. A., Smit, A. L., Thomeer, H. G. X. M., Topsakal, V., & Grolman, W. (2017). Stable benefits of bilateral over unilateral cochlear implantation after two years : A randomized controlled trial. *The Laryngoscope*, 127(5), 1161-1168.
- 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. *2019 International Conference on 3D Immersion (IC3D)*, 1-8.
- Vermeire, K., Punte, A. K., & Van de Heyning, P. (2010). Better speech recognition in noise with the fine structure processing coding strategy. *ORL; Journal for Oto-Rhino-Laryngology and Its Related Specialties*, 72(6), 305-311.

- Verschuur, C. A., Lutman, M. E., Ramsden, R., Greenham, P., & O'Driscoll, M. (2005). Auditory localization abilities in bilateral cochlear implant recipients. *Otology & Neurotology: Official Publication of the American Otological Society, American Neurotology Society [and] European Academy of Otology and Neurotology*, 26(5), 965-971.
- Wallach, H. (1938). On Sound Localization. *J Acoust Soc Am* 10, 83–83.
- Wallach, H. (1940). The role of head movements and vestibular and visual cues in sound localization. *J Experimental Psychol*, 27: 339 68
- Wang, S., Lin, M., Sun, L., Chen, X., Fu, X., Yan, L., Li, C., & Zhang, X. (2020). Neural Mechanisms of Hearing Recovery for Cochlear-Implanted Patients: An Electroencephalogram Follow-Up Study. *Frontiers in Neuroscience*, 14, 624484.
- Warren, D.H., (1970). Intermodality interactions in spatial localization. *Cogn. Psychol.* 1 (2), 114–133.
- Weinberger, N. M. (2004). Specific long-term memory traces in primary auditory cortex. *Nature Reviews. Neuroscience*, 5(4), 279-290.
- WHO QOL Group. (1993). Study protocol for the World Health Organization project to develop a Quality of Life assessment instrument (WHOQOL). *Qual Life Res*, 2,153–9.
- WHO. (2018). Addressing the prevalence of hearing loss. World Health Organization. ISBN 978-92-4-155026-0.
- Wiebe, S., Guyatt, G., Weaver, B., Matijevic, S., & Sidwell, C. (2003). Comparative responsiveness of generic and specific quality-of-life instruments. *Journal of Clinical Epidemiology*, 56(1), 52-60.
- Wightman, F. L., & Kistler, D. J. (1989). Headphone simulation of free-field listening. II : Psychophysical validation. *The Journal of the Acoustical Society of America*, 85(2), 868-878.
- 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.
- Wright, B. A., & Zhang, Y. (2009). A review of the generalization of auditory learning. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 364(1515), 301-311.
- Yost, W. A., & Zhong, X. (2014). Sound source localization identification accuracy : Bandwidth dependencies. *The Journal of the Acoustical Society of America*, 136(5), 2737-2746.

- Zahorik, P. (2002). Assessing auditory distance perception using virtual acoustics. *The Journal of the Acoustical Society of America*, *111*(4), 1832-1846.
- Zahorik, P., Bangayan, P., Sundareswaran, V., Wang, K., & Tam, C. (2006). Perceptual recalibration in human sound localization : Learning to remediate front-back reversals. *The Journal of the Acoustical Society of America*, *120*(1), 343-359.
- 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. *American Journal of Audiology*, *24*(3), 384-397.
- Zheng, Y., Godar, S. P., & Litovsky, R. Y. (2015). Development of Sound Localization Strategies in Children with Bilateral Cochlear Implants. *PloS One*, *10*(8), e0135790.
- Zhong, X., & Yost, W. A. (2017). How many images are in an auditory scene? *The Journal of the Acoustical Society of America*, *141*(4), 2882.
- Zirn, S., Arndt, S., Aschendorff, A., Laszig, R., & Wesarg, T. (2016). Perception of Interaural Phase Differences With Envelope and Fine Structure Coding Strategies in Bilateral Cochlear Implant Users. *Trends in Hearing*, *20*, 2331216516665608.
- Zonooz, B., Arani, E., & Opstal, A. J. V. (2018). Learning to localise weakly-informative sound spectra with and without feedback. *Scientific Reports*, *8*(1), 17933.

ANNEXES

Autres publications en rapport avec les travaux de thèse

- Annexe A : Validation du système de réalité virtuelle VIVE comme outil de localisation spatiale

- Annexe B : SPHERE – évaluation des performances de localisation spatiale en 3D chez des adultes normo-entendant et implantés cochléaires

Annexe A : Validation du système de réalité virtuelle VIVE comme outil de localisation spatiale

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 (9mm static, 9.4mm 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. 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 [3]. 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 [4].

With this tracking system, we can expect HTC Vive to have a submillimeter jitter and around the cm accuracy in a

tracking volume of $5 \times 5 \times 2 \text{ m}^3$ for the HMD [2,5]. 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.

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.

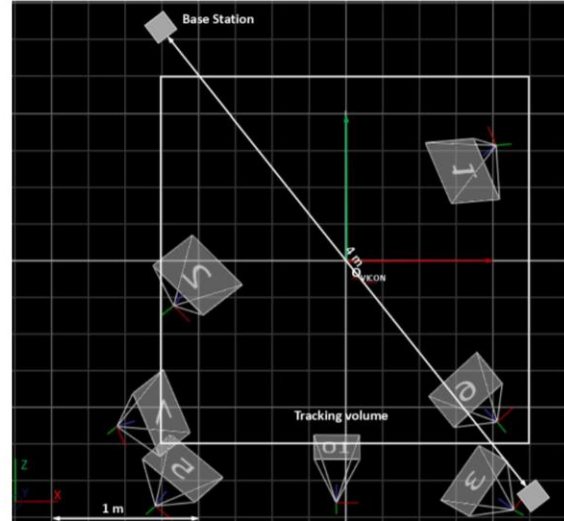


Fig. 2. Tracking volume and position of BONITA 10 and Base stations.

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.

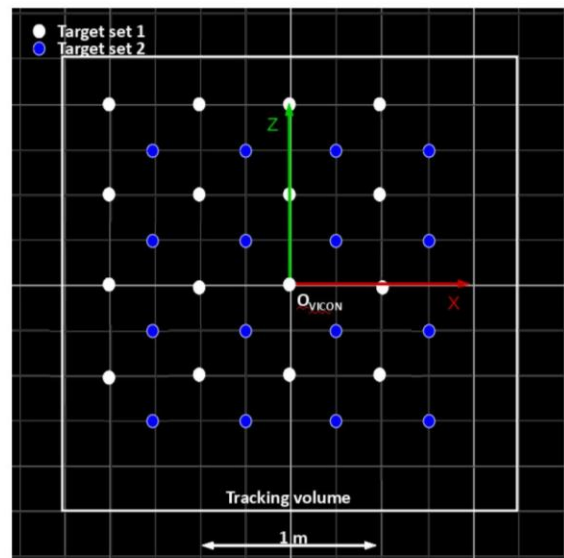


Fig. 3. Target positions in the tracking volume for the static analysis. We represented 2 different sets of static points.

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:

$$Eq. 1: \quad Dispersion = \frac{\sum_{i=1}^N \|p_i - centroid(p)\|}{N}$$

To estimate the 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:

$$Eq. 2: \quad p'_i = Rp_i + T + N_i$$

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. [7], 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 $\{p'_i\}$:

$$Eq. 3: \quad \Sigma^2 = \sum_{i=1}^N \|p'_i - (Rp_i + T)\|^2$$

To minimize least-square, we computed the matrix 3×3 H as “Eq. 4”:

$$Eq. 4: \quad H = \sum_{i=1}^N q_i q_i^t$$

$$q_i = p_i - centroid(p)$$

$$q'_i = p'_i - centroid(p')$$

Then we calculated the singular value decomposition (SVD) “Eq. 5”:

$$Eq. 5: \quad H = UAV^t$$

Finally, the R matrix was determined according to the determinant of VU^t “Eq. 6”, we found the vector T as seen above:

$$Eq. 6: \quad \begin{cases} \text{If } \det(VU^t) = +1, \text{ then } R = VU^t \\ \text{If } \det(VU^t) = -1, \text{ then } R = V'U^t \end{cases}$$

$$\text{with } V' = [v_1, v_2, -v_3]$$

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:

$$[V'_x; V'_y; V'_z] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \rightarrow [V_x, V_y, V_z] = \begin{bmatrix} V_{x1} & V_{y1} & V_{z1} \\ V_{x2} & V_{y2} & V_{z2} \\ V_{x3} & V_{y3} & V_{z3} \end{bmatrix}$$

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:

$$M = \begin{bmatrix} V_{x1} & V_{y1} & V_{z1} \\ V_{x2} & V_{y2} & V_{z2} \\ V_{x3} & V_{y3} & V_{z3} \end{bmatrix}$$

$$T = P_0 M - P'_0$$

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.

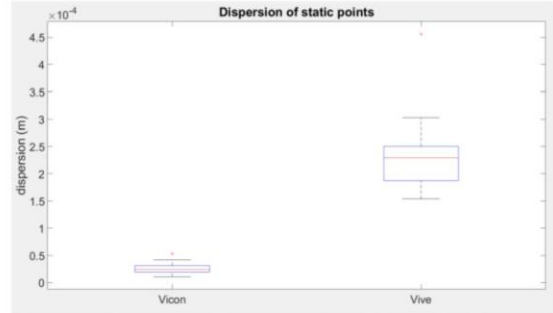


Fig. 4. Dispersion of static points. N = 14; dispersion VICON = 0.027 +/- 0.011 mm; dispersion Vive = 0.237 +/- 0.076 mm.

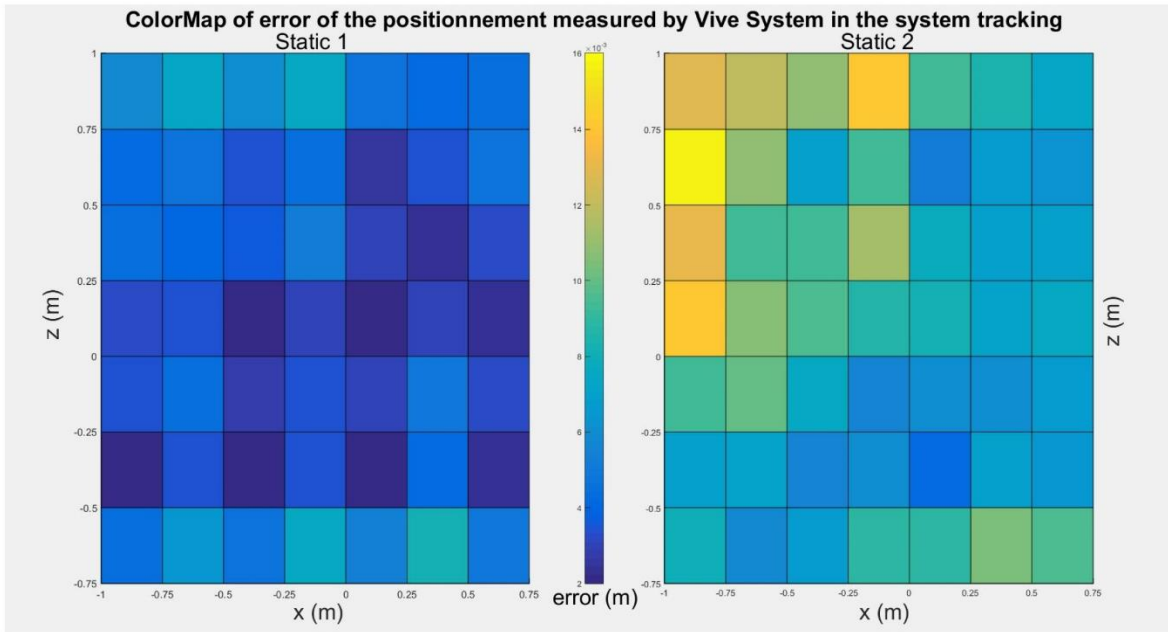


Fig. 5. Color map of error measured between Vive and VICON system to the position of object tracked across the tracking volume. Static 1 shows the fluctuation of the error as measured with method 1. The nearer the target to the center the greater the accuracy. Static 2 shows the fluctuation of the error measured with the method 2. Accuracy decreases around the left top corner.

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.

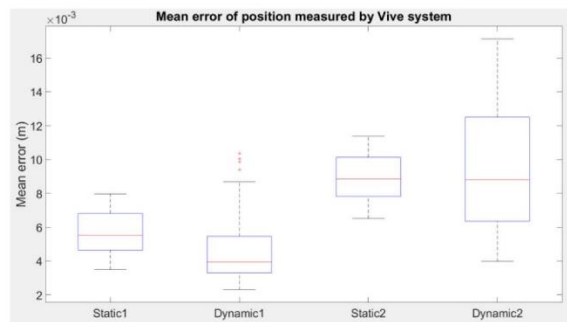


Fig. 6. Box plot of the mean error of position measured by the Vive system compared with the VICON position for both methods and for static and dynamic conditions. Static 1 represents the mean error of each trial in static experience using method 1 to change the reference frame, N = 14; error = 5.7 +/- 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 = 9 +/- 1.5 mm. Dynamic 2 represents the mean error of each trial in dynamic experience using method 2 to change the reference frame, N = 26; error = 9.4 +/- 4 mm.

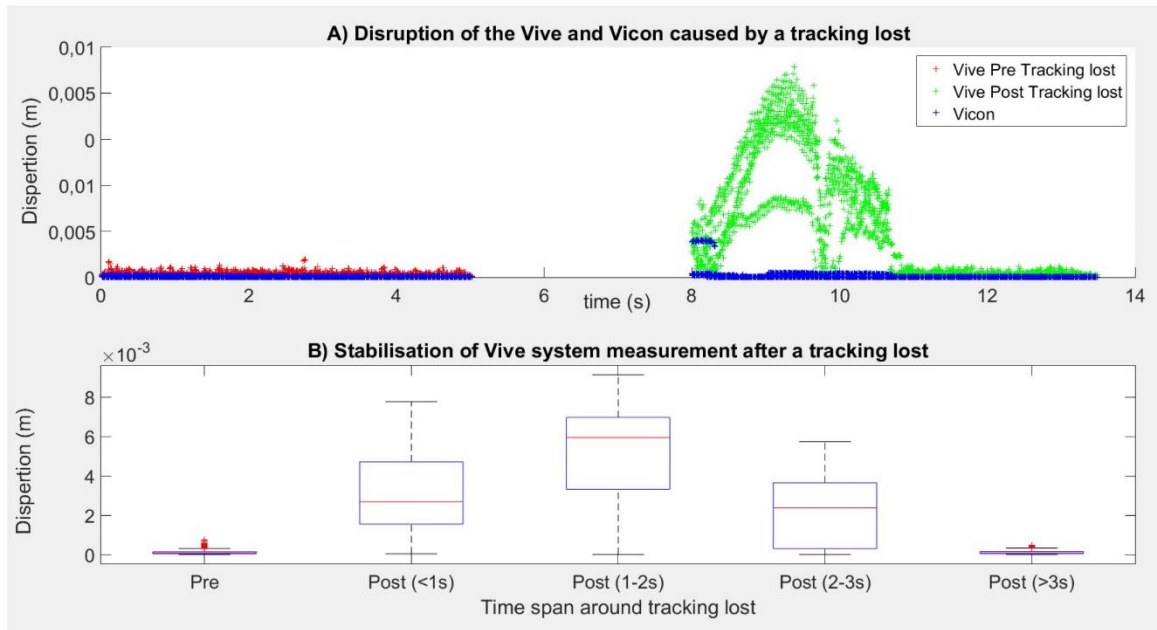


Fig. 7. Representation of the impact of a tracking lost on the Vive measure. **A.** Disruption of the Vive and the VICON system caused by the tracking lost. The blue crosses show that VICON is not disrupted by the tracking lost whereas the green crosses show the Vive tracking system after a tracking 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.

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 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 [2,5].

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. For a non-research purpose device, the Vive has indeed a well acceptable precision. 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. 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. 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.
 - [3] A. Malventano. SteamVR HTC Vive In-depth – Lighthouse Tracking System Dissected and Explored. *PC Perspective*. 2016.
 - [4] Skarredghost. All you need to know about SteamVR Tracking 2.0... will it be the foundation of Vive 2? *The Ghost Howls*. 2017.
 - [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. Bouteau, P. Vasseur, and X. Savatier. A study of Vicon system positioning performance. *Sensors*, 17:1591-1609. July 2017.
 - [8] A.B. Smith, C.D. Jones, and E.F. Roberts. Article Title. *Journal*, 62:291-294, January 1920.
- New refs from the discussion to be added to the ref list:
- [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. Høring, 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).

Annexe B

SPHERE: A novel approach to 3D and active sound localization

V. GAVEAU^{1 2}, A. COUDERT^{1 2 6}, R. SALEMME^{1 2 3}, E. KOUN^{1 2}, C. DESOCHE^{1 2 3}, E. TRUY^{1 2 5 6}, A. FARNE^{1 2 3}, F. PAVANI^{1 2 4}

doi: <https://doi.org/10.1101/2020.03.19.998906>

in bioRxiv

1 Integrative Multisensory Perception Action & Cognition Team—ImpAct, Lyon Neuroscience Research Center, INSERM U1028, CNRS U5292, 16 Avenue Doyen Lepine, Bron 69676, France*

2 University of Lyon 1, 43 Boulevard du 11 Novembre 1918, Villeurbanne 69100, France*

3 Hospices Civils de Lyon, Neuro-immersion Platform, 16 Avenue Doyen Lepine, Bron 69676, France*

4 Center for Mind/Brain Sciences, University of Trento, 31 Corso Bettini, Rovereto 38068, Italy*

5 ENT Department, Herriot Hospital, Hospices Civils de Lyon, 5 place d'Arsonval, 69003 Lyon, France*

6 Pediatric ENT Department, Woman Mother Children, Hospital, Hospices Civils de Lyon, 59 Bvd Pinel, 69500 Bron, France*

Corresponding author:

V. GAVEAU

Integrative Multisensory Perception Action & Cognition Team—ImpAct, Lyon Neuroscience Research Center, INSERM U1028, CNRS U5292, 16 Avenue Doyen Lepine, Bron 69676, France

+33 4 72 91 34 21

valerie.gaveau@inserm.fr

Running head: ACTIVE HEARING 3D

Abstract

In everyday life, localizing a sound source in free-field entails more than the sole extraction of monaural and binaural auditory cues to define its location in the three-dimensions (azimuth, elevation and distance). In spatial hearing, we also take into account all the available visual information (e.g., cues to sound position, cues to the structure of the environment), and we resolve perceptual ambiguities through active listening behavior, exploring the auditory environment with head or/and body movements. Here we introduce a novel approach to sound localization in 3D named SPHERE (European patent n° WO2017203028A1), which exploits a commercially available Virtual Reality Head-mounted display system with real-time kinematic tracking to combine all of these elements (controlled positioning of a real sound source and recording of participants' responses in 3D, controlled visual stimulations and active listening behavior). We prove that SPHERE allows accurate sampling of the 3D spatial hearing abilities of normal hearing adults, and it allowed detecting and quantifying the contribution of active listening. Specifically, comparing static vs. free head-motion during sound emission we found an improvement of sound localization accuracy and precisions. By combining visual virtual reality, real-time kinematic tracking and real-sound delivery we have achieved a novel approach to the study of spatial hearing, with the potentials to capture real-life behaviors in laboratory conditions. Furthermore, our new approach also paves the way for clinical and industrial applications that will leverage the full potentials of active listening and multisensory stimulation intrinsic to the SPHERE approach for the purpose rehabilitation and product assessment.

Keywords: SPATIAL HEARING, VIRTUAL REALITY, HEAD-MOVEMENTS, EYE-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 behaviour, 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 estimate azimuth, elevation and distance of sounds. We perceive the visual context in which they occur and often also the event that generated them (Kumpik et al. 2019). Most importantly, to resolve perceptual ambiguity in sound localization, we explore the auditory environment with head and body movements (Andéol & Simpson, 2016).

When spatial hearing abilities are investigated in laboratory or clinical settings, however, most of these naturalistic aspects of spatial hearing are considerably constrained. In the laboratory, participants are typically presented with sounds that originate from a limited set of positions,

typically varying only along the azimuthal plane (i.e., all having fixed elevation and distance; Haber et al., 1993; Grantham et al., 2007). Few studies manipulated azimuth and elevation jointly (e.g. Oldfield & Parker 1984; Wightman & Kistler 1999; Ahrens et al., 2019), and even fewer have modified azimuth, elevation and distance across trials (e.g. Haber et al., 1993; Brungart et al., 1999). In addition, although sounds were delivered in 3D -- hence, supposedly perceived in 3D space -- the response was often limited to one dimension (e.g., in studies using a rotating dial to indicate angular sound position; Haber et al., 1993) or two dimensions (e.g., in studies using a remote display to indicate sound direction; Andéol et al., 2014). These approaches limit our understanding of perceived auditory space, and might overestimate sound localization performance by reducing stimulation and response uncertainty.

Multisensory contributions to sound localization are also often poorly controlled. Some studies allow full vision of the loudspeakers (e.g., van Hoesel & Tyler 2003), whereas others hide loudspeakers behind visible barriers (e.g., Nava et al., 2009; Pavani et al., 2001; Pavani et al., 2003). Yet, any visible barrier provides relevant cues for estimating sound elevation and distance. When visual cues are completely eliminated by placing subjects in total darkness (Pavani et al., 2008), the listening experience lacks entirely the multisensory context that characterizes everyday auditory environments. In addition, whenever participants are blindfolded (Ahrens et al., 2019), or are required to keep their eyes closed (Brungart et al., 1999), eye-position cannot be measured. But in natural conditions, eye-orienting responses permit encoding of sound position in retinocentric coordinates (Bulkin & Groh 2006; Pavani et al., 2008) and allow for associations between heard sounds and plausible visual sources. In addition, static and dynamic eye-position influence sound localization (e.g., Lewald & Ehrenstein 1996; Groh & Sparks 1992; Pavani et al., 2008).

Most importantly, to ensure reproducibility of sound characteristics across trials, conditions and experimental sessions, it is common practice to constrain head-movements in sound localization studies. Participants are required to place their head on a chin-rest (e.g., Brungart et al., 1999; Pavani et al., 2008) or keep their head still during each trial (Távora-Vieira et al., 2015; Litovsky et al., 2009; van Hoesel & Tyler 2003). Although the importance of head-movements for sound localization has been emphasized since the 1940s (Wallach, 1940), the contribution of spontaneous head-movement has been largely overlooked. Until recently, researchers have tracked head-movements more as a way to collect behavioural measures for sound localization accuracy (Slattery & Middlebrooks, 1994), rather than a way to study spontaneous head-movements during sound localization (for notable exceptions see Brimijoin et al., 2010; 2012).

In sum, currently dominant approaches to sound localization typically do not stimulate nor measure responses in 3D auditory space, cannot fully control for multisensory information and, crucially, do not allow for active listening through head-movements. Here, we report on a novel approach that circumvents all these limitations and allows deeper insights into spatial hearing by using spatially reproducible free-field sounds, full control over visual stimuli and unconstrained, yet real-time controlled, head-, eye- and hand-movements.

Starting from the pioneering work by Brungart and colleagues (1999), we have implemented a method to guide a real loudspeaker to precisely pre-determined head-centered coordinates in

each trial. We combined real-time tracking of the loudspeaker's position with continuously monitoring of the listener's head-position. To this aim, a head-mounted display (HMD) worn by the participant allowed to determine speaker position in each trial. Furthermore, we took advantage of the HMD to measure eye-movements, and to achieve complete control over the available visual stimulation. Using virtual reality (VR) we also instructed participants as to desired postures (e.g., straight ahead orienting at the beginning of each trial). Finally, we tracked in real-time hand motion and obtained full description of hand kinematic responses to sound source position in 3D. We provide full-detail description of this new method -- named SPHERE -- and report on one experiment testing 3D sound localization in normal hearing participants. To examine the validity of our active listening approach we compared one condition in which participants were required to remain still during the sound delivery (static listening condition), with a condition in which they were free to move (active listening condition). We predicted improved sound localization performance in the active compared to static listening (Brimijoin et al., 2010; 2012). Our results provide validation of the SPHERE system and paradigm, additionally showing the contribution of head-movements to sound localization accuracy.

METHODS

Participants

20 participants (mean age = 46, SD = 18; 12 females; 18 right-handed) were recruited through advertisement (e-mail or flyer) to take part in the study. All reported normal or corrected-to-normal vision, and no history of hearing deficits. Participants were informed that they were about to 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 was conducted in accordance with the ethical standards of the 1964 Declaration of Helsinki.

Apparatus and stimuli

A schematic description of the SPHERE apparatus is shown in Figure 1. The left side of the figure illustrates equipment in the control room, whereas the right side illustrates equipment in the testing room.

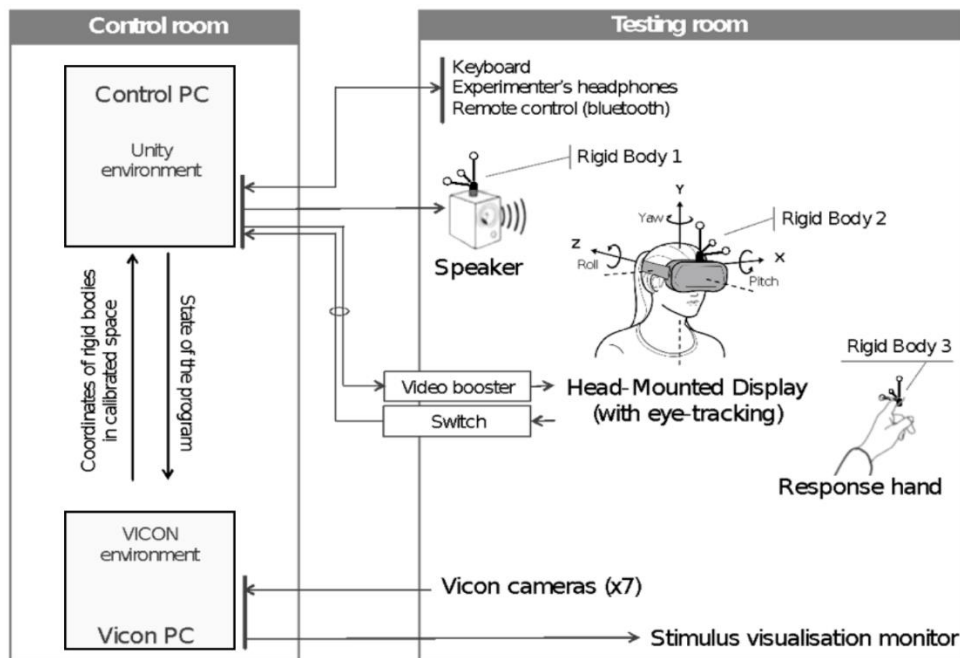


Figure 1. Schematics of the apparatus used for the study.

The control room hosted two desktop computers. The first computer, from now on ‘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 collections 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 ‘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 x64, Vicon Motion Systems LTD, Oxford, UK), and it ran a custom-made script written in Unity that served for stimulus visualization (see below).

The setting comprised Vicon cameras for motion capture, three rigid bodies for real-time object-tracking, the head-mounted display (HMD) incorporating 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 it was interfaced with the Control and Vicon PCs.

Vicon motion capture. The Vicon motion capture system comprised 7 infra-red cameras (Bonita 10: Frame rate 250 fps, Resolution 1024x1024, Vicon®, Oxford, UK) mounted on the walls of the testing room. The elevation (195-205 cm) and semi-circular 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 x 4 meters volume). Then we placed the HMD on the floor in a straight-ahead position to recorded 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 mounting 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 served for tracking 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 served for tracking HMD and the head center positions; the third rigid body (rigid body 3; radius 75 mm) served 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 x 1080 (960 x 1080 per eye), maximal refresh of 75Hz, dimensions L x W x H: 1.3 x 14.7 x 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, Berlin, Germany; www.smivision.com; 60Hz frequency and 0.5 degrees spatial precision). In our setup, the HMD served two purposes: (1) it conveyed visual instructions to the participant; (2) it allowed continuous monitoring of the participant's eye movements.

Loudspeaker. A loudspeaker (JBL GO Portable, 68.3 x 82.7 x 30.8 mm, Output Power 3.0W, 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 broad-band bursts lasting 3 seconds (see Supplementary Figure S1 for spectrograms).

A keyboard, a remote control (Targus®, Laser Presentation Remote) and a monitor (DELL 19'' 5:4, resolution 1280 x 1024), completed the equipment in the testing room. All devices were connected to the control PC, except the stimulus visualization monitor that streamed a copy of the screen of the VICON PC to be seen inside the testing room. The function of these four pieces of equipment is described in details in the procedure section below.

Procedure

Before starting the experiment, participants were introduced to the task and to the virtual reality equipment using a picturized information sheet and a custom-made video that showed one experimenter wearing the HMD and performing several sound localization trials. It was made explicit to all participants that their task was to listen carefully to the sound and indicate its location in space using the pointer held in their hand (i.e., rigid body) after the sound terminated.

Participants were told that sounds could be delivered everywhere in the 3D space around them (e.g., above and below ear level, in front and back space, at multiple azimuthal positions). They were also informed that all sounds would be delivered within their arm reaching distance. They were also told they would perform the sound localization task under two conditions. One (static listening) in which they would have to keep their head still in the initial position throughout sound presentation and another condition (active listening) in which they were prompted to actively search for the sound during its presentation, by freely moving their head. In both

conditions, they were free to move their head and body as soon as the sound terminated (e.g., rotation on the chair to reach for sounds in back space). After receiving the instructions, participants were conducted in the testing room, seated on a rotating office chair (the center of the head being aligned with the center of the chair rotation axis), and prepared for the task with the HMD and the hand-held rigid body.

The experiment began with eye- and head-center calibrations: (i) Eye-calibration was performed using a 5-points calibration grid (smart recorder of SMI Eye tracking software) and it 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 origin for the polar coordinate system that included speaker, hand, head and cyclopean gaze positions. Both eye and head calibrations were carried out each time the HMD was displaced (e.g., when participant took breaks during the experimental session).

The loudspeaker position in 3D space was calculated 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. Twelve pre-predetermined positions were used throughout the experiment, resulting from the combination of 4 different azimuths (-30° , 30° , -150° or 150°), 3 different depths (35 cm, 55 cm or 75 cm) and a single elevation (0° , i.e., ear-level). Figure 2 shows a bird-eye view of the 12 positions around the participant's head.

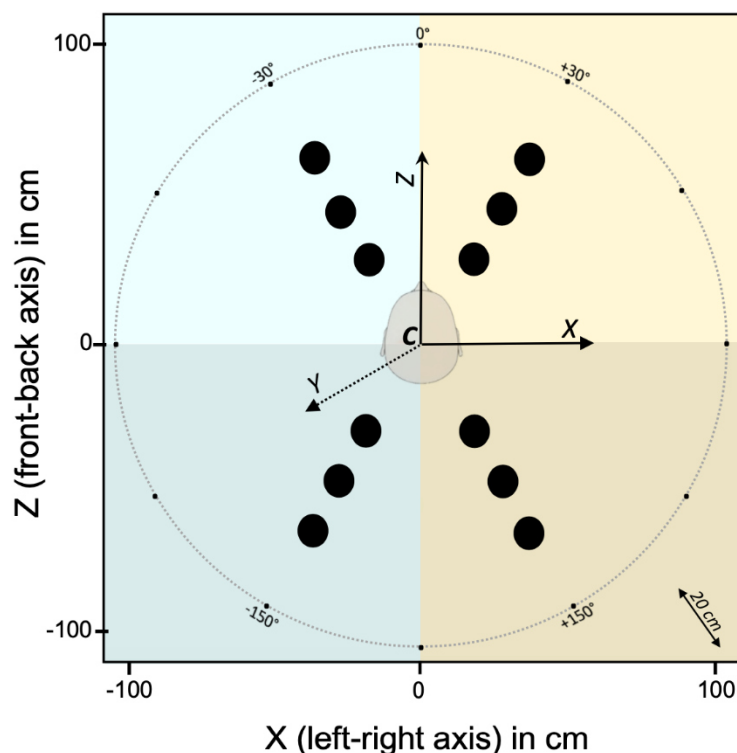


Figure 2. Bird-eye view of all pre-determined loudspeaker positions in head-centered coordinates. The origin of the polar reference-frame used for sound positioning was the participant head-center, measured in real-time (marked by C in the image). The 3D axes of this reference frame were the azimuth (X), the elevation (Y) and the depth (Z). Twelve pre-

determined loudspeaker locations (white circles) were arranged around the participant's head: at 4 different azimuths (X-values: $\pm 30^\circ$ and $\pm 150^\circ$), 3 different depths (Z-values: 35, 55 or 75 cm), all at ear-level (Y-values: 0°).

In each trial, two sets of instructions generated in real-time by the computer informed the experimenter as to where to position the loudspeaker in the 3D space surrounding the participant (i.e., one of the 12 predetermined positions detailed above). The stimulus visualization monitor (see Figure 1) displayed in real-time the actual position of the loudspeaker and its desired position for the upcoming trial. This allowed the experimenter to rapidly reach for each computer-determined position, keeping the speaker at ear-level. Information about the precise elevation positioning of the speaker was given to the experimenter via one in-ear headphone (i.e., non-audible by the participant) and consisted in an echo radar sound, which increased in frequency and intensity as the speaker approached the target position. The system software considered the loudspeaker to be correctly positioned when it entered a sphere of 5 cm diameter centered on the pre-determined sound position and delivered the sound only when three criteria were concurrently 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 actively complied with criterion 2 and 3 by aiming their head and eyes to align two crosshairs displayed in the HMD. Figure 3 shows the instructions visualized in the HMD (bottom part of each panel) and a cartoon of the participant's corresponding head- and eye-posture (upper part of each panel), in the phases that preceded sound delivery. As shown in Figure 3A, at the beginning of each trial two white crosshairs were presented to the participant: a bold cross with a ball inside, indicating the desired position of the head and eyes, respectively. A thinner cross provided participant with visual feedback of their actual head-position. Participants were instructed to move their head to align the two crosshairs. When the alignment was achieved, the bold cross turned blue (Figure 3B). Likewise, participants were instructed to gaze the ball inside the bold cross. When fixation was achieved, the ball turned blue (Figure 3C). Once all the three criteria were met, all visual stimuli disappeared (i.e., the HMD display turned black) and sound emission started (Figure 3D). Participants were instructed to respond only after sound emission ended, by bringing the hand-held rigid body where they perceived the sound to originate and hold it there for two seconds. The experimenter validated this position by pressing a button on the remote control, which also terminated the trial. After trial completion, no feedback on performance was ever provided. Each trial lasted approximately 10-15 seconds, with the speaker positioning phase lasting about 3-5 seconds, depending on the pre-determined position. Recall that multiple aspects contributed to trial duration: events before sound delivery (the participant actively moved head and eyes to the desired initial posture, the experimenter manually brought the loudspeaker to the pre-determined position), the sound delivery itself, and the participant's full-body motor response to the target after sound delivery.

The experimental session was organized in 4 successive blocks, with a pause between each block in which the HMD was removed. 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 comprised 48 randomized trials (i.e., 4 trials for each of the 12 pre-determined positions), resulting in a total of 192 trials (i.e., 8 trials for each pre-determined position in each listening condition). The entire experimental protocol lasted approximately 40 minutes.

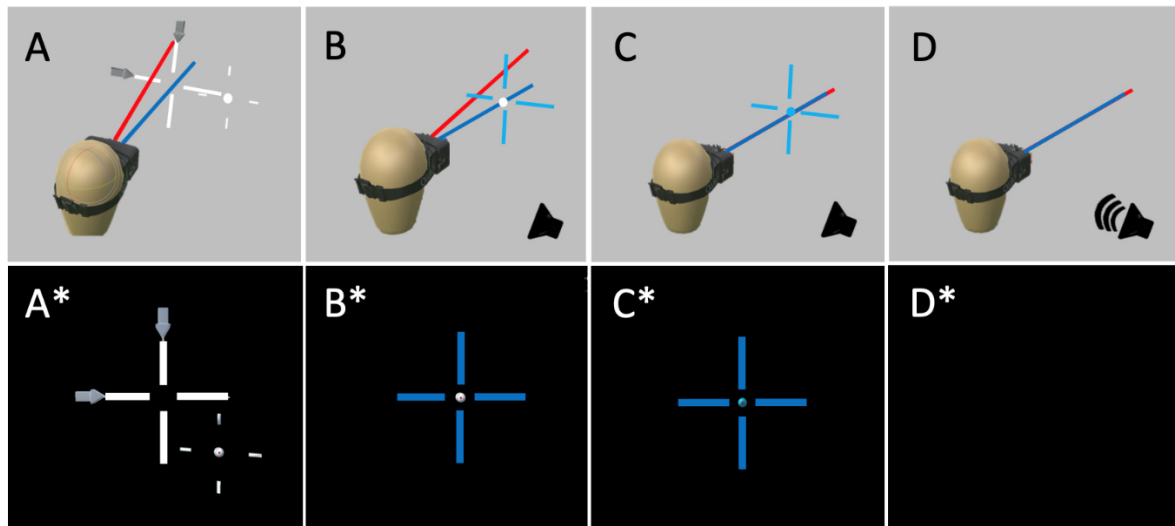


Figure 3. Pre-stimulation alignment of head and eyes. (A) At the beginning of each trial, the participant wearing the HMD was free to move his head (symbolized head direction: blue line) and eyes (symbolized cyclopean eye direction: red line). (A) In the HMD display, the participant saw a bold white cross (indicating the actual position of the head), and a thin white cross with a central ball (indicating the desired position of the head and eyes, respectively). Two grey arrows flanked the bold white cross, showing the participant in which direction to move his head to achieve the desired initial position (in this example, the participant had to move his head to the right and down to ensure alignment). (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 (i.e., head-position straight ahead, eye-position straight ahead, and speaker within a sphere around the pre-determined position) all visual stimulations were removed, (D*) the scene became entirely dark, and the sound was delivered.*

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 analysis of head and hand were analyzed and inspected for each trial by a custom-made software running on MATLAB R2013a. 3D position signal was filtered (50 Hz cut-off frequency, finite impulse response filter FIR) and velocities were computed from the filtered position signal using a two-point central difference derivative algorithm (Bahill & McDonald, 1983). In order 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 (80 mm/s). The results of this automatic procedure were then inspected off-line and corrected manually, if necessary. This procedure served to obtain the spatio-temporal profile of head and hand behavior, and to extract relevant parameters for subsequent analyses (number of head movements during sound emission, onset of the first head movement, onset of the head and hand response). 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 delivery, or head movement erroneously produced in the static listening condition).

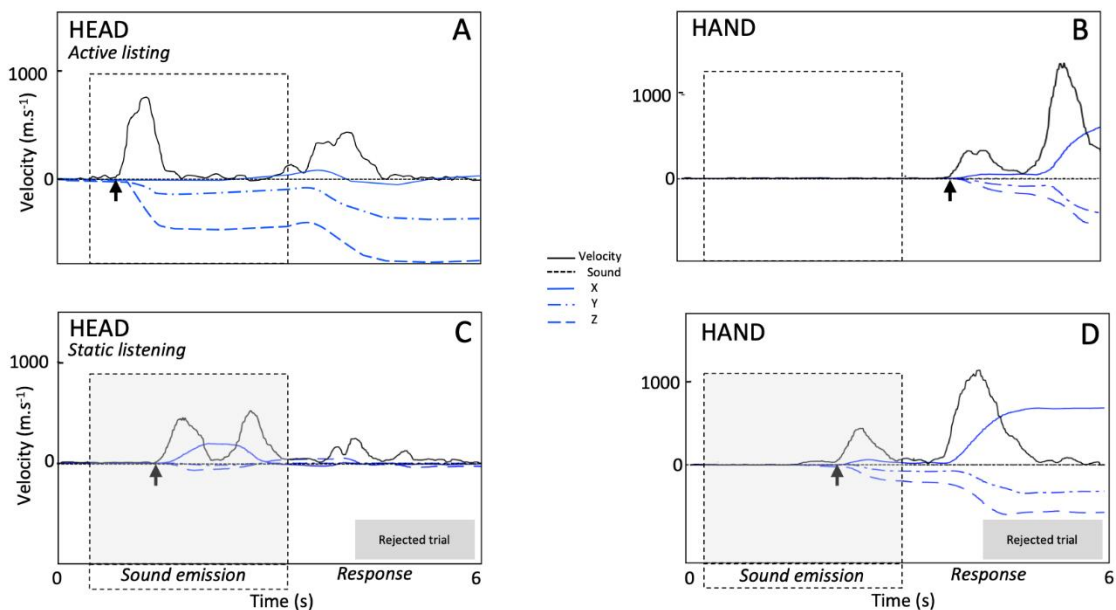


Figure 4. Description of spatio-temporal behavior of the head (A, C) and the hand (B, D), as shown by our custom-made software, during representative trials. The dotted black line shows the time of sound emission, the bold black line and the blue lines respectively shows the velocity profile and the x, y, z positions of the tracked element (head or hand). The arrow marks the beginning of the head or hand movement. Examples illustrate correctly executed trials (A and B), and rejected trials in which instructions were violated (C: head movement initiated during sound emission in a static listening trial; D, hand movement initiated during sound emission).

An example of this procedure for a correctly executed trial in the active listening condition is shown in Figure 4A and 4B, for head and hand, respectively. In this example, one single head movement is visible during sound delivery (Figure 4A), and the head and hand response movements are clearly identified in the response phase (Figures 4A and 4B).

An example of this procedure for two rejected trials is shown in Figures 4C and 4D. In Figure 4C, two head movements are visible during sound delivery in a trial which required static listening. In Figure 4D, the hand moves during sound emission.

All statistical analyses and data visualizations were performed using R and R-studio environment. Unless otherwise indicated, means \pm standard errors are reported in the text. We planned ANOVAs or t-test, and the Greenhouse-Geisser sphericity correction has been applied to analyses of variance, when appropriate. For all post-hoc comparisons, the FDR (False Discovery Rate) method implemented in R (Benjamini & Hochberg, 1995) has been adopted to correct for multiple comparisons when needed.

RESULTS

Positioning of speakers at pre-determined locations

The apparatus we developed poses minimal constraints on the participants' body movements: movements were unrestrained in all three dimensions and, after sound delivery, participants were free to rotate on the chair to explore sounds in back space. Yet, at the beginning of each trial, participants were instructed to align their head and eye with respect to straight ahead position. On-line kinematic tracking allowed to deliver sounds only when the required eye and head posture criterions were matched. Moreover, the participant's compliance with the instructions was further examined through off-line kinematic analyses using a custom-made software. Trials in which instructions were not followed 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.

The absence of physical constraints on the participants' posture poses potentially a problem of reproducibility of sound source position across experimental trials and participants. This because the head returns always to different initial positions after the response (hand-pointing to sounds delivered all around the participant). In the next paragraphs, we show how head-centered positioning of the loudspeaker effectively solved this issue.

We started by testing if the pre-determined loudspeaker locations remained constant across trials and participants. Figure 5 shows initial head position (in black) and 12 pre-determined locations (in gray) for all participants across 192 trials in VICON reference frame (Fig. 5A and C, in bird-eye and lateral view respectively). A substantial variability is observed, which however reduces dramatically when all positions are referenced to head-center, i.e., converted to a head-centered reference frame (Fig. 5B and D). Figure 5E summarizes the effect of head-referencing by comparing mean changes in standard deviation (with standard errors) across participants in VICON vs. head-centered coordinates (x: 1.08 ± 0.24 cm vs. 0.10 ± 0.02 cm; y: 0.63 ± 0.14 cm vs. 0.10 cm; depth: 1.89 ± 0.42 cm vs. 0.15 ± 0.03 cm)

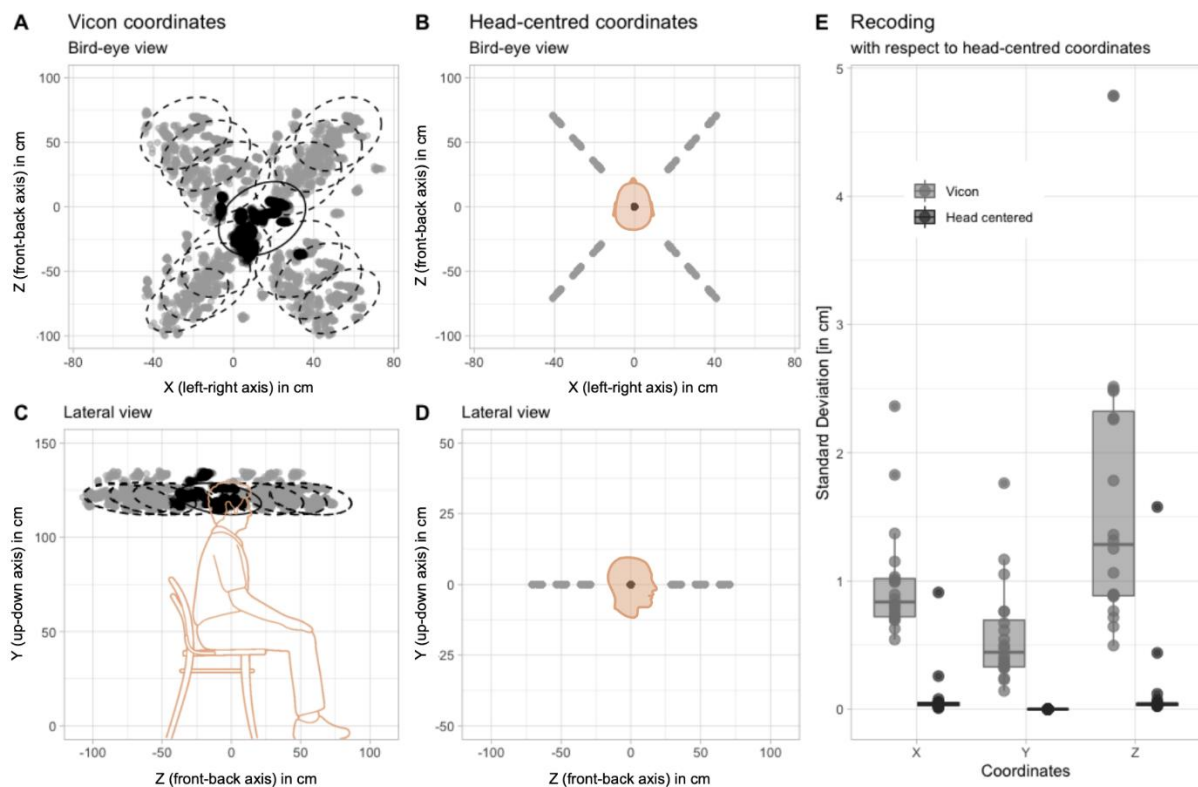


Figure 5. Normalization to head-centered coordinates. Bird-eye and lateral view of initial head position (in black) and 12 pre-determined locations (in gray) for all participants (192 trials each), in VICON reference frame (A-C) and head-centered coordinates (B-D). Variability of pre-determined locations averaged across 12 positions for each participant in the two reference frames, as a function of coordinates x, y, z (E).

Next, we tested if the variability of loudspeaker actual location around the pre-determined position was within the established tolerance (i.e., a sphere with a 5 cm diameter around the pre-determined position). Figure 6 shows all 192 stimulations for all participants, when the 12 pre-determined locations are re-aligned to a single coordinate, centered on the origin of the axes. Dashed ellipses in the 2D panels of Figure 6 represent 95% confidence interval of the distribution, and indicate that for all these trials stimulation remained within 2.5 cm from the pre-determined location (mean differences with SE between pre-determined 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). (mean differences between pre-determined and actual location: $x = 0.87 \pm 0.6$ cm, $y = 1.15 \pm 0.8$ cm, $z = 0.93 \pm 0.7$ cm).

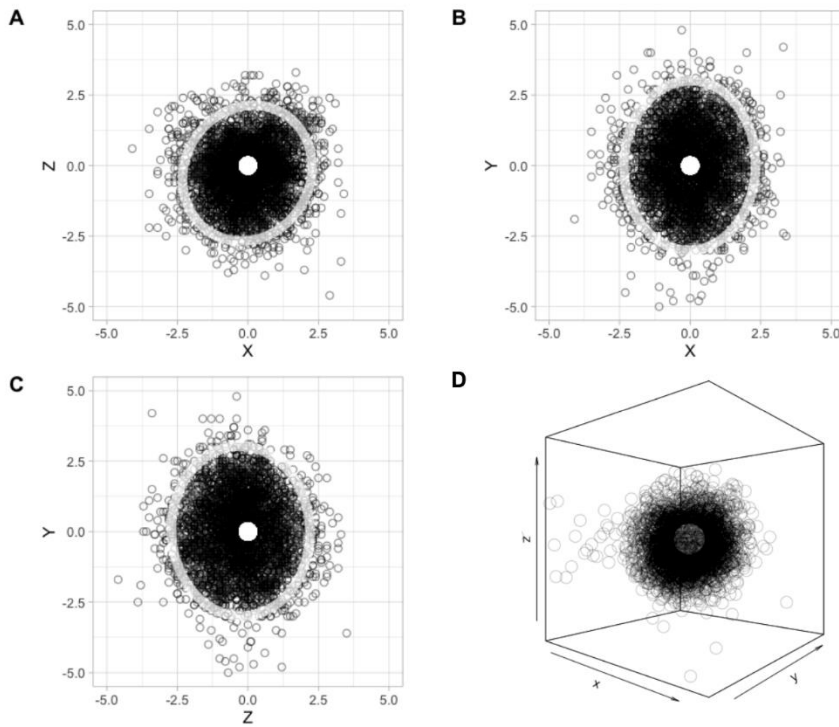


Figure 6. Actual speaker location with respect to pre-determined location, in centimeters. Stimulations delivered to all participants, when the 12 pre-determined locations are re-aligned to a single coordinate, centered on the origin of the axes. (A) top view; (B) front view; (C) lateral view; (D) 3D rendering. Ellipses in the 2D panels represent 95% confidence interval of the distribution.

Hand pointing to sounds in static listening

Sound localization performance during static listening is shown in Figure 7, for azimuth, elevation and distance separately. Dispersion plots show hand-pointing responses for all trials and participants (with 95% confidence intervals shown by ellipses), color-coded as a function of sound distance (far, middle and near). Bar plots summarize means and 95% confidence intervals for absolute and variable errors in each of the three dimensions (for performance details, see Table 1).

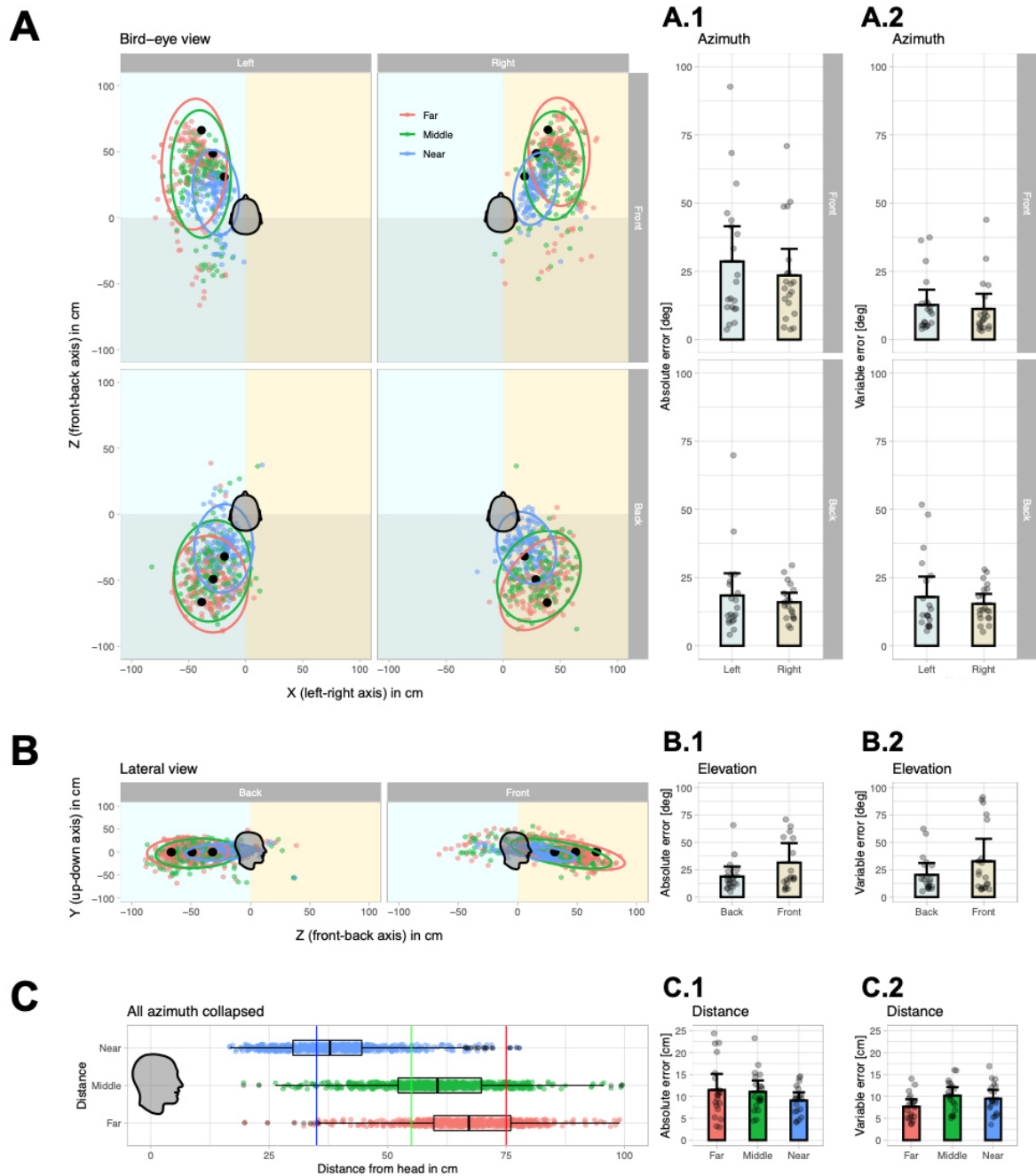


Figure 7. Sound localization performance during static listening. (A) Bird-eye view of all target positions (black dots) and hand-pointing responses (colored dots) in all trials and participants. Ellipses indicate 95% confidence interval of responses within each quadrant and distance. Responses and ellipses are color-coded as a function of target distance: red for far, green for middle, and blue for near targets. Absolute (A1) and variable (A2) error in azimuth for each participant (dots), as a function of side (left-right) and antero-posterior position (front-back) of target sounds. Error bars indicate the 95% confidence intervals, adjusted for within-subjects' designs using method from Morey (2008). (B) Lateral view of all target positions and responses of all participants. Absolute (B1) and variable (B2) error in elevation for each participant (dots) as a function of antero-posterior position of target sounds. Responses for each participant are averaged across side (left or right) and distance. (C) Lateral view of

responses (with boxplots) as a function of sound distance. Absolute (C1) and variable (C2) errors in depth for each participant (dots).

	Azimuth (in degrees)			Elevation (in degrees)			Distance (in centimetre)			3d error (in cm)
	Absolute error	Signed error	Variable error	Absolute error	Signed error	Variable error	Absolute error	Signed error	Variable error	D
Static posture										
Front	Mean (SE)	Mean (SE)	Mean (SE)	Mean (SE)	Mean (SE)	Mean (SE)	Mean (SE)	Mean (SE)	Mean (SE)	Mean (SE)
Left	28.6 (5.3)	-28.2 (5.4)	12.7 (2.3)	34.2 (8.3)	-2.9 (7.9)	30.4 (8.7)	10.0 (0.7)	-5.4 (1.6)	8.7 (0.5)	34.7 (4.3)
Right	23.5 (4.0)	23.2 (4.1)	11.2 (2.3)	28.4 (6.1)	-5.1 (7.2)	23.5 (6.1)	9.6 (0.7)	2.0 (1.7)	9.6 (0.6)	32.6 (3.8)
Mean	26.1 (4.0)	-2.4 (2.6)	29.4 (4.4)	31.4 (6.0)	-4.2 (6.9)	32.7 (6.9)	9.8 (0.6)	-1.7 (1.5)	10.2 (0.6)	34.5 (3.7)
Back										
Left	18.4 (3.4)	5.8 (4.3)	17.9 (3.1)	20.7 (5.0)	4.7 (5.7)	20.4 (4.5)	11.6 (1.4)	-3.8 (2.4)	10.4 (0.7)	27.1 (2.2)
Right	16.0 (1.4)	-6.0 (2.7)	15.4 (1.5)	17.3 (2.0)	-1.6 (4.0)	16.2 (2.2)	10.9 (0.8)	1.4 (1.8)	11.5 (0.7)	28.5 (1.8)
Mean	17.1 (2.2)	0.2 (2.2)	20.5 (2.7)	18.7 (3.1)	1.6 (4.3)	20.3 (3.6)	11.3 (1.0)	-1.4 (2.0)	11.5 (0.7)	27.9 (1.7)
Overall	21.6 (2.3)	-1.1 (2.2)	27.6 (2.9)	25.1 (3.4)	-1.3 (3.4)	35.5 (5.0)	10.5 (0.7)	-1.5 (1.6)	11.3 (0.6)	32.2 (2.3)
Active posture										
Front										
Left	18.5 (5.0)	-17.9 (5.1)	8.1 (1.3)	21.8 (6.3)	1.1 (7.7)	12.5 (1.9)	10.0 (0.8)	-2.9 (1.8)	9.2 (0.4)	27.1 (3.5)
Right	13.3 (3.0)	11.8 (3.1)	9.5 (2.4)	20.5 (4.6)	2.2 (5.5)	14.7 (4.0)	10.0 (0.9)	3.7 (1.7)	10.0 (0.6)	27.4 (3.6)
Mean	16.0 (3.8)	-3.2 (1.8)	18.3 (4.2)	21.2 (5.2)	1.6 (6.4)	15.4 (3.3)	10.0 (0.6)	0.4 (1.7)	10.4 (0.5)	27.6 (3.5)
Back										
Left	17.8 (3.1)	2.9 (4.3)	16.9 (3.5)	21.7 (4.6)	-2.7 (4.5)	21.6 (5.8)	11.6 (1.4)	-5.0 (2.4)	10.6 (0.8)	28.0 (2.3)
Right	15.1 (2.0)	-6.4 (3.1)	12.3 (1.3)	17.0 (2.3)	-4.5 (3.8)	16.0 (2.7)	10.0 (0.8)	0.3 (1.7)	10.3 (0.7)	27.9 (2.1)
Mean	16.7 (2.4)	-1.6 (2.3)	20.0 (3.1)	19.7 (3.1)	-3.5 (4.1)	21.3 (4.3)	10.9 (1.0)	-2.5 (1.9)	11.3 (0.7)	28.4 (2.0)
Overall	16.2 (2.5)	-2.3 (1.9)	20.9 (3.0)	20.2 (3.2)	-0.8 (2.8)	26.3 (4.0)	10.4 (0.7)	-1.0 (1.7)	11.3 (0.5)	28.5 (2.4)

TABLE 1. Mean absolute, signed and variable errors for the three dimensions separately (azimuth, elevation and distance), and for all dimensions (3D-error). The mean errors and standard deviation were calculated separately for each participant and 12 sound positions, then combined to generate the values in the table. Mean errors were segregated in respect to listening postures (Static and Active posture), front and backspace, and left and right sound sources. The errors in azimuth and elevation are reported in degrees, and the distance errors are in centimeters, the 3D-error is a computation of all dimension errors (see paragraph ‘3D-error’). (standard errors were put in parenthesis).

The bird-eye view in Figure 7A shows that the vast majority of trials fell within the correct quadrant and were distributed radially along the stimulation axis. Responses to front sounds were more eccentric than actual sound position, whereas this pattern was not evident for back sounds. To confirm this observation statistically, we entered the absolute (Figure 7A1), variable (Figure 7A2) and signed error (not shown) in the azimuth dimension into separate analysis of variance (ANOVA) with side (left and right) and antero-posterior sector (front and back) as within-participants factors. The analysis on signed error revealed a two-way interaction ($F(1,19) = 54.60, p < .001, \eta_g^2 = .42$) that confirmed larger lateral biases in pointing responses in front compared to backspace, for right (front: 23.2 ± 4.7 degrees; back: -6.0 ± 3.1 degrees; $p < 0.001$) and left target sounds (front: -28.2 ± 6.3 degrees; back: 5.8 ± 5.0 degrees; $p < 0.001$). The difference between front than back space also approached significance when absolute errors were considered ($F(1,19) = 3.58, p = .074, \eta_g^2 = .07$). No other main effect or interaction reached or approached significance.

The lateral view in Figure 7B shows that minimal front-back confusions occurred (back to front = $2.6\% \pm 1.9$; front to back = 10.7 ± 3.5) and that vertical dispersion was rather limited (variable error front: 32.7 ± 9.8 degrees; back: 20.3 ± 5.2 degrees). To study performance in elevation as a function of stimulation quadrant we entered absolute, variable and signed error into separate ANOVAs similar to one described above. Although absolute (Figure 7B1) and variable (Figure 7B2) errors in elevation were numerically larger for targets in front compared to backspace only marginally significant differences emerged (absolute error, main effect for antero-posterior sector: $F(1, 19) = 3.35$, $p = .083$, $\eta_g^2 = .06$).

Finally, Figure 7C shows that participants were able to distinguish the three target distances, yet with an average absolute error of 10.5 ± 1.3 cm. ANOVAs with target distance (near, middle and far) as within-participants factor, showed that target distance did not affect the absolute error (Figure 7C1), but significantly impacted on signed ($F(1.15, 21.94) = 30.23$, $p < .001$, $\eta_g^2 = .27$) and variable errors in Figure 7C2 ($F(1.54, 29.22) = 5.72$, $p = .013$, $\eta_g^2 = .10$). Participants underestimated far targets (-8.4 ± 2.5 cm) compared to middle (3.3 ± 2.5 cm; $p < 0.0001$) and near targets (0.2 ± 1.7 cm; $p < 0.0001$). In addition, dispersion of responses in depth was smaller for far targets (7.7 ± 0.8 cm) compared to middle (10.2 ± 0.9 cm; $p < 0.0001$) and near targets (9.4 ± 1.0 cm; $p < 0.0001$).

Hand pointing to sounds in active listening

Having characterized sound localization responses in azimuth, elevation and depth in the static listening condition, we turn to investigate whether active listening (i.e., free head-movements during sound presentation) changes spatial hearing. Dispersion plots on the left side of Figure 8 show hand-pointing responses (averaged across trials delivered from each loudspeaker position, separately for each participant), color-coded as a function of listening condition (black: static; red: active). Line plots on the right side of Figure 8 depict changes in absolute localization errors between listening conditions, separately for each participant (the bold horizontal line indicates the mean).

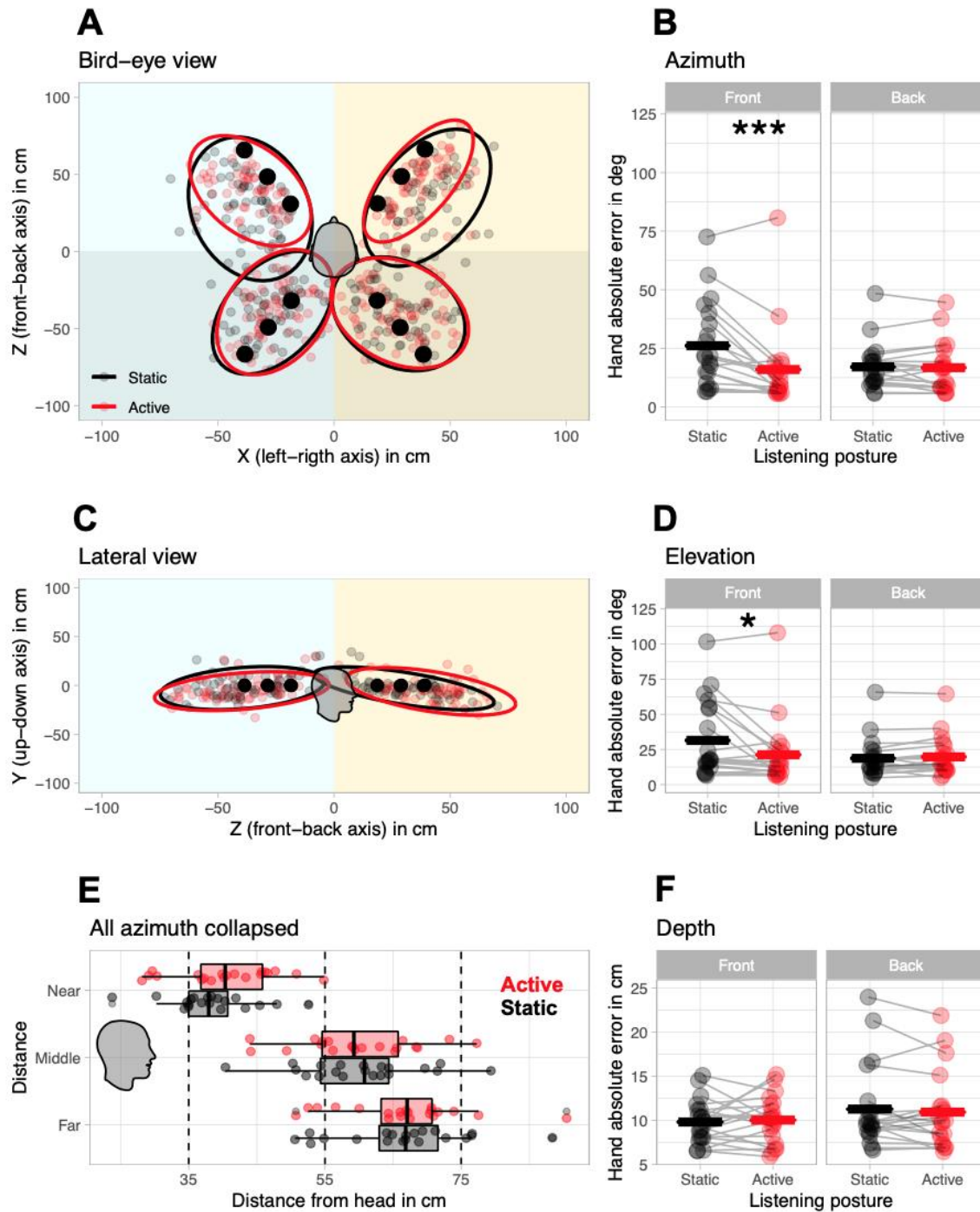


Figure 8. Effects of static and active listening on sounds localization. (A) Bird-eye view of all target positions (black dots) and hand pointing responses (smaller gray and red circles) for each participant, averaged across trials in a quadrant (i.e. front-left, front-right, back-left, back-right) irrespective of sound distance. Ellipses indicate 95% confidence interval of the responses across participants within each quadrant. Ellipses are color-coded as a function of listening condition (black: static listening; red: active listening). (B) Hand absolute error in azimuth for each participant, as a function of listening condition and antero-posterior position of target sounds. Bold horizontal lines indicate the mean for all participants. (C) Lateral view of all target positions and responses. Responses for each participant are averaged across side (left or right) and distance (near, middle or far). (D) Hand absolute error in elevation for each

participant. (E) Lateral view of responses in depth (black boxplot: static listening; red boxplot: active listening). (F) Hand absolute error in depth for each participant. Asterisks indicate significant differences (* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$).

The bird-eye view in Figure 8A shows that 95% confidence ellipses reduces in active compared to static listening, in front space selectively. 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-participants' factors. The analysis on absolute errors revealed the expected 2-way interaction ($F(1, 19) = 12.93$, $p = .002$, $\eta_g^2 = .03$), Azimuth absolute error in front space reduced for active (15.9 ± 3.8 degrees) compared to static listening (26.1 ± 4.0 degrees, $p = 0.0005$), whereas no such change occurred in back space (static: 17.2 ± 2.2 degrees; active: 16.5 ± 2.3 degrees; $p = 0.4$; Figure 8B). Instead, azimuth variable error reduced in active ($11.7.9 \pm 1.4$) compared to static listening (14.3 ± 1.3), irrespective of whether stimuli were in front or back space (main effect of antero-posterior sector, $F(1, 19) = 4.55$, $p = .046$, $\eta_g^2 = .03$).

A convergent result emerged for elevation (Figure 8C). When absolute and variable elevation errors were entered into an ANOVA similar to the one described above, the 2-way interaction between antero-posterior sector and listening condition reached significance for absolute errors ($F(1, 19) = 6.54$, $p = .019$, $\eta_g^2 = .02$). Absolute errors in front space reduced for active (21.2 ± 5.1 degrees) compared to static listening (31.4 ± 6.0 degrees, $p = 0.02$), whereas no such change occurred in backspace (static: 19.0 ± 3.1 degrees; active: 19.3 ± 3.0 degrees; $p = 0.7$; Figure 8D). Again, variable error 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 antero-posterior sector, $F(1, 19) = 4.59$, $p = .045$, $\eta_g^2 = .03$).

By contrast, active listening did not affect depth estimation (Figure 8E). 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-participants factors. No significant main effect or interaction involving listening posture emerged (all $F < 2.06$). Likewise, no main effect or interaction involving listening condition emerged for variable errors in depth.

Head movements during active listening

Kinematic tracking of the HMD allowed detailed investigation of head movements during sound in the active listening condition. Despite active listening allowed free head-movements during sounds, participants moved their head in 66.6% of trials on average (SD = 31.4%; median = 75.7, IQR = [58.3 90.9]). In fact, the distribution of percent head-movements (Supplementary Figure S2A) revealed 2 outliers: one participant that never moved his head and

another that moved only on 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; Figure S2B). Head-movements occurred both for targets in front and back space (75.7% and 72.9% of trials, respectively). For targets in front space they were mostly directed to the target. On average, for sounds at +30 degrees the first head movement was directed to 23.6 ± 1.8 degrees, whereas for sounds at -30 degrees it was directed to -27.0 ± 2.5 degrees (Figure S2C). For targets in back space head movements were distributed within the entire stimulated hemispace (Figure S2D). They were either directed to the front quadrant on the same side as the target (e.g., left front quadrant for targets at -150 degrees) or were aimed directly at the back target (in this case involving a trunk movement). On average, for sounds at +150 degrees the first head movement was directed to 107.5 ± 6.3 degrees, whereas for sounds located at -150 degrees it was directed to -118.0 ± 6.6 .

3D-error

As a final step, we quantified overall sound localization performance in 3D (i.e., across azimuth, elevation and depth, and irrespective of sound position) and studied how this cumulative index was affected by listening condition. To this aim, we adapted the error system introduced by Rakerd & Hartmann (1986; see also Grantham et al., 2008), which combines into a single measure the absolute constant error (referred to as \underline{C}) and the random error (referred to as \underline{S}) as follows: $\sqrt{\underline{C}^2 + \underline{S}^2}$. To obtain overall error in 3D we calculated in each trial i the norm of the vector $\underline{C}i$ (as shown in Figure 9A). This is the distance in 3D space between the participant's response (i.e., the coordinates of 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 $\underline{C}i$ extracted for each participant were then averaged irrespective of sound position. The random error \underline{S} for each participant was computed as the standard deviation of the responses at each sound position, averaged across all sound. Figure 9B shows change in 3D-error in the two listening conditions. Considering all participants and trials, the improvement in sound localization in active (28.5 ± 2.4) compared to static listening (32.2 ± 2.3) emerged as marginally significant ($F(1, 19) = 4.34, p = .051, \eta_g^2 = .03$). However, when the difference was studied as a function of mean number of head movements a positive correlation emerged ($r = 0.36, p = 0.027$). The higher the proportion of trials with head movements during sound emission, the larger the performance improvement in active compared to static listening. A convergent correlation emerged also between mean number of head-movements and 3D-error ($r = 0.34, p = 0.038$).

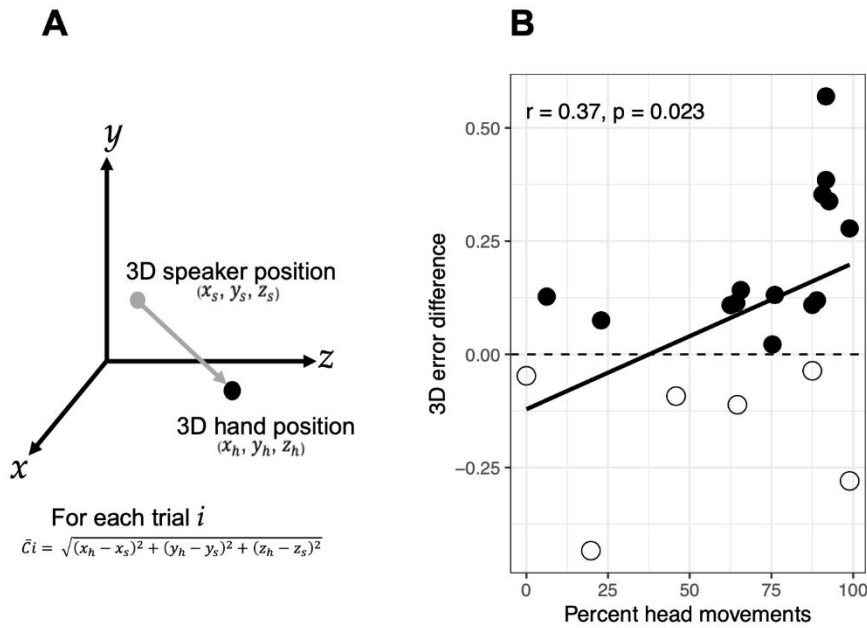


Figure 9. 3D-error. (A) Schematics for the calculation of the 3D vector. (B) Scatterplot of the difference in 3D-error between active and static listening (normalized difference based on static listening performance), as a function of percent head movements. Filled circles indicate participants who improved in active compared to static listening; empty circles indicate participants whose sound localization performance decreased in active compared to static listening.

DISCUSSION

In the present study, we aimed to report on a newly developed system for measuring sound localization, that allows replicable delivery of real sounds in 3D with respect to the head, while leaving participants free in their head-movement behavior. To validate our system, we examined to what extent voluntary head-movements improved sound localization in 3D – i.e., azimuth, elevation, and depth – by comparing static vs. active listening postures. We showed that our system is effective for controlling the delivery of sounds in 3D while measuring in real-time the participant’s behavior (hand, head and eye position), without imposing a physically constrained posture and with minimal training required. In addition, we found that head-behavior changes sound localization in 3D, improving specifically azimuth and elevation perception.

Sound localization in 3D

Despite minimal physical constraints, we succeeded in presenting free-field sounds at controlled 3D locations, across trials, participants, and sessions led by different experimenters.

When referenced to the center of the head, the average error across target positions for all participants and sessions of recording was below 1 cm. Notably, all experimenters achieved accurate and fast speaker 3D positioning with a few minutes training (< 5 minutes). Most importantly, participants required no procedural training to perform the task, with less than 7% of trials rejected for no-compliance to instructions.

Previous studies that examined sound localization in 3D were based on one of two approaches. On the one hand, some studies have used several real speakers in external space (free field sounds), which occupied fixed locations with respect to the participant (Ahrens et al., 2019), or a single speaker, which is moved and tracked in space (Brungart et al., 1999). On the other hand, other studies have used virtual auditory stimuli, either obtained by recording real sounds from the ear-canal of individual participants, or generated synthetically using Head-Related Transfer Functions HRTFs (individualized or not).

In the experimental approaches based on free field sounds, it is practically mandatory that the participant keeps a fixed posture, because this is the only way to establish off-line the position of target sounds with respect to the head (Wightman & Kistler 1992; Oldfield & Parker 1984; Brungart et al., 1999; Seeber et al., 2004). In addition, because physical sound sources are present, these approaches face the problem of controlling the contribution of visual cues to sound localization, as any visual prior will contribute to the interpretation of auditory cues (Jackson & Morton 1984; Makous & Middlebrooks 1990). Participants are therefore blind-folded, sometimes from the moment they enter the experimental room (e.g., Ahrens and al., 2019), or are instructed to close their eyes at specific moments during the task (e.g., Brungart et al., 1999), or face speakers hidden behind a fabric panel of some sort (e.g., Rabini et al., 2019).

Although studies using this approach typically used speakers at fixed positions in the environment, this needs not be the case, as first proposed by Brungart and colleagues (Brungart et al., 1999). In their pioneering study, participants were required to indicate the position of the sound source around their head while the experimenter manually placed the speaker during each trial. To ensure reproducibility of sound source coordinates across trials and participants, the participant's head was immobilized with a chin rest and the experimenter received verbal instructions about the pre-determined speaker location (i.e., a set of three numbers spoken through headphones, from 1 to 6, to indicate azimuth, elevation and distance according to learned correspondences). Although the speaker position was somewhat approximate, its actual location was recorded at the end of each trial using a position-sensing system mounted on the chin rest. Using this real sound method, Brungart and colleagues (1999) succeeded in placing sound sources in 3D space. While innovative, this experimental setup was complex and time consuming and, unfortunately, it was not followed-up by other investigators. Participants had to familiarize with the procedure before data collection. Moreover, the method had intrinsic limitations for the study of sound perception. 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, subjects had to close their eyes during sound positioning, thus limiting most oculomotor information that could have enhanced sound localization abilities (Maddox et al., 2014) and preventing any eye-movement monitoring.

With respect to the pioneering approach proposed by Brungart and colleagues (Brungart et al., 1999), our novel methodology allows: (1) accurate guiding of the speaker to pre-determined coordinates, that need not be interpreted by the experimenter; (2) anchoring of all coordinates to head position, without using a chin-rest; (3) control of visible cues about speaker position; and (4) control of eye position throughout the experiment. In addition, with respect to more classic approaches based on fixed speakers, our paradigm allowed sampling positions all around the participant in 3D. This is particularly evident when studying distance: if fixed speakers are used, placing them at different distances would interfere with the sound field generated by one another.

In the experimental approaches based on virtual auditory stimuli, two alternative strategies have been employed: in-ear sounds encoding and synthetic sound encoding. In in-ear sounds encoding, real sounds originating from pre-determined locations are recorded from the ear-canal of individual participants or from microphones in a dummy-head. The head has to remain motionless because it is impossible to modify the sound scene after having encoded each sound. In synthetic sound encoding, sound positions are computed in post-production via specialization techniques (HRTFs or Binaural Room Impulse Response) including several cues related to body features or room. Although a considerable amount of fundamental knowledge about HRTFs has accumulated over the years (for review, see Lida, 2019), synthetic sound encoding is not as powerful as the approach based on real speakers because several features must be implemented in the delivered sound.

First, the HRTFs depends on each individual's morphology (size of the head, shape of the pinna, etc.) which varies from one listener to another. The use of HRTFs without taking into account the individual characteristics of the subject's head (non-individualized HRTF), leads to the occurrence of front-back confusions, distortions of spatial and timbre perception (Seeber & Fastl, 2003; Wenzel et al., 1993; Middlebrooks, 1999), and synthetic sounds sources delivered by ear-phones are perceived inside the head, thus without distance perception (Begault et al., 2001). Second, modelling sound space with HRTFs is still a major challenge, particularly in relation to distance perception. Modelling sound distance requires a combination of a variety of acoustic cues (for review, see Zahorik et al., 2005): intensity of sound, direct-to-reverberant energy ratio, frequency spectrum, binaural cues. In addition, non-acoustic visual cues are often added to improve distance perception. To add to the complexity, different HRTFs models may be necessary when dealing with acoustic sources in far vs. near acoustic spaces (Brungart & Rabinowitz 1999). This is probably the reason why few of synthetic sound studies have dealt with sound localization in the 3D space reachable by the hand and close to the head. More generally, it may explain why distance is implemented separately with respect to the other dimensions: changes in azimuth and elevation are typically implemented at fixed distances (e.g., Rajendran & Gamper 2019; Lubeck et al., 2019), whereas changes in distance are typically implemented with fixed azimuth and elevation (e.g., Zahorik, 2002); (3) Third, use of synthetic sounds becomes extremely problematic when dealing with people with hearing deficit, individuals using hearing aids or cochlear implants, and children -- for whom HRTFs are likely to change faster over time. Yet, notable attempts to study these populations with HRTFs exist in the literature (e.g., Majdak et al., 2011, for cochlear implants users). In the

Majdak and colleagues study (2011), they delivered synthetic sounds directly to hearing aids of participants. To allow comparable sound perception across cochlear implant participant, the levels of synthetic sounds were adapted on an individual basis. However, despite this procedure, the authors faced the problem of variability of loudness perception over sessions. Finally, when compared to real-sound stimulation, the authors pointed that their “artificial testing conditions overestimated the real-life localization performance of our listener” (Majdak et al., 2011, page 207).

Several aspects of our approach to sound localization are notable with respect to the approach based on virtual auditory stimuli: (1) our procedure based on real sounds does not require measuring individual HRTF or recording from the individual’s ear canal; (2) the three dimensions can easily be changed at the same time in the whole space, particularly in reaching space; (3) our 3D sound localization approach can easily be adopted in individuals who use hearing aids or cochlear implants, children and elderly, for which synthetic acoustic approach is possible yet extremely time consuming and challenging.

Using the SPHERE system to characterize sound localization performance of normal hearing subject in static listening posture, we were able to show that participants discriminated azimuth, elevation and distance. Absolute average errors were 21.6 degrees, 25.1 degrees and 10.5 cm respectively. Differences emerged across along the medial dimension, with poorer performance in azimuth for front (26.1 degrees) compared to back targets (17.1 degrees), which emerged as a bias to point to more eccentric positions for frontal sources. For elevation, inaccuracy appeared for frontal sources compared to back ones (even if did not reach significant difference). This overall angular error measured in our study is numerically greater than those obtained by others studies in free field sound stimulation. For comparable elevation position, Brungart and colleagues (1999) obtained a mean angular error of 16.3 degrees, and Wightman & Kistler (1999), a mean error of 21.1 degrees. While the performance of our normal-hearing adults may appear poor, three aspects of our paradigm may contribute to this outcome. First, 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. Second, these errors combine uncertainty across the third dimension (i.e. distance is also unknown for participants). Third, they had no visual prior that could help sound localization.

Head-movements

Space is not represented directly on inner ear and spatial cues must be derived by the nervous system from incoming sound and dynamic auditory cues (i.e. head motion). Head motion is spontaneous in natural listening (Kim et al., 2013), and its key role in sound localization has been proposed since the last century (Wallach, 1940; see also Wightman & Kistler, 1999). It is generally assumed that head movements are accounted for during the computation of sound-source coordinates (Goossens & Van Opstal, 1999), as suggested by the fact that neck muscle stimulation has been shown to produce a shift of auditory localization toward the side of stimulation (Lewald et al., 1999). Converging evidence suggests that continuous integration of

head motion feedback on auditory input provides more stable sound source position (Vliegen et al., 2004). Several psychoacoustic studies demonstrate that head motion increases sound localization in humans (Pollack & Rose 1967; Perrett & Noble, 1997a; 1997b; Wightman & Kistler 1999; Vliegen et al., 2004; Brimijoin et al., 2013; Honda et al., 2013; McAnally & Martin, 2014), and monkeys (Populin, 2006). Speech perception improvements have also been documented (Munhall et al., 2004), even on patients with hearing aids (Mueller et al., 2014). Finally, head movements also proved important for the experience of the surrounding ambient sound (Suzuki et al., 2011) and are starting to be implemented in machine-hearing systems (Ma et al., 2015).

Notwithstanding these shared considerations, the study of head-movements in sound localization remained largely overlooked. Searching with the descriptor ‘sound localization’ in the MEDLINE repository since the work of Wallace (1940) returns 6014 entries, but when including the additional descriptor ‘head-movement’ reduces the entries to 271 articles related to spatial hearing performance (4.5%). One reason for this discrepancy may reside in the fact that taking into account head-movement has been problematic for most approaches to sound localization. In free-field approaches, head-movements have mostly been prevented, to ensure reproducibility of sound source position across trials and participants, or remained uncontrolled. In in-ear approaches, head tracking enables the signals fed to the two ears to change accordingly to head movements, so that the perception of the virtual sound source remains while moving the head. Indeed, some psychoacoustic studies have implemented in their model a dynamic modification of HRTFs cues caused by head motion (Brimijoin et al., 2013; Hendrickx et al., 2017; Pöntynen & Salminen 2019). Motion allows listeners to take advantage of dynamic localization cues, i.e. the changes in acoustic input caused by the movement of the head with respect to the sound source. This dynamic change caused by head movements can enhance sound perception: the accuracy of sound localization is better than that when HRTFs are not switched. Despite advances in 3D sound modelling taking particular account of head motion, mostly HRTF applications did not use 3D head tracking displacement (i.e. rotation in azimuth, tipping and tilting) and limit HRTF updating to lateral head movements only (Ma et al., 2015).

Our novel approach overcomes previous limitations by measuring head-position in real-time, thus allowing accurate and reproducible positioning of sounds even without physical constraints to head-position, and a full description of head behavior before, during and after sound presentation. Notably, in the active listening posture, no specific head-movement strategy was imposed on participants, and they were free to move or not their head to explore their surrounding environment during sound emission. As a proof-of-concept of our methodology, we chose to compare static vs. active listening posture for sounds lasting 3 seconds, a duration for which participants can benefit of acoustic dynamic cues from head-motion (Pollack & Rose, 1967; review by Middlebrooks & Green 1991).

Active listening improves 3D sound localization

We found that head motion during sound emission improved both their azimuth and elevation performances. Improvements concerned azimuth and elevation of front sources localizations (both absolute and variable errors) and azimuth and elevation of back sources localizations (variable error). For the third dimension (depth), active listening did not modify performances.

Previous studies have shown that head movements help normal hearing listeners to distinguish between sounds coming from front and rear positions (Makous & Middlebrooks 1990; Mueller et al., 2014; Bronkhorst 1995; Mackensen, 2004; Perrett & Noble 1997a; Wallach 1940; Wenzel et al., 1993; Wightman & Kistler, 1999). Front-back confusion is removed if head movements were larger than 5° (Perrett & Noble, 1997a), and if sounds were long enough (Muller et al., 2014; Perrett & Noble, 1997a). For a 3-sec. sound duration, in Perrett and Noble's study (1997a), a few confusions occurred but only in situations where the listeners did not rotate their heads. In our study with a sound duration of 3 sec., the subjects had enough time to improve their performance in localizing sound sources by head motion.

Current limitation of our approach

Sound localization studies with head motionless reported a more accurate performance for sources located in the frontal space than in back space (e.g. Oldfield & Parker, 1984; Brungart et al., 1999). In the present study, when head motion was not allowed, participants were more accurate in localizing sounds presented in the back (17.1 degrees) compared to front space (26.1 degrees). However, their performance increased when head motion was allowed (back: 16.7 degrees; front: 16.0 degrees). There are two likely reasons for this finding. First, several studies (Etchemendy et al., 2018; Brungart et al., 2000; Haber et al., 1993) showed superior accuracy and lower variability with direct-location response methods for sound localization (e.g., pointing with body parts) and also helps front-back discrimination (Aggius-Vella and al., 2017). In Oldfield and Parker study (1984), participant were head fixed and used a gun to localize the sound direction. The error in azimuth localization increased in the regions behind the head, particularly for azimuth positions 120° to 160°. When pointing at sounds behind the head subjects were instructed to turn the gun backwards and to 'shoot' from the front through their head to the sound source. In our direct-method, subjects are free to orient their body and their arm towards the rear field to respond, leading to decrease errors in the back space. Here we add to this direct-location panoply of methods an ecological, sensitive and rather intuitive (see Valzolgher et al., under revision) direct-location response to study 3D human spatial audition in the reaching space. Second, it has been reported that wearing and HMD may alter sound localization cues (Gupta et al., 2018; Ahrens et al., 2019; Genovese et al., 2018). Here we add that it may also affect sound source localization accuracy. In the Ahrens and colleagues study (2019), which explored the effect of HMD on sound localization presented in a frontal plane, a larger azimuthal error was found for lateral sound sources when participants wore HMD than without. In this study, sound duration was very short (240 ms) to limit the effect of head movement during sound presentation. These findings are in keeping with those observed here in static listening conditions (i.e., head motionless) for the frontal sources and could be due to a binaural disparity caused by the HMD. This possibility is supported by the participants

azimuthal and elevation improvement for frontal targets during the active listening condition. Accordingly, the lateral error increase due to HMD does not exist for the rear sources in either static or active listening, as the HMD is no longer offer an obstacle for binaural processing of rear sounds. In regard to depth, participants succeed in distinguish the three distances sound sources with an underestimation for far targets compared to near ones during static listening. The commonness of errors in distance localization is well-established (Brungart, 1999; Kearney et al., 2010; Zahorik 2002; 2005; Zahorik & Wightman 2001; Middlebrooks & Green, 1991; Parseihian et al., 2014). Sound distance cues such as binaural cue and interaural level difference (Brungart, 1999) are important for distance localization in the near-head acoustic field and not affected by the HMD.

CONCLUSION

SPHERE (European patent n° WO2017203028A1) is a valid tool to accurately sample the spatial abilities in auditory perception all around the subject, with minimal constraints to participant and experimenter. Most interestingly, SPHERE proved sensitive to allow detecting and quantifying the contribution of active listening (here, free head motion during sound emission) to improve sound localization accuracy and precision. This system also 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. Indeed, sound localization is not only a purely acoustic phenomenon, but a combination of multisensory information processing, especially relying on visual environmental information. HMDs offer the possibility of a 3D, fully immersive visual experience that combined to 3D sound simulation and head-body motion tracking will offer a highly versatile opportunity to assess normal and pathological sound localization performance. Beyond the range of performance of normal hearing people, it is a tool of choice for patients with cochlear implant and / or hearing aids (Avan et al. 2015).

Acknowledgements

This work was supported by LabEx CORTEX (<https://labex-cortex.universite-lyon.fr>), the Medisite Foundation (<https://www.fondationdefrance.org/fr/fondation/fondation-medisite>), ITMO New technologies for Neuroscience, IDEX Lyon 'Senses In Space Lab', ANR (ANR-16-CE17-0016-01), CNRS-PICS (Programme PHC Galilée Univ Franco-Italienne), and from IHU CeSaMe ANR-10-IBHU-0003.

REFERENCES

- 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., Kasper Duemose, L., Marschall, M., & Dau, T. (2019). Sound Source Localization with Varying Amount of Visual Information in Virtual Reality. *PLoS ONE* 29 (14): 1. <https://doi.org/10.1371/journal.pone.0214603>
- Andéol, G., Savel, S., & Guillaume, A. (2014). Perceptual factors contribute more than acoustical factors to sound localization abilities with virtual sources. *Front Neurosci.* 2014; 8: 451. doi: 10.3389/fnins.2014.00451
- 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*, 10:36. <https://doi.org/10.3389/fnins.2016>.
- Avan, P., Giraudet, F., & Büki, B. (2015). Importance of Binaural Hearing ». *Audiology and Neurotology* 20 (1): 3-6. <https://doi.org/10.1159/000380741>
- 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 Trans Biomed Eng.* 30(3):191-4. doi: 10.1109/tbme.1983.325108
- Begault, D.R., Wenzel, E.M., & Anderson, M.R. (2001). Direct Comparison of the Impact of Head Tracking, Reverberation, and Individualized Head-Related Transfer Functions on the Spatial Perception of a Virtual Speech Source. *J. Audio Eng. Soc.* 49: 14.
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B (Methodological)*. 57(1): 289–300.
- 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): 3678-88. <https://doi.org/10.1121/1.3409488>.
- 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): 162-68. <https://doi.org/10.1016/j.heares.2011.10.009>.
- 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. (1995). Localization of Real and Virtual Sound Sources ». *The Journal of the Acoustical Society of America* 98 (5): 2542-53. <https://doi.org/10.1121/1.413219>.

Brungart, D.S. (1999). Auditory localization of nearby sources. III. Stimulus effects. *J Acoust Soc Am.* 1999 Dec;106(6):3589-602. <https://doi.org/10.1121/1.428212>

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): 1956-68. <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): 1465-79. <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>.

Bulkin, D.A, & Groh, J.M. (2006). Seeing Sounds: Visual and Auditory Interactions in the Brain. *Current Opinion in Neurobiology* 16 (4): 415-19. <https://doi.org/10.1016/j.conb.2006.06.008>.

Etchemendy, P.E., Spiousas, I., Calcagno, E.R., Abregú, E., Eguia, M.C., & Vergara, R.O. (2018). Direct-Location versus Verbal Report Methods for Measuring Auditory Distance Perception in the Far Field. *Behavior Research Methods* 50 (3): 1234-47. <https://doi.org/10.3758/s13428-017-0939-x>.

Genovese, A., Zalles, G., Reardon, G., & Roginska, A. 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. Audio Engineering Society.

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): 2720-36. <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. (2007). Horizontal-Plane Localization of Noise and Speech Signals by Postlingually Deafened Adults Fitted With Bilateral Cochlear Implants. *Ear and Hearing* 28 (4): 524-41. <https://doi.org/10.1097/AUD.0b013e31806dc21a>.

Grantham, D.W., Ricketts, T.A., Ashmead, D.H., Labadie, R.F., & Haynes, D.D. (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>.

Gupta, A., Ranjan, R., He, J., & Gan, W-S. (2018). Investigation of Effect of VR/AR Headgear on Head Related Transfer Functions for Natural Listening. AES International Conference on Audio for Virtual and Augmented Reality (August 2018). <http://www.aes.org/e-lib/browse.cfm?elib=19697>

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>.

Hendrickx, E., Stitt, P., Messonnier, J-C., Lyzwa, J-M., Katz, B.F.G , & 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): 2011-23. <https://doi.org/10.1121/1.4978612>.

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): 1617-30. <https://doi.org/10.1121/1.1539520>.

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): 253-64. <https://doi.org/10.1068/i0522>.

Jackson, A., Morton, J. (1984). Facilitation of Auditory Word Recognition. *Memory & Cognition* 12 (6): 568-74. <https://doi.org/10.3758/BF03213345>.

Kearney, G., Gorzel, M., Boland, F., & Rice, H., (2010). Depth perception in interactive virtual acoustic environments using higher order ambisonic soundfields. Proc. of the 2nd International Symposium on Ambisonics and Spherical Acoustics.

Kim, C., Mason, R., & Brookes, T. (2013). Head Movements Made by Listeners in Experimental and Real-Life Listening Activities. *Journal of the Audio Engineering Society*, 2013, sect. 61(6). <http://www.aes.org/e-lib/browse.cfm?elib=16833>

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>

Lewald, J., & Ehrenstein, W.H. (1996). The Effect of Eye Position on Auditory Lateralization. *Experimental Brain Research*, no 108: 473-85. doi 10.1007/bf00227270

Lewald, J., Karnath, H-O., & Ehrenstein, W.H. (1999). Neck-Proprioceptive Influence on Auditory Lateralization. *Experimental Brain Research* 125 (4): 389-96. <https://doi.org/10.1007/s002210050695>.

Lida, K. (2019). Head-Related Transfer Functions And Acoustic Virtual Reality. Eds Springer Nature. ISBN 9789811397455

Litovsky, R., Parkinson, A., & Arcaroli, J. (2009). Spatial Hearing and Speech Intelligibility in Bilateral Cochlear Implant Users. *Ear and Hearing* 30 (4): 419-31. <https://doi.org/10.1097/AUD.0b013e3181a165be>.

Lubeck, T., Arend, J.M., & Porschmann, C. (2019). HMD-Based Virtual Environments for Localization Experiments. Conference DAGA 2019 Rostock, pp. 1116-1119.

Ma, N., May, T., Wierstorf, H., & Brown, G.J. (2015). A Machine-Hearing System Exploiting Head Movements for Binaural Sound Localisation in Reverberant Conditions. In *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2699-2703. South Brisbane, Queensland, Australia: IEEE. <https://doi.org/10.1109/ICASSP.2015.7178461>.

Mackensen, P. (2004). Auditive Localization. Head Movements, an Additional Cue in Localization. PhD Thesis, Berlin: Technische Universität Berlin.

Maddox, R.K., Pospisil, D.A., Stecker, G.C., & Lee, A.K.C. (2014). Directing Eye Gaze Enhances Auditory Spatial Cue Discrimination. *Current Biology* 24 (7): 748-52. <https://doi.org/10.1016/j.cub.2014.02.021>.

Majdak, P., Goupell, M.J. & Laback, B. (2011). Two-Dimensional Localization of Virtual Sound Sources in Cochlear-Implant Listeners. *Ear and Hearing* 32 (2): 198-208. <https://doi.org/10.1097/AUD.0b013e3181f4dfe9>.

Makous, J.C., & Middlebrooks, J.C. (1990). Two-dimensional Sound Localization by Human Listeners. *The Journal of the Acoustical Society of America* 87 (5): 2188-2200. <https://doi.org/10.1121/1.399186>.

McAnally, K.I., & Martin, R.L. (2014). Sound Localization with Head Movement: Implications for 3-d Audio Displays. *Frontiers in Neuroscience* 8 :210. <https://doi.org/10.3389/fnins.2014.00210>.

Middlebrooks, J.C, & Green, D.M. (1991). Sound Localization by Human Listeners. *Annu Rev Psychol* 42: 135-59. doi: 10.1146/annurev.ps.42.020191.001031

Middlebrooks, J.C. (1999). Individual Differences in External-Ear Transfer Functions Reduced by Scaling in Frequency. *The Journal of the Acoustical Society of America* 106 (3): 1480-92. <https://doi.org/10.1121/1.427176>.

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 International* 15 (1): 36-42. <https://doi.org/10.1179/1754762813Y.0000000040>.

Munhall, K.G., Jeffery J.A., Callan, D.E., Kuratate, T., & Vatikiotis-Bateson, E. (2004). Visual Prosody and Speech Intelligibility: Head Movement Improves Auditory Speech Perception. *Psychological Science* 15 (2): 133-37. <https://doi.org/10.1111/j.0963-7214.2004.01502010.x>.

Nava, E., Bottari, D., Bonfioli, F., Beltrame, M.A., & Pavani, F. (2009). Spatial Hearing with a Single Cochlear Implant in Late-Implanted Adults. *Hearing Research* 255 (1-2): 91-98. <https://doi.org/10.1016/j.heares.2009.06.007>.

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: 269. <https://doi.org/10.3389/fnins.2014.00269>.

Pavani, F., Meneghello F., Làdavas E. (2001). Deficit of auditory space perception in patients with visuospatial neglect. *Neuropsychologia*. 39(13):1401-9. Doi. 10.1016/s0028-3932(01)00060-4

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. *Experimental Brain Research* 189 (4): 435-49. <https://doi.org/10.1007/s00221-008-1440-7>.

Perrett, S., & Noble, W. (1997a). The Contribution of Head Motion Cues to Localization of Low-Pass Noise. *Perception & Psychophysics* 59 (7): 1018-26. <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): 2325-32. <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. *Perception & Psychophysics* 2 (12): 591-96. <https://doi.org/10.3758/BF03210274>.

Pöntynen, H., & Salminen, N.H. (2019). Resolving Front-Back Ambiguity with Head Rotation: The Role of Level Dynamics. *Hearing Research* 377 : 196-207. <https://doi.org/10.1016/j.heares.2019.03.020>.

Populin, L.C. (2006). Monkey Sound Localization: Head-Restrained versus Head-Unrestricted Orienting. *Journal of Neuroscience* 26 (38): 9820-32. <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.* 2019; 9: 7892. doi: 10.1038/s41598-019-44267-3

Rajendran, V.G., & Gamper, H. (2019). Spectral Manipulation Improves Elevation Perception with Non-Individualized Head-Related Transfer Functions. *The Journal of the Acoustical Society of America* 145 (3): EL222-28. <https://doi.org/10.1121/1.5093641>.

Rakerd, B., & Hartmann, W.M. (1986). Localization of Sound in Rooms, III: Onset and Duration Effects. *The Journal of the Acoustical Society of America* 80 (6): 1695-1706. <https://doi.org/10.1121/1.394282>.

Seeber, B.U., & Fastl, H. (2003). Subjective Selection of Non-Individual Head-Related Transfer Functions. Proceedings of the 2003 International Conference on Auditory Display, Boston, MA, USA, July 6-9. <http://hdl.handle.net/1853/50488>

Seeber, B.U., Baumann, U., & Fastl, H. (2004). Localization Ability with Bimodal Hearing Aids and Bilateral Cochlear Implants. *The Journal of the Acoustical Society of America* 116 (3): 1698-1709. <https://doi.org/10.1121/1.1776192>.

Slattery, W.H., & Middlebrooks, J.C. (1994). Monaural Sound Localization: Acute versus Chronic Unilateral Impairment I ». *Hearing Research* 75(1-2): 38-46. doi. 10.1016/0378-5955(94)90053-1

Suzuki, Y., Brungart, D., Iwaya, Y., Iida, K., Cabrera, D., & Kato, H. (2011). *Principles and Applications of Spatial Hearing*. Eds. World Scientific. <https://doi.org/10.1142/7674>.

Távora-Vieira, D., De Ceulaer, G., Govaerts, P.J., & Rajan, G.P. (2015). Cochlear Implantation Improves Localization Ability in Patients With Unilateral Deafness. *Ear and Hearing* 36 (3): e93-98. <https://doi.org/10.1097/AUD.000000000000130>.

Valzolgher, C., Verdelet, G., Salemme, R., Lombardi, L., Farnè, A., Pavani, P. Reaching to sounds in virtual reality: A multisensory-motor approach to re-learn sound localisation. Under review

Vliegen, J., van Grootel, T.J., & van Opstal, A. J. (2004). Dynamic Sound Localization during Rapid Eye-Head Gaze Shifts ». *Journal of Neuroscience* 24 (42): 9291-9302. <https://doi.org/10.1523/JNEUROSCI.2671-04.2004>.

Wallach, H. (1940). The Role of Head Movements and Vestibular and Visual Cues in Sound Localization. *Journal of Experimental Psychology* 27 (4): 339-68. <https://doi.org/10.1037/h0054629>.

Wenzel, E.M., Arruda, M., Kistler, D.J , & Wightman, F.L. (1993). Localization Using Nonindividualized Head-related Transfer Functions ». *The Journal of the Acoustical Society of America* 94 (1): 111-23. <https://doi.org/10.1121/1.407089>.

Wightman, F.L., & Kistler, D.J. (1992). The Dominant Role of Low-frequency Interaural Time Differences in Sound Localization ». *The Journal of the Acoustical Society of America* 91 (3): 1648-61. <https://doi.org/10.1121/1.402445>.

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): 2841-53. <https://doi.org/10.1121/1.426899>.

Zahorik, P., & Wightman, F.L. (2001). Loudness Constancy with Varying Sound Source Distance. *Nature Neuroscience* 4 (1): 78-83. <https://doi.org/10.1038/82931>.

Zahorik, P. 2002. Assessing Auditory Distance Perception Using Virtual Acoustics. *The Journal of the Acoustical Society of America* 111 (4): 1832-46. <https://doi.org/10.1121/1.1458027>.

Zahorik, P., Brungart, D.S., Bronkhorst, A.W. (2005). Auditory Distance Perception in Humans: A Summary of Past and Present Research. *Acat acustica United with Acustica*. vol. 91(3): 409-420