



Sound localization mechanisms and their plasticity in the human auditory cortex

Régis Trapeau

► To cite this version:

Régis Trapeau. Sound localization mechanisms and their plasticity in the human auditory cortex. Psychology. Université de Montréal; University of Montreal, 2016. English. NNT : . tel-01740564v1

HAL Id: tel-01740564

<https://hal.science/tel-01740564v1>

Submitted on 22 Mar 2018 (v1), last revised 23 Apr 2021 (v2)

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.



Distributed under a Creative Commons Attribution - NonCommercial - ShareAlike 4.0 International License

Université de Montréal

**ÉTUDE DES MÉCANISMES DE LOCALISATION AUDITIVE
ET DE LEUR PLASTICITÉ DANS LE CORTEX AUDITIF HUMAIN**

par
Régis Trapeau

Département de psychologie
Faculté des arts et des sciences

Thèse présentée à la Faculté des études supérieures
en vue de l'obtention du grade de Philosophiæ Doctor (Ph.D.)
en psychologie

décembre, 2015

© Régis Trapeau, 2015.

Université de Montréal
Faculté des études supérieures

Cette thèse intitulée:

**ÉTUDE DES MÉCANISMES DE LOCALISATION AUDITIVE
ET DE LEUR PLASTICITÉ DANS LE CORTEX AUDITIF HUMAIN**

présentée par:

Régis Trapeau

a été évaluée par un jury composé des personnes suivantes:

Pierre Jolicœur,	président-rapporteur
Marc Schönwiesner,	directeur de recherche
Karim Jerbi,	membre du jury
Étienne De Villers-Sidani,	examineur externe
Milton Campos,	représentant du doyen de la FES

Thèse acceptée le 22 mars 2016

RÉSUMÉ

Pouvoir déterminer la provenance des sons est fondamental pour bien interagir avec notre environnement. La localisation auditive est une faculté importante et complexe du système auditif humain. Le cerveau doit décoder le signal acoustique pour en extraire les indices qui lui permettent de localiser une source sonore. Ces indices de localisation auditive dépendent en partie de propriétés morphologiques et environnementales qui ne peuvent être anticipées par l'encodage génétique. Le traitement de ces indices doit donc être ajusté par l'expérience durant la période de développement. À l'âge adulte, la plasticité en localisation auditive existe encore. Cette plasticité a été étudiée au niveau comportemental, mais on ne connaît que très peu ses corrélats et mécanismes neuronaux. La présente recherche avait pour objectif d'examiner cette plasticité, ainsi que les mécanismes d'encodage des indices de localisation auditive, tant sur le plan comportemental, qu'à travers les corrélats neuronaux du comportement observé. Dans les deux premières études, nous avons imposé un décalage perceptif de l'espace auditif horizontal à l'aide de bouchons d'oreille numériques. Nous avons montré que de jeunes adultes peuvent rapidement s'adapter à un décalage perceptif important. Au moyen de l'IRM fonctionnelle haute résolution, nous avons observé des changements de l'activité corticale auditive accompagnant cette adaptation, en termes de latéralisation hémisphérique. Nous avons également pu confirmer l'hypothèse de codage par hémichamp comme représentation de l'espace auditif horizontal. Dans une troisième étude, nous avons modifié l'indice auditif le plus important pour la perception de l'espace vertical à l'aide de moulages en silicone. Nous avons montré que l'adaptation à cette modification n'était suivie d'aucun effet consécutif au retrait des moulages, même lors de la toute première présentation d'un stimulus sonore. Ce résultat concorde avec l'hypothèse d'un mécanisme dit de *many-to-one mapping*, à travers lequel plusieurs profils spectraux peuvent être associés à une même position spatiale. Dans une quatrième étude, au moyen de l'IRM fonctionnelle et en tirant profit de l'adaptation aux moulages de silicone, nous avons révélé l'encodage de l'élévation sonore dans le cortex auditif humain.

Mots clés: cortex auditif, localisation auditive, plasticité, IRMf.

ABSTRACT

Spatial hearing is an important but complex capacity of the auditory system. The human auditory system infers the location of a sound source from a variety of acoustic cues, known as auditory localization cues. Because these cues depend to some degree on morphological and environmental factors that cannot be predicted by the genetic makeup, their processing has to be fine-tuned during development. Even in adulthood, some plasticity in the processing of localization cues remains. This plasticity has been studied behaviorally, but very little is known about its neural correlates and mechanisms. The present research aimed to investigate this plasticity, as well as the encoding mechanisms of the auditory localization cues, using both behavioral and neuroimaging techniques. In the first two studies, we shifted the perception of horizontal auditory space using digital earplugs. We showed that young adults rapidly adapt to a large perceived shift and that adaptation is accompanied by changes in hemispheric lateralization of auditory cortex activity, as observed with high-resolution functional MRI. We also confirmed the hypothesis of a hemifield code for horizontal sound source location representation in the human auditory cortex. In a third study, we modified the major cue for vertical space perception using silicone earmolds and showed that the adaptation to this modification was not followed by any aftereffect upon earmolds removal, even at the very first sound presentation. This result is consistent with the hypothesis of a “many-to-one mapping” mechanism in which several spectral profiles can become associated with a given spatial direction. In a fourth study, using functional MRI and taking advantage of the adaptation to silicone earmolds, we revealed the encoding of sound source elevation in the human auditory cortex.

Keywords: auditory cortex, spatial hearing, plasticity, fMRI.

TABLE DES MATIÈRES

Résumé	iii
Abstract	iv
Liste des figures	viii
Liste des sigles	ix
CHAPITRE 1 : INTRODUCTION	1
1.1 Les mécanismes de localisation auditive	3
1.1.1 Mécanismes perceptifs	3
1.1.2 Physiologie de l’audition spatiale	7
1.1.3 Encodage des indices acoustiques	10
1.2 Plasticité en localisation auditive	15
1.2.1 Plasticité développementale	16
1.2.2 Existence d’une période critique ?	19
1.2.3 Importance du contexte comportemental	21
1.2.4 Implication du cortex auditif	22
1.2.5 Études chez l’humain	23
1.3 Objectifs et méthodes	30
CHAPITRE 2 : ARTICLE 1 : RELEARNING SOUND LOCALIZATION WITH DIGITAL EAR-PLUGS	35
2.1 Introduction	35
2.2 Methods	36
2.2.1 Participants	36
2.2.2 Plugs	36
2.2.3 Design and stimuli	36
2.2.4 Procedure	37
2.2.5 Analysis	37
2.3 Results	38
2.4 Discussion and conclusions	39
CHAPITRE 3 : ARTICLE 2 : ADAPTATION TO SHIFTED INTERAURAL TIME DIFFERENCES CHANGES ENCODING OF SOUND LOCATION IN HUMAN AUDITORY CORTEX	41
3.1 Introduction	42
3.2 Materials and Methods	44
3.2.1 Participants	44
3.2.2 Digital earplugs	44
3.2.3 Overall procedure	45
3.2.4 Stimuli	47
3.2.5 Behavioral testing	48
3.2.6 FMRI data acquisition	49
3.2.7 FMRI data analysis	50

3.2.8	Modeling	52
3.3	Results	54
3.3.1	Behavioral results	54
3.3.2	Neuroimaging Results	58
3.3.3	Modeling	63
3.4	Discussion	64
3.4.1	Behavioral recalibration and implications of missing aftereffect	64
3.4.2	Correlations between horizontal and vertical localization performance	66
3.4.3	Neural correlates of recalibration	67
3.4.4	Modeling	69
3.5	Conclusions	70
3.6	Acknowledgments	70

**CHAPITRE 4 : ARTICLE 3 : FAST AND PERSISTENT ADAPTATION TO NEW SPECTRAL CUES
 FOR SOUND LOCALIZATION SUGGESTS A MANY-TO-ONE MAPPING ME-
 CHANISM**

		71
4.1	Introduction	72
4.2	Methods	74
4.2.1	Participants	74
4.2.2	Earmolds	74
4.2.3	Overall procedure	74
4.2.4	Stimuli	76
4.2.5	Localization tasks	77
4.2.6	Control experiments	78
4.2.7	Training	79
4.2.8	Statistical analysis	80
4.2.9	Binaural recordings	81
4.2.10	Directional transfer functions	81
4.3	Results	83
4.3.1	Relation between VSI (Vertical Spectral Information) and ears free perfor- mance	83
4.3.2	Acoustical effect of the molds	84
4.3.3	Behavioral effects of the molds	86
4.3.4	Relation between behavioral effects of the molds and VSI dissimilarity	87
4.3.5	Adaptation	88
4.3.6	Aftereffect	90
4.3.7	Persistence tests	91
4.4	Discussion	91
4.4.1	Acoustic factors of behavioral performance	92
4.4.2	Effect of the molds on horizontal sound localization	94
4.4.3	Individual differences in adaptation to modified spectral cues	94
4.4.4	Aftereffect	95
4.4.5	Persistence	97

CHAPITRE 5 :	ARTICLE 4 : THE ENCODING OF SOUND SOURCE ELEVATION IN THE HU-	
	MAN AUDITORY CORTEX	99
5.1	Significance statement	99
5.2	Introduction	100
5.3	Methods	102
5.3.1	Participants	102
5.3.2	Procedure overview	102
5.3.3	Apparatus	103
5.3.4	Detailed procedure	104
5.3.5	Stimuli	105
5.3.6	Localization tasks	108
5.3.7	Training tasks	109
5.3.8	FMRI data acquisition	111
5.3.9	FMRI data analysis	112
5.4	Results	114
5.4.1	Behavioral results	114
5.4.2	Elevation tuning curves with free ears	116
5.4.3	Effect of the molds and adaptation on elevation tuning	121
5.4.4	Secondary narrow tuning to elevation	122
5.5	Discussion	125
5.5.1	Behavioral adaptation to modified spectral cues	125
5.5.2	Elevation tuning in auditory cortex	125
5.5.3	Limitations	128
5.5.4	Outlook	129
5.5.5	Acknowledgments	129
CHAPITRE 6:	DISCUSSION GÉNÉRALE	130
6.1	Encodage des indices binauraux et plasticité	130
6.2	Encodage des indices spectraux et plasticité	136
6.3	Rapidité d'adaptation, absence d'effet consécutif et persistance	138
6.4	Variabilité interindividuelle	141
6.5	Conclusion	144
BIBLIOGRAPHIE		169

LISTE DES FIGURES

1.1	Schématisation des indices binauraux	3
1.2	Cône de confusion et fonctions de transfert directionnelles	6
1.3	Voie auditive ascendante	9
1.4	Modèle de Jeffress	12
1.5	Le <i>reversing pseudophone</i> de Paul Thomas Young	15
1.6	Bouchons d'oreille numériques et moulage en silicone	31
2.1	Evolution of <i>P3</i> 's MSE	38
2.2	Raw results of <i>P3</i>	39
3.1	Earplugs delay and attenuation	45
3.2	Procedure timeline	46
3.3	Illustration of the model parameters	53
3.4	Time course of sound localization performance	55
3.5	Effect of the digital earplugs on horizontal sound localization performance	56
3.6	Effect of the digital earplugs on vertical sound localization performance	57
3.7	Correlation between azimuth and elevation performance	58
3.8	Directional tuning curves	59
3.9	Observed and modeled distributions of the center of gravity and maximum of directional tuning curves	60
3.10	Distribution of tuning curve center of gravity across auditory cortex	62
3.11	Sound vs. silence contrasts in both hemispheres and lateralization ratio	63
3.12	Modeled directional tuning curves	65
4.1	Illustration of the acoustic effects of an earmold on the DTFs of one participant	83
4.2	Mean VSI across octave bands	84
4.3	Mean acoustic effect of an earmold on the DTFs	85
4.4	Effects of the molds on the VSI	86
4.5	Correlations between behavioral results and acoustical metrics	88
4.6	Time course of sound localization performance	89
4.7	Trial-by-trial localization error in the aftereffect tests	91
5.1	Silicone mold and array of loudspeakers	103
5.2	Stimuli locations in the different tasks	106
5.3	Spectra and difference spectra of fMRI stimuli	107
5.4	Sound localization performance	115
5.5	Mean elevation tuning curve of all sound-responsive voxels in the right hemisphere	117
5.6	Acoustic elevation gain and mean elevation tuning curves in control experiment	119
5.7	Evolution of the mean elevation tuning curve and of the tuning curve EG with swapped stimuli	120
5.8	Distribution of tuning curve elevation gain across auditory cortex	121
5.9	Behavioral and tuning curve results across sessions	123
5.10	Mean tuning curves after subtracting the global mean	124
6.1	Courbes d'accord directionnelles moyennes créées à partir du modèle	133

LISTE DES SIGLES

ADC	Analog to Digital Converter / Convertisseur Analogique-Numérique
Az	Azimuth / Azimut
BOLD	Blood-Oxygen-Level Dependent / Dépendant du Niveau d'Oxygène Sanguin
CN	Cochlear Nucleus / Noyau Cochléaire
COG	Center of Gravity / Centre de Gravité
DAC	Digital to Analog Converter / Convertisseur Numérique-Analogique
dB HL	Decibel Hearing Level / Décibel Niveau d'Audition
dB SPL	Decibel Sound Pressure Level / Décibel Niveau de Pression Acoustique
DTF	Directional Transfer Function / Fonction de Transfert Directionnelle
DSP	Digital Signal Processor / Processeur de Signal Numérique
EEG	Electroencephalography / Électroencéphalographie
EG	Elevation Gain / Gain en Élévation
El	Elevation / Élévation
fMRI	Functional Magnetic Resonance Imaging / Imagerie par Résonance Magnétique Fonctionnelle
HRTF	Head-Related Transfer Function / Fonction de Transfert Relative à la Tête
IC	Inferior Colliculus / Colliculus Inférieur
ILD	Interaural Level difference / Différence Interaurale d'Intensité
ITD	Interaural Time difference / Différence Interaurale de Temps
KEMAR	Knowles Electronics Mannequin for Acoustic Research
LED	Light-Emitting Diode / Diode électroluminescente
LSO	Lateral Superior Olive / Olive Supérieure Latérale
MEG	Magnetoencephalography / Magnétoencéphalographie
MGB	Medial Geniculate Nucleus / Corps Genouillé Médian
MNTB	Medial Nucleus of the Trapezoid Body / Noyau Médian du Corps Trapézoïde
MSE	Mean Signed Error / Erreur Signée Moyenne
MSO	Medial Superior Olive / Olive Supérieure Médiane
OT	Optic Tectum / Tectum Optique
SOC	Superior Olivary Complex / Complexe Olivaire Supérieur
TR	Repetition Time / Temps de Répétition
VSI	Vertical Spectral Information / Information Spectrale Verticale

CHAPITRE 1

INTRODUCTION

L'ouïe joue un rôle primordial dans le déchiffrement du monde qui nous entoure. Quelques instants nous suffisent pour analyser les fines variations de pression atmosphérique qui constituent un événement sonore et en déduire la nature de sa source, sa potentielle dangerosité ou encore sa provenance. Pouvoir déterminer de manière précise la position d'une source sonore est fondamental pour bien percevoir et interagir avec notre environnement. La localisation auditive permet en effet de guider l'attention (Broadbent, 1954 ; Scharf, 1998), d'améliorer la détection, la ségrégation et la reconnaissance de phénomènes acoustiques divers (Bregman, 1994 ; Yost & Fay, 2007), et pour un grand nombre d'espèces, c'est un trait sélectif crucial pour la localisation de partenaires sexuels, de proies ou d'éventuels prédateurs (Masterton & Diamond, 1973). Cette faculté joue également un rôle de premier plan dans l'intelligibilité et la compréhension de la parole dans un environnement bruyant (Dirks & Wilson, 1969 ; Hirsh, 1950 ; Kock, 1950 ; MacKeith & Coles, 1971 ; Roman, Wang, & Brown, 2003), ou lorsqu'il s'agit de prêter attention à un seul interlocuteur quand plusieurs personnes parlent en même temps ("*cocktail party effect*", Bronkhorst, 2000 ; Cherry, 1953 ; Kidd, Arbogast, Mason, & Gallun, 2005).

Non seulement la localisation auditive est l'une des plus importantes, mais c'est aussi l'une des facultés les plus complexes du système auditif. Contrairement au système visuel ou somatosensoriel, les récepteurs sensoriels de l'oreille interne ne sont pas organisés topographiquement, mais tonotopiquement. Le système auditif doit alors décoder le signal acoustique pour en extraire les indices qui lui permettent de localiser une source sonore (cf. 1.1. Les mécanismes de localisation auditive). La relation entre ces indices et la position d'une source sonore dépend fortement de la morphologie de l'auditeur. Par conséquent, les indices de localisation varient d'un individu à l'autre et sont constamment modifiés pendant la période de développement. Les mécanismes de traitement de ces indices ne peuvent alors être fixés dès la naissance mais doivent être plastiques. Tout au long du développement, le système central doit tirer profit

de l'expérience sensorielle quotidienne et de l'interaction multimodale, pour améliorer, puis maintenir, ses aptitudes de localisation auditive. Il a été démontré que cette plasticité, nécessaire afin d'ajuster les mécanismes de localisation auditive au cours du développement, existe encore à l'âge adulte (cf. 1.2. **Plasticité en localisation auditive**). Les limites et le mode opératoire de cette plasticité, ainsi que le rôle joué par l'expérience sensorielle restent en grande partie indéterminés à ce jour.

Outre l'apport de nouveaux éléments pour la compréhension des mécanismes généraux de la plasticité du système nerveux central, l'étude de la plasticité du système auditif par l'altération des indices de localisation auditive est un puissant outil de sondage des mécanismes de traitement de l'information auditive. Elle peut en effet permettre de révéler des mécanismes neuronaux qui ne reflètent pas seulement les caractéristiques physiques des stimuli utilisés, mais la perception de ces stimuli. Ce champ d'études a également de l'importance d'un point de vue clinique. Ses résultats peuvent indiquer dans quelle mesure et par quels moyens, une ré-acquisition des aptitudes de localisation auditive est possible suite à un traumatisme auditif ou suite à l'équipement d'appareils auditifs ou d'implants cochléaires.

Ce chapitre débute par un état de l'art des connaissances psychoacoustiques et neuronales des mécanismes de localisation auditive, suivi d'une recension des études de la plasticité en localisation auditive. Il se terminera par l'énoncé des objectifs et des méthodes utilisées pour mener à bien les différentes études qui composent cette recherche.

1.1 LES MÉCANISMES DE LOCALISATION AUDITIVE

1.1.1 Mécanismes perceptifs

Le système auditif déduit la position d'une source sonore grâce à des indices acoustiques qui dépendent et sont créés par des propriétés anatomiques du corps humain. Ces indices sont appelés indices de localisation auditive et sont généralement regroupés en plusieurs classes, à savoir les indices binauraux, spectraux et dynamiques.

1.1.1.1 Les indices binauraux

Les indices binauraux sont assurés par la présence de deux oreilles, placées de chaque côté du crâne. La tête produit un effet de diffraction des ondes sonores et crée une ombre acoustique au niveau de l'oreille la plus éloignée de la source sonore. Par conséquent, lorsqu'une source sonore ne se trouve pas sur le plan sagittal médian, l'intensité du son qu'elle produit est plus importante dans une oreille que dans l'autre (voir figure 1.1). Les premières recherches en localisation auditive ont très vite conclu que cette différence d'intensité procurait un indice de localisation (Steinhauser, 1879 ; Thompson, 1877). Cet indice varie suivant la position d'une source sonore et on parle de différence interaurale d'intensité (ou ILD¹ pour *Interaural Level Difference*).

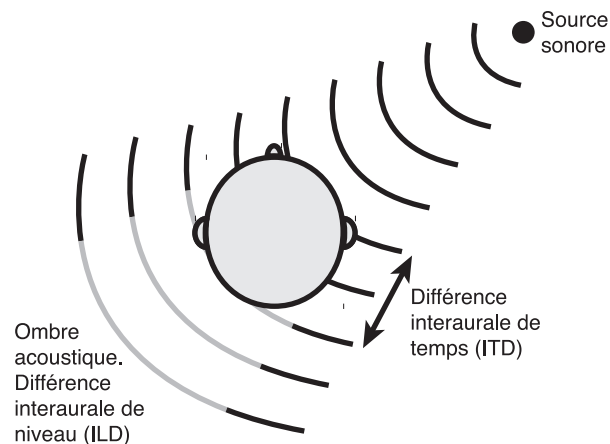


Figure 1.1 – Schématisation des indices binauraux.

¹Dans un souci de cohérence avec les articles de la thèse, ainsi qu'avec la littérature scientifique internationale, seuls les sigles anglais seront utilisés dans ce texte.

Quand la fréquence d'un son diminue, sa longueur d'onde augmente et peut atteindre plusieurs fois la largeur de la tête. L'ombre acoustique devient alors inexistante et les ILDs sont considérées comme négligeables pour des fréquences inférieures à environ 1 kHz (Akeroyd, 2006 ; Kuhn, 1987). En dessous de 1.5 kHz, les ILDs sont insuffisantes pour procurer un indice de localisation exploitable (Middlebrooks & Green, 1991), mais nous arrivons cependant très bien à localiser les sons de cette gamme de fréquences. Ceci est dû à l'existence d'un deuxième indice binaural, produit cette fois-ci par la différence de temps qui sépare l'arrivée d'une onde sonore d'une oreille à l'autre (voir figure 1.1). Cet indice, appelé différence interaurale de temps (ou ITD pour *Interaural Time Difference*), est efficace pour des sons basses fréquences et devient ambigu à partir de fréquences supérieures à environ 1.2 kHz (Middlebrooks & Green, 1991). C'est Lord Rayleigh (Strutt (Lord Rayleigh), 1907), qui le premier a démontré que les infimes différences temporelles entre les deux oreilles pouvaient permettre de localiser les sons à basses fréquences. Cette faculté du système auditif est remarquable quand on considère que la différence temporelle maximale que produit l'écartement des oreilles de l'être humain est environ 700 μs et que nous pouvons discriminer des ITDs d'environ 15 μs (Mills, 1958 ; Yost, 1974). La découverte de cette faculté prouva l'existence de deux mécanismes complémentaires qui permettent au système auditif humain de localiser respectivement les sons de hautes et basses fréquences. Cette dichotomie des mécanismes de localisation auditive proposée par Lord Rayleigh est connue sous le nom de *duplex theory of localization*. Des études ultérieures ont confirmé cette théorie et montré qu'aux fréquences intermédiaires des deux mécanismes (entre environ 1 et 2 kHz), les indices sont plus ambigus et les performances de localisation moindres (Mills, 1958 ; Stevens & Newman, 1936).

Les ITDs peuvent procurer un indice secondaire de localisation pour des sons hautes fréquences sur la base de leur enveloppe temporelle (Henning, 1974 ; McFadden & Pasanen, 1976). Cette particularité attribue aux ITDs le rôle de dominant parmi les indices binauraux (Macpherson & Middlebrooks, 2002). De plus, les ITDs restent constantes en fonction de l'éloignement de la source sonore de l'auditeur, ce qui n'est pas le cas pour les ILDs. En effet, dans un rayon inférieur à 1 m, plus la source sonore est proche de l'auditeur, plus les ILDs augmentent,

ce qui donne lieu à une perception biaisée de la position de la source (Brungart & Rabinowitz, 1999).

Les indices binauraux sont de robustes porteurs d'information pour la résolution de l'azimut, c'est-à-dire la direction angulaire sur le plan horizontal. Cependant, ils sont insuffisants pour pleinement déterminer la position d'une source sonore, en particulier sur le plan vertical (élévation) et pour différencier les sons provenant de devant ou de derrière. Différentes positions spatiales, situées sur la surface d'un cône imaginaire appelé "cône de confusion" (Wallach, 1939), peuvent en effet produire les mêmes valeurs d'ITD et d'ILD (voir figure 1.2.A). Afin de différencier les positions qui se trouvent sur ce cône de confusion, une autre classe d'indices est nécessaire, les indices spectraux.

1.1.1.2 Les indices spectraux

Le relief anatomique du corps humain produit des phénomènes de diffraction des ondes sonores et ainsi, le torse, la tête et en particulier le pavillon (oreille externe), agissent comme un filtre spectral des phénomènes acoustiques qui parviennent au canal auditif (Batteau, 1967; Gardner, 1973). Pour chaque direction de l'espace auditif, le filtre créé par ces diffractions a une fonction de transfert unique. L'ensemble des fonctions de transfert relatives à un individu est appelé HRTFs pour *Head-Related Transfer Functions* et lorsque les composantes non-directionnelles en sont soustraites, on parle de fonctions de transfert directionnelles, ou DTFs pour *Directional Transfer Functions*. La morphologie des oreilles, de la tête et du corps en général, est propre à chaque individu et par conséquent les DTFs aussi sont spécifiques à chacun (Middlebrooks, 1999). La figure 1.2.B–D montre les DTFs de trois individus, pour des positions spatiales placées à $\pm 45^\circ$ sur le plan vertical frontal.

Ces diffractions ont pour effet d'amplifier ou diminuer certaines régions du spectre acoustique, résultant en des fonctions de transfert uniques pour chaque position de l'espace. Les réflexions au niveau du torse modifient le spectre sonore autour de 2–3 kHz (Algazi, Avendano, & Duda, 2001), alors que le pavillon affecte les plus hautes fréquences, au-delà de 4 kHz (Asano, Suzuki, & Sone, 1990). Le pavillon est la partie du corps humain qui procure les indices spectraux

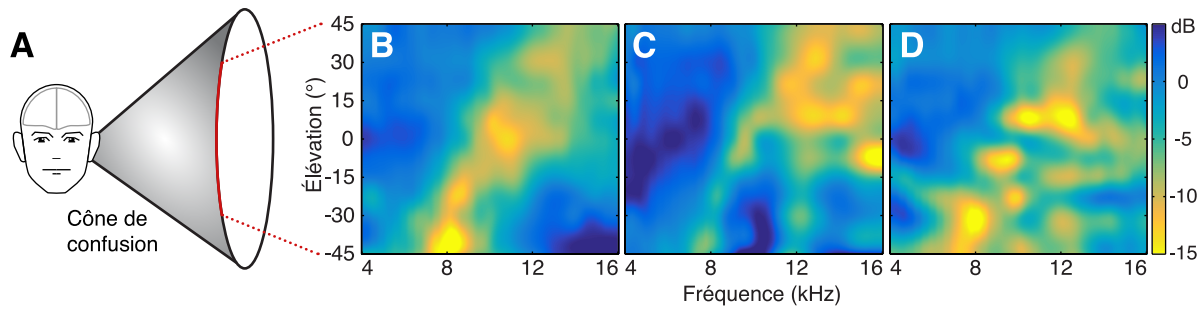


Figure 1.2 – Cône de confusion et fonctions de transfert directionnelles. A) Schématisation du principe de cône de confusion. B–D) Exemples de DTFs de trois individus différents, pour des positions spatiales placées à $\pm 45^\circ$ sur le plan vertical frontal.

les plus robustes. Une occlusion des différentes cavités du pavillon (sans bloquer le canal auditif) annihile la localisation sur le plan vertical et la différenciation entre les sons provenant de devant ou derrière (Gardner & Gardner, 1973 ; S. R. Oldfield & Parker, 1984a). Langendijk et Bronkhorst (2002) ont montré que les indices les plus primordiaux pour la localisation sur le plan vertical se trouvent autour des fréquences 6–12 kHz, alors que les indices importants pour la discrimination devant-derrrière couvrent tout le spectre au-dessus de 4 kHz.

Parce qu’une oreille procure des indices spectraux à elle seule, ces indices ont également été appelés indices monauraux. Il a cependant été démontré que la localisation monaurale sur le plan vertical est moins précise que lorsque les informations provenant des deux oreilles sont disponibles (S. R. Oldfield & Parker, 1986 ; Slattery & Middlebrooks, 1994). Hofman et Van Opstal (2003) ont proposé que la perception de l’élévation est en partie attribuable à des interactions binaurales et à la perception de l’azimut. Cette hypothèse a été confirmée par une étude d’adaptation à des indices spectraux modifiés unilatéralement (M. M. V. Van Wanrooij & Van Opstal, 2005).

1.1.1.3 Les indices dynamiques

En plus des indices binauraux et spectraux, il existe des indices dynamiques qui naissent des micro-mouvements de la tête de l’auditeur. En 1939, Hans Wallach avait déjà formulé l’hypothèse selon laquelle ces mouvements de tête aident à résoudre certaines ambiguïtés du cône de confusion. Cette hypothèse a été vérifiée et il a été montré que les indices dynamiques

contribuent à discriminer les sons provenant de devant ou de derrière (Perrett & Noble, 1997a ; Wightman & Kistler, 1999), ainsi que sur le plan vertical (Perrett & Noble, 1997b). En l'absence d'indices spectraux, les ITDs dynamiques procurent un indice saillant pour cette discrimination devant-derrière (Macpherson, 2013). Wightman et Kistler (1999) ont montré que ces indices dynamiques ne sont pertinents que s'ils sont initiés par l'auditeur lui-même.

1.1.2 Physiologie de l'audition spatiale

1.1.2.1 Structures sous-corticales

Un signal auditif quittant l'oreille interne et la cochlée par le nerf auditif, atteint le noyau cochléaire (CN, *Cochlear Nucleus*), où différents types de cellules vont projeter vers différents noyaux. Les cellules de type IV du noyau cochléaire dorsal sont sensibles à des variations spectrales à bande étroite (*notch-sensitive neurons* : Imig, Bibikov, Poirier, & Samson, 2000 ; E. D. Young, Spirou, Rice, Voigt, & Rees, 1992) et projettent vers les cellules de type O du noyau central du colliculus inférieur (IC, *Inferior Colliculus*). Il semblerait que ces groupes de cellules constituent une première étape du traitement des indices spectraux (Davis, Ramachandran, & May, 2003). Les cellules touffues du noyau cochléaire ventral quant à elles, déchargent en phase avec le signal auditif (*phase-locking*) et projettent vers le complexe supérieur olivaire (SOC, *Superior Olivary Complex*), structure spécialisée dans le traitement des indices binauraux.

Le SOC reçoit des afférences des deux CNs, les afférences controlatérales passant par le corps trapézoïde. La dichotomie définie au niveau psychophysique entre ITDs et ILDs se retrouve au niveau neuronal dans le SOC, où deux noyaux, l'olive supérieure médiale (MSO, *Medial Superior Olive*) et l'olive supérieure latérale (LSO, *Lateral Superior Olive*) sont sensibles à chacun de ces indices respectivement (Brand, Behrend, Marquardt, McAlpine, & Grothe, 2002 ; Irvine, Park, & McCormick, 2001 ; Masterton, Diamond, Harrison, & Beecher, 1967 ; Park, 1998). Le traitement des ILDs dans le LSO est possible grâce aux afférences excitatrices qu'il reçoit du CN ipsilatéral et inhibitrices depuis le noyau médian du corps trapézoïde controlatéral (MNTB, *Medial Nucleus of the Trapezoid Body*). Or, l'existence du MNTB chez l'être humain est controversée (Bazwinsky,

Hilbig, Bidmon, & Rübsamen, 2003 ; Hilbig, Beil, Hilbig, Call, & Bidmon, 2009 ; Moore, 2000), mettant en doute un traitement des ILDs au niveau du LSO, par ailleurs plus petit chez l'humain que chez d'autres mammifères (Hilbig et al., 2009). De récentes études soutiennent cependant l'existence de ces structures dans le tronc cérébral humain (Kulesza & Grothe, 2015).

Depuis le SOC, les voies parallèles de traitement des ITDs et ILDs convergent vers le colliculus inférieur via le lemniscus latéral. L'IC est considéré comme étant une structure centrale du traitement des indices de localisation, ainsi qu'un centre intégrateur de ces indices (Chase & Young, 2006). Les neurones de l'IC sont sensibles aux différents indices binauraux ainsi qu'aux indices monauraux (Casseday & Covey, 1987 ; Casseday, Fremouw, & Covey, 2002 ; Delgutte, Joris, Litovsky, & Yin, 1995). La dernière étape de la voie ascendante auditive est le corps genouillé médian du thalamus (MGB, *Medial Geniculate Nucleus*). Le MGB constitue le point d'entrée de l'information auditive vers le cortex. La voie auditive ascendante est représentée schématiquement dans la figure 1.3.

Le colliculus supérieur ne fait pas partie de la voie auditive ascendante, mais reçoit cependant des projections descendantes du cortex auditif et projette lui-même vers le colliculus inférieur. Une carte topographique de l'information auditive a été reportée à plusieurs reprises dans les couches profondes du colliculus supérieur (Gaese & Johnen, 2000 ; King & Hutchings, 1987 ; Middlebrooks & Knudsen, 1984 ; Palmer & King, 1982).

1.1.2.2 *Cortex auditif*

Bien qu'une partie importante du traitement des informations auditives pour la localisation semble prendre place dans des aires sous-corticales, il a été observé que des lésions au niveau du cortex auditif primaire dégradent les performances de localisation auditive (Jenkins & Merzenich, 1984 ; Jenkins & Masterson, 1982). Les plus importantes dégradations de performances sont observées lorsque les aires voisines du cortex auditif primaires sont également affectées (Harrington, Stecker, Macpherson, & Middlebrooks, 2008 ; Malhotra, Stecker, Middlebrooks, & Lomber, 2008 ; Miller & Recanzone, 2009 ; Nodal et al., 2010). Le traitement de l'espace auditif

n'est donc pas limité au cortex auditif primaire et ces études soulignent l'importance des aires voisines.

Les études de lésions unilatérales ont démontré une importante contralatéralité de la représentation de l'espace auditif au niveau du cortex. Une lésion au niveau du cortex auditif primaire entraîne en effet des déficits en localisation auditive dans l'hémichamp contralatéral à la lésion, chez des mammifères comme les primates, les chats ou les furets (Beitel & Kaas, 1993; Heffner & Heffner, 1990; Kavanagh & Kelly, 1987). Une inactivation réversible du cortex auditif, par voie pharmacologique ou cryogénique, entraîne une diminution des aptitudes de localisation sonore chez l'animal sain, en accord avec les études sur les lésions (Malhotra, Hall, & Lomber, 2004; Smith et al., 2004). Cependant la contralatéralité reportée par les études physiologiques est plus mitigée dans les études d'imagerie cérébrale chez l'humain, certaines études l'ayant rapportée (Krumbholz, Schönwiesner, Cramon, et al., 2005; Palomäki, Tiitinen, Mäkinen, May, & Alku, 2005; Pavani, Macaluso, Warren, Driver, & Griffiths, 2002; D. L. Woods et al., 2009), et d'autres non (Brunetti et al., 2005; Jäncke, Wüstenberg, Schulze, & Heinze, 2002; Woldorff et al., 1999; Zimmer, Lewald, Erb, & Karnath, 2006).

Une hypothèse de dominance de l'hémisphère droit pour le traitement de l'espace auditif a également été proposée chez l'humain (Zatorre & Penhune, 2001). Cette hypothèse est corroborée par des études montrant une forte contralatéralité de l'hémisphère gauche pour le traitement des ITDs, alors que l'hémisphère gauche répondrait aux deux hémichamps (Kaiser, Lutzenberger, Preissl, Ackermann, & Birbaumer, 2000; Salminen, Tiitinen, Miettinen, Alku, & May, 2010; Schönwiesner, Krumbholz, Rübsamen, Fink, & von Cramon, 2007).

1.1.3 Encodage des indices acoustiques

1.1.3.1 Indices binauraux

La façon dont les ILDs sont encodés au niveau du LSO est facile à concevoir. Le LSO reçoit des afférences excitatrices ipsilatérales et des afférences inhibitrices contralatérales. La somme des deux signaux à la sortie du LSO sera donc d'autant plus importante que la source sonore

provenait du côté de l'oreille ipsilatérale (Irvine et al., 2001). Le LSO encode donc les ILDs par un codage fréquentiel (fréquence de décharge des neurones) et montre une préférence pour l'espace auditif ipsilatéral. Ce traitement ipsilatéral des ILDs au niveau du tronc cérébral devient contralatéral dès le mésencéphale étant donné les projections excitatrices majoritairement contralatérales du LSO vers l'IC (Glendenning, Baker, Hutson, & Masterton, 1992).

L'encodage des ITDs est plus délicat. Le système auditif doit en effet mesurer et comparer des différences d'arrivée des signaux acoustiques entre les deux oreilles de l'ordre de dizaines de microsecondes. Lloyd Jeffress proposa en 1948 une théorie d'encodage de fines différences temporelles, connue sous le nom de *Jeffress model* (Jeffress, 1948). Ce modèle présuppose l'existence de neurones qui recevraient des afférences des deux oreilles et qui ne déchargeraient que lorsque ces afférences sont synchrones. Ces neurones devraient être alignés pour former un réseau et agiraient comme des détecteurs de coïncidence (*coincidence-detector neurones*). Le modèle présuppose également l'existence d'un agencement précis de fibres nerveuses de longueurs différentes qui formeraient deux groupes de "lignes à retard" opposés (*delay-lines*). Lorsqu'un son arrive aux deux oreilles, le modèle fonctionne ainsi : le signal auditif provenant d'une oreille parcourt une ligne à retard et atteint les différents détecteurs de coïncidence à des intervalles de temps différents ; le signal provenant de l'autre oreille parcourt la ligne à retard opposée et atteint successivement les détecteurs de coïncidence dans l'ordre inverse de celui dans lequel le signal de la première oreille les atteint. Seul le neurone où les deux signaux ont convergé de manière synchrone va décharger des potentiels d'action. Différentes ITDs sont alors à l'origine de la décharge de différents neurones, ayant chacun leur ITD préférée. La disposition de ces neurones, les détecteurs de coïncidence, formerait alors une carte topographique des ITDs et l'encodage des ITDs serait alors spatial (*place code*). Le modèle de Jeffress est représenté sur la figure 1.4.

Quarante ans après la proposition de Jeffress, des études chez la chouette effraie ont montré que les axones des cellules du noyau magnocellularis² procurent des lignes à retard et que les

²Homologue aviaire du noyau cochléaire ventral.

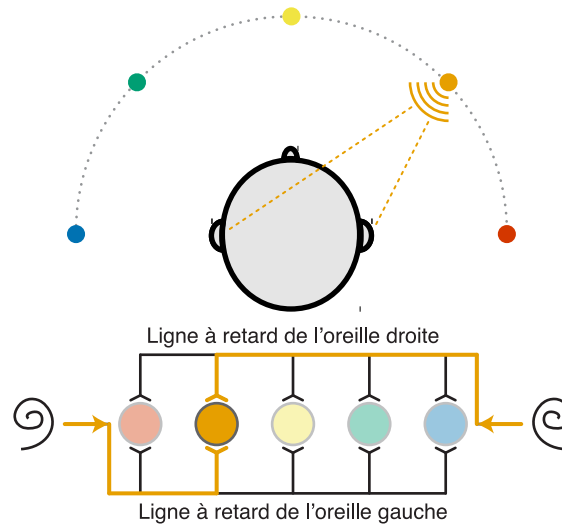


Figure 1.4 – Modèle de Jeffress. Dans cet exemple, les signaux auditifs correspondant à l'arrivée à chaque oreille d'un son provenant de 45° vers la droite, se propagent à travers des lignes à retard opposées. Les signaux se rejoignent de manière synchrone au niveau d'un seul détecteur de coïncidence. De par son emplacement, la décharge de ce neurone indique alors la provenance du son.

cellules du noyau laminaris³ agissent comme des détecteurs de coïncidence et sont étroitement sensibles aux ITDs (Carr & Konishi, 1988, 1990). Ces études ainsi que des résultats chez la poule (Seidl, Rubel, & Harris, 2010), indiquent fortement l'existence d'un modèle proche de celui de Jeffress chez les oiseaux.

Il a alors longtemps été considéré que le MSO traitait les ITDs selon la vision élégante de Jeffress. Les résultats d'études chez le mammifère présentent en revanche des incohérences avec ce modèle et remettent en question l'existence de ce modèle chez l'homme. Les neurones du MSO du cochon d'inde (McAlpine, Jiang, & Palmer, 2001), du chat (Hancock & Delgutte, 2004), ou encore de la gerbille (Brand et al., 2002 ; Siveke, Pecka, Seidl, Baudoux, & Grothe, 2006), sont bien sensibles aux ITDs, mais leur taux de décharge maximal correspond pour la plupart d'entre eux à des ITDs trop grands pour être produits par l'écartement des oreilles de ces animaux. Aucune carte topographique des sensibilités aux ITDs n'a été révélée, les neurones répondant majoritairement de manière équivalente. Ces observations ne satisfaisant pas l'hypothèse d'un codage spatial, il a été proposé que les ITDs seraient représentées par un codage fréquentiel chez ces espèces. Codage pour lequel la pente, plutôt que le pic, des courbes d'accords aux ITDs

³Homologue aviaire du MSO.

(*tuning curves*) serait porteuse d'information (Harper & McAlpine, 2004 ; Lesica, Lingner, & Grothe, 2010). David McAlpine (2005) propose que la représentation de l'espace auditif chez le mammifère soit alors effectuée par la comparaison de l'activité neuronale des deux hémisphères, chaque hémisphère comprenant une population de neurones à sélectivité large et préférant l'espace controlatéral.

Des observations similaires ont été effectuées au niveau du cortex auditif du chat (Middlebrooks, Xu, Eddins, & Green, 1998 ; Stecker, Harrington, & Middlebrooks, 2005) et du singe (Werner-Reiss & Groh, 2008 ; Miller & Recanzone, 2009), ce qui supposerait que cette représentation est également valable au niveau cortical. Ces études montrent également que la controlatéralité n'est pas évidente au niveau cortical. Ceci implique que la représentation de l'espace auditif n'est pas forcément effectuée par la comparaison de l'activité neuronale des deux hémisphères, mais plutôt par celle de deux populations de neurones, préférant chacune un hémichamp, et présentes dans chaque hémisphère.

Cette hypothèse d'une représentation de l'espace auditif horizontal par deux populations neuronales à préférences directionnelles opposées (*population rate code of two opponent populations*) est démontrée chez l'humain par des études en électroencéphalographie (EEG : Magezi & Krumbholz, 2010) et en magnétoencéphalographie (MEG : Salminen, Tiitinen, Yrttiaho, & May, 2010 ; Salminen, May, Alku, & Tiitinen, 2009). Le terme "codage par hémichamp" (*hemifield code*) est proposé pour caractériser cette représentation (Salminen, Tiitinen, & May, 2012). Les techniques d'EEG et de MEG ont une résolution spatiale limitée et ne permettent pas d'établir des courbes d'accord de la sélectivité aux directions spatiales. Ces études n'ont alors pas directement observé ce codage, mais l'on déduit à partir de paradigmes d'adaptation à un stimulus spécifique.

1.1.3.2 Indices spectraux

Outre l'implication des cellules de type IV du noyau cochléaire dorsal et de leurs projections vers les cellules de type O de l'IC (Davis et al., 2003), la façon dont sont encodés les indices spectraux reste inconnue à ce jour. Des études électrophysiologiques chez l'animal ont montré

que des cellules du cortex auditif primaire (Bizley, Nodal, Parsons, & King, 2007 ; Brugge et al., 1994 ; Mrsic-Flogel, King, & Schnupp, 2005), ainsi que des cellules d'aires corticales plus élevées (Xu, Furukawa, & Middlebrooks, 1998), étaient sensibles à différentes élévations sonores. Ces études suggèrent également que les neurones du cortex auditif sont sensibles à plusieurs indices de localisation à la fois. La façon dont cette sensibilité se manifeste n'a cependant pas été révélée à ce jour.



Figure 1.5 – Le *reversing pseudophone* de Paul Thomas Young.

1.2 PLASTICITÉ EN LOCALISATION AUDITIVE

La question de savoir si les mécanismes de traitement des indices de localisation auditive pouvaient être modifiés, s'adapter à des contraintes par l'apprentissage de nouvelles associations entre les informations acoustiques et l'espace auditif, a été posée il y a déjà longtemps. En 1928, Paul Thomas Young passa 18 jours entre chez lui, son laboratoire et les rues de Berlin, équipé périodiquement du dispositif représenté figure 1.5, le *reversing pseudophone*, qui inversait l'arrivée des signaux acoustiques entre les deux oreilles. Après avoir porté le dispositif plus de 80 heures, Young n'a reporté aucune adaptation à la manipulation acoustique, sa perception sur le plan horizontal restant inversée jusqu'à la fin de l'expérience. Ce résultat a été confirmé plus de 70 ans plus tard par une réplique moderne de l'expérience (Hofman, Vlaming, Termeer, & Van Opstal, 2002). Même si elle peut paraître anecdotique, cette manipulation drastique des informations acoustiques montre une des limites de la plasticité du système auditif et suggère que certains de ses mécanismes sont solidement ancrés.

Bien que l'expérience radicale de Young ne fournisse pas de preuve de la plasticité du système auditif en localisation spatiale, la capacité du cerveau à s'adapter à des indices de localisation auditive biaisés existe et est nécessaire. Nous avons vu dans la première partie de cette introduction, que les indices de localisation dépendent fortement de la morphologie

de l'auditeur et que le traitement neuronal de ces indices pour aboutir à la perception d'un espace auditif était complexe. Les variations anatomiques entre individus rendent impossible une connaissance *a priori* des associations entre indices acoustiques et positions spatiales, car elles ne peuvent pas être anticipées par l'encodage génétique (Rauschecker, 1999). La localisation auditive doit donc être apprise par l'expérience. De plus, la taille de la tête ainsi que la forme des pavillons sont constamment modifiés pendant le développement, rendant l'ajustement des mécanismes de localisation obligatoires durant cette période (Clifton, Gwiazda, Bauer, Clarkson, & Held, 1988 ; Hartley & King, 2010). Une aptitude naturelle à l'adaptation du traitement des indices de localisation auditive doit alors exister.

Les bases neurophysiologiques de cette plasticité développementale ont été étudiées chez l'animal, en particulier la chouette effraie, qui chasse ses proies de nuit grâce à d'exceptionnelles capacités de localisation auditive ; et le furet, dont la vision limitée l'oblige à se fier principalement à ses capacités de localisation auditive et son odorat pour se diriger. Ces études ont permis une meilleure compréhension de la maturation du système auditif et de la construction d'une représentation de l'espace auditif. L'étude des phénomènes de plasticité en localisation auditive chez l'animal adulte pose un cadre de réflexion pour l'interprétation des stratégies d'adaptation chez l'être humain.

1.2.1 Plasticité développementale

Les travaux qui étudient la plasticité développementale observent le comportement et la maturation du système auditif d'animaux élevés dans des conditions anormales. Deux types de manipulations ont classiquement été employées pour l'étude du développement de la représentation de l'espace auditif : la déviation du champ visuel par des prismes ; la perturbation des indices acoustiques.

1.2.1.1 Déviation du champ visuel

Knudsen et Knudsen (1989) ont élevé des chouettes effraies équipées de prismes qui déplaçaient leur champ de vision vers la droite. Une fois en âge d'être testées comportementalement,

ces chouettes orientaient la tête vers la droite de cibles visuelles ou auditives, d'un angle équivalent à celui des prismes qu'elles portaient. Les auteurs ont alors conclu que ces animaux avaient développé une association entre les indices de localisation auditive et leur perception de l'espace auditif qui concordait avec leur perception de l'espace visuel, bien que celle-ci soit biaisée. Une fois les prismes retirés, les chouettes orientaient la tête directement vers les cibles visuelles, mais le décalage persistait avec des stimuli auditifs.

Les corrélats neuronaux de cette plasticité ont été étudiés en comparant les réponses électrophysiologiques pour un groupe de chouettes élevées avec des prismes et un groupe de chouettes élevées normalement (Brainard & Knudsen, 1993). Les auteurs ont comparé la sélectivité neuronale aux ITDs dans différentes structures sous-corticales : l'IC et le tectum optique⁴ (OT, *Optic Tectum*). L'OT renferme des neurones organisés topographiquement, qui sont à la fois sensibles aux directions spatiales visuelles et auditives. Une représentation topographique de l'espace auditif existe dans le noyau externe de l'IC, noyau qui partage des projections ascendantes et descendantes avec l'OT. Les auteurs ont observé un décalage de la représentation des ITDs à la fois dans l'OT et dans le noyau externe de l'IC. Par exemple, chez une chouette élevée avec des prismes déviant le champ visuel de 23° vers la gauche, une cellule qui répondait préférentiellement à un stimulus visuel à 0° sans prisme, répondait de manière maximale à une ITD de +52 μ s, ITD correspondant à un angle de 20° vers la droite. Ces résultats ont démontré que la représentation neuronale des indices de localisation auditive est affectée par l'expérience et la perception visuelle. Des résultats similaires ont été observés au niveau du SC de furets dont l'orientation d'un oeil avait été déviée de manière chirurgicale au début de leur vie (King, Hutchings, Moore, & Blakemore, 1988).

1.2.1.2 Perturbation des indices acoustiques

La perturbation des indices acoustiques est une manière plus directe d'étudier la plasticité en localisation auditive. Le moyen de perturber les indices acoustiques le plus simple et le plus utilisé est l'insertion unilatérale d'un bouchon d'oreille. La manipulation a pour effet d'atténuer

⁴Homologue aviaire du colliculus supérieur.

le signal acoustique à l'entrée d'une oreille, et modifie alors les indices binauraux, en particulier le ILDs. Les ITDs ne sont modifiés que de façon minime. Au niveau comportemental, l'insertion du bouchon va initialement biaiser la perception auditive du côté de l'oreille libre.

Il a été montré que de jeunes chouettes équipées d'un tel bouchon, pouvaient retrouver leur capacité à localiser les sons de manière précise après quelques semaines (Knudsen, Esterly, & Knudsen, 1984). Au cours des premiers tests effectués après avoir retiré le bouchon, les performances de localisation des chouettes étaient déviées dans la direction opposée de la déviation initialement induite par le bouchon. Cet effet consécutif (*aftereffect*) disparut progressivement après quelques jours passés sans le bouchon.

Mogdans et Knudsen (1992) ont comparé la sélectivité neuronale de l'OT de chouettes élevées normalement ou avec un bouchon d'oreille. L'étude donna des résultats analogues aux études utilisant des prismes. Les neurones d'OT ayant un champ récepteur visuel correspondant à une direction neutre par rapport à la tête de la chouette, répondaient de façon maximale à des ITDs et ILDs déviés dans la direction de l'oreille libre. Cette déviation de sélectivité neuronale a également été observée au niveau du noyau externe de l'IC (Mogdans & Knudsen, 1993).

Des études similaires avec des furets ont donné des résultats étonnamment différents. Les furets eux aussi développent des facultés de localisation auditive normale quand ils sont élevés avec un bouchon d'oreille, cependant aucun effet consécutif n'est observé après le retrait du bouchon (King, Parsons, & Moore, 2000). De plus, les enregistrements électrophysiologiques montrent que ces mêmes furets développent une représentation de l'espace auditif proche de la normale au niveau de SC. Contrairement aux chouettes, aucun décalage entre les sélectivités auditives et visuelles n'a été observé.

1.2.1.3 Construction d'une représentation de l'espace auditif

Cet ensemble de résultats met en exergue le rôle de guide que semble jouer la vue pour la calibration de l'espace auditif. Lorsque l'espace visuel ou auditif est perturbé, la carte de l'espace auditif semble se développer de façon à être en accord avec la carte de l'espace visuel. Mais la vue ne semble jamais être calibrée par l'audition. De par son organisation entièrement

topographique, il est probable que la vue procure une vérification fiable de l'exactitude des associations entre indices acoustiques et positions des sources sonores. La vue pourrait ainsi guider la plasticité quand les univers auditif et visuel sont décalés. Knudsen et Mogdans (1992) ont cependant montré que la vue ne peut être le seul guide de cette recalibration. Les auteurs ont répété l'expérience d'occlusion monaurale avec des chouettes auxquelles on avait suturé les paupières. Comme chez les chouettes pour qui la vue était préservée, les chouettes privées de vision présentaient un décalage de la sélectivité neuronale des indices binauraux. Bien que ce décalage était moindre que pour les chouettes qui bénéficiaient de la vue, ce résultat montra que le système auditif avait pu procéder à des ajustements adaptatifs sans l'influence de cette dernière.

Le fait que les furets élevés avec un bouchon d'oreille développent une carte de l'espace auditif presque normale dans SC, associé avec leur acquisition de capacités de localisation auditive normales elles aussi, semble indiquer que cette plasticité n'est pas le résultat d'un changement de sélectivité neuronale de l'espace auditif comme observé chez la chouette. Des stratégies d'adaptation différentes, pouvant refléter les différentes stratégies d'encodage des indices acoustiques entre oiseaux et mammifères (cf. 1.1.3. Encodage des indices acoustiques), semblent exister chez les deux espèces. L'absence à la fois, d'une modification de la sélectivité neuronale dans SC, et d'un effet consécutif au niveau comportemental, plaide en faveur d'une hypothèse de repondération des indices de localisation auditive comme stratégie d'adaptation chez le furet (King et al., 2011). Cette hypothèse suppose que les indices perturbés seraient mis de côté dans le calcul de la position des sources sonores, au profit d'indices non altérés. Dans le cas d'une perturbation par l'insertion d'un bouchon d'oreille, les ITDs constituent l'indice le moins affecté, et une repondération en faveur de cet indice a été proposée (Kacelnik, Nodal, Parsons, & King, 2006).

1.2.2 Existence d'une période critique ?

La plasticité du cerveau pendant la période de développement est considérable et permet aux individus de s'adapter à certaines contraintes environnementales qui ne peuvent être prédites par

le patrimoine génétique (Rauschecker, 1999). La période pendant laquelle cette plasticité est à son apogée est appelée période critique ou période sensible (voir : Hensch, 2004, pour une revue de littérature). L'existence de cette période critique a été démontrée pour différentes fonctions du système auditif (Chang & Merzenich, 2003 ; de Villers-Sidani, Chang, Bao, & Merzenich, 2007 ; Nakahara, Zhang, & Merzenich, 2004), mais l'apprentissage est encore possible après cette période (Polley, Steinberg, & Merzenich, 2006 ; Recanzone, Schreiner, & Merzenich, 1993). Dans le cadre de la localisation auditive, les premiers travaux ont fortement suggéré l'existence d'une période critique. Les résultats d'études plus récentes ont cependant montré que la plasticité en localisation auditive était aussi possible à l'âge adulte.

L'âge de l'animal au moment où on lui impose des prismes ou un bouchon d'oreille, joue fortement sur sa capacité d'adaptation à la manipulation ou de récupération post-manipulation. L'âge auquel on retire le bouchon d'oreille avec lequel des chouettes ont grandi, influe par exemple sur leur capacité à retrouver des performances normales sans bouchons (Knudsen, Knudsen, & Esterly, 1984). Plus le bouchon est retiré tard, moins la récupération est rapide, et si l'âge des chouettes est supérieur à 38–42 semaines au moment du retrait du bouchon, aucune récupération n'est observée. Ceci suggère que la récupération⁵ ne peut avoir lieu qu'au cours d'une période critique, pendant le développement. Il est intéressant de noter que les chouettes qui avaient eu une certaine expérience de la situation sans bouchon avant qu'un bouchon ne leur fût imposé, étaient capables de recouvrer des capacités de localisation normales, même si le retrait du bouchon était effectué après la fin de la période critique. L'effet de l'âge a également été démontré au niveau des modifications de sélectivité neuronales dans OT chez les chouettes portant des prismes (Knudsen, 1998). L'imposition de prismes à l'âge adulte ne décale pas la sélectivité aux ITDs. Cependant, si une chouette a grandi avec des prismes, s'est réadaptée à la situation sans prismes et montre alors une sélectivité neuronale aux ITDs normale, une réinstallation des prismes à l'âge adulte entraînera à nouveau une modification de la sélectivité neuronale. Ces résultats, ainsi que ceux de Knudsen, Knudsen et Esterly (1984), suggèrent que pour qu'une adaptation à l'âge adulte soit possible, il faut que les connexions nécessaires à cette

⁵Qui peut dans ce cas être considérée comme une réadaptation à la situation sans bouchon.

adaptation aient déjà été construites pendant le développement. Ces études plaident fortement pour l'existence d'une période critique ou sensible pour la plasticité en localisation auditive, mais des travaux ultérieurs vont montrer que cette plasticité est possible à l'âge adulte, même sans exposition préalable aux conditions altérées pendant le développement.

Si plutôt que d'imposer directement des prismes à 23° à la chouette adulte⁶, on lui installe d'abord des prismes à 6° les premières semaines, puis 11° les semaines suivantes, 17°, et enfin 23°, la chouette s'adapte, et un décalage de sélectivité neuronale comparable à celui observé chez les jeunes chouettes se produit (Linkenhoker & Knudsen, 2002). L'imposition graduelle de petites divergences entre les indices acoustiques et visuels permet donc une adaptation à l'âge adulte. Les auteurs ont également montré qu'après avoir retiré les prismes et laissé le temps aux chouettes de recouvrer une sélectivité neuronale neutre, les chouettes se réadaptaient rapidement à une imposition directe de prismes à 23°, et la sélectivité neuronale était à nouveau modifiée. En plus de montrer que la plasticité à l'âge adulte est possible, cette étude indique donc que les associations préalablement apprises ne sont pas oubliées et peuvent refaire surface si besoin.

1.2.3 Importance du contexte comportemental

Des études plus récentes ont montré l'importance du contexte comportemental lors de l'adaptation à des indices altérés. Les chouettes qui ont fait l'objet des études recensées précédemment sont en général gardées captives dans des cages et nourries d'animaux morts⁷. Bergan, Ro, Ro et Knudsen (2005) ont avancé l'idée que forcer les chouettes à utiliser leurs aptitudes de localisation auditive pour survivre devrait favoriser l'adaptation aux prismes. Ils ont en effet montré que dix semaines après l'installation de prismes à 17°, des chouettes adultes devant, pour survivre, chasser des proies vivantes dans la pénombre, présentaient une déviation de la sélectivité neuronale aux ITDs dans OT cinq fois plus grande que des chouettes qu'on nourrissait

⁶Prismes d'un "angle d'ouverture" de 23°.

⁷Elles ont toutefois l'occasion de voler de temps en temps dans de grands espaces.

de souris mortes. En outre, l'importance de la déviation corrélait avec les performances de chasse retrouvées pendant la période d'adaptation.

Tout comme les chouettes, les furets s'adaptent mieux à l'insertion d'un bouchon d'oreille lorsqu'ils sont jeunes, qu'une fois adultes (King et al., 2000). Mais comme chez la chouette, le contexte comportemental du furet est important. L'entraînement facilite en effet le processus d'adaptation. Kacelnik, Nodal, Parsons et King (2006) ont montré une forte corrélation entre adaptation et fréquence d'entraînements à une tâche de localisation auditive. Ils ont observé une adaptation beaucoup plus importante et plus rapide chez un groupe entraîné tous les jours, par rapport à un groupe entraîné une fois tous les six jours, ou un groupe n'ayant pas suivi d'entraînement. Les auteurs ont également étudié l'adaptation à un bouchon d'oreille pour un groupe auquel on avait supprimé la vue et ont montré que ce groupe réussissait aussi bien à s'adapter que les autres. Ce résultat suggère que bien que la vue ait un rôle de guide lorsqu'elle est présente, d'autres modalités semblent pouvoir procurer un *feedback* suffisant pour que l'adaptation puisse se mettre en place.

1.2.4 Implication du cortex auditif

La plasticité et l'implication du cortex auditif dans différents types d'apprentissage perceptif auditif ont été démontrées à maintes reprises (de Villers-Sidani et al., 2007 ; Nakahara et al., 2004 ; Polley et al., 2006 ; Recanzone et al., 1993). Nous avons précédemment mentionné l'importance du cortex auditif pour la perception de l'espace auditif (cf. 1.1.2.2. *Cortex auditif*), il s'avère crucial pour sa recalibration. Il a récemment été montré que l'adaptation à une occlusion monaurale qu'on observe chez le furet adulte sain, est inexistante si le cortex auditif est retiré par ablation (Nodal et al., 2010). En effet, un furet auquel une partie du cortex auditif a été retirée peut encore localiser des sons de longue durée, mais sera incapable de les localiser après l'insertion d'un bouchon d'oreille et plusieurs semaines d'entraînement. Ce constat était identique que la lésion soit limitée au cortex auditif primaire ou qu'elle touche également les régions voisines.

Le rôle que joue le cortex pour la plasticité en localisation auditive est probablement de

guider la recalibration au niveau sous-cortical. Les projections corticofugales jouent un rôle critique dans différents types de plasticité auditive (Suga, Xiao, Ma, & Ji, 2002 ; Yan & Suga, 1998), et une étude récente a montré que l’ablation de ces projections chez le furet empêchait toute adaptation à une occlusion monaurale (Bajo, Nodal, Moore, & King, 2010).

1.2.5 Études chez l’humain

Les études sur les animaux peuvent permettre de faire des inférences sur le modèle humain, mais il faut garder à l’esprit que les mécanismes auditifs varient d’une espèce à l’autre, rendant certaines comparaisons ardues. Alors que les études de la plasticité en localisation auditive chez l’animal n’ont utilisé que des manipulations qui avaient une incidence sur le plan horizontal, les études chez l’humain ont plus souvent manipulé la perception du plan vertical. Plusieurs études ont cependant cherché à savoir si l’adulte pouvait s’adapter à des perturbations des indices binauraux.

1.2.5.1 Perturbation des indices binauraux

Le paradigme expérimental d’occlusion monaurale par un bouchon d’oreille utilisé avec l’animal, l’est également avec l’humain. Il permet de simuler une perte auditive unilatérale, situation qui peut survenir après un traumatisme acoustique ou physique, après le développement d’un neurinome de l’acoustique⁸, suite à une labyrinthite, l’apparition de la maladie de Menière, etc... Il a été montré que l’insertion d’un bouchon déviait les réponses des participants du côté opposé au bouchon, et qu’une adaptation modérée est possible lorsque le bouchon est porté continûment pendant plusieurs jours (Bauer, Matuzsa, Blackmer, & Glucksberg, 1966 ; Florentine, 1976). Lorsque le protocole de l’étude comprend des sessions d’entraînement discrètes, plutôt que le port continu du bouchon, l’adaptation est moins nette, voire absente, et semble être spécifique au stimulus d’entraînement (Butler, 1987 ; McPartland, Culling, & Moore, 1997). Ce constat reflète l’hypothèse d’importance du contexte comportemental vu chez l’animal, et montre que la

⁸Tumeur bénigne du nerf vestibulocochléaire.

plasticité est mieux observée si les participants ont l'occasion de s'adapter dans des situations naturelles.

Kumpik, Kacelnik et King (2010) ont montré qu'en plus de porter le bouchon continûment, un entraînement quotidien améliorerait les chances d'adaptation. Les auteurs émettent l'hypothèse que l'adaptation à l'occlusion monaurale est en partie due à une dépendance accrue des indices spectraux. C'est l'hypothèse de repondération vers des indices plus fiables que nous avons vue avec le furet (cf. 1.2.1.3. *Construction d'une représentation de l'espace auditif* et King et al., 2011). Cette hypothèse est à prendre avec précaution car l'étude comprend certaines limitations importantes. Lors des tests de localisation, la façon dont les participants indiquaient d'où provenait les sons qui leurs étaient présentés, était de choisir entre 12 haut-parleurs couvrant les 360° du plan horizontal. Cette méthode ne reflète pas entièrement la perception de l'auditeur puisqu'il est contraint de choisir entre plusieurs directions imposées. Si un son est entendu entre deux haut-parleurs, l'auditeur devra donc indiquer le haut-parleur le plus proche de son percept et ne sera pas libre de pointer la position d'où il a effectivement entendu provenir le son. La résolution spatiale que procurent les 12 haut-parleurs est seulement de 30°. Cette faible résolution empêche de détecter des déviations induites par le bouchon d'oreille si elles sont inférieures à 15°. Il semble que cela ait été le cas, car cette étude ne montre aucun déplacement des réponses des participants dans la direction opposée à l'oreille dans laquelle est inséré le bouchon, mais seulement une hausse du nombre d'erreurs de localisation, principalement en matière de confusions entre l'avant et l'arrière. Les auteurs rapportent en effet que les améliorations de performance observées pendant la période d'adaptation sont surtout dues à une baisse des confusions entre l'avant et l'arrière. Comme les indices spectraux permettent la résolution de ce type d'erreurs, les auteurs concluent que l'adaptation à une occlusion monaurale doit être effectuée par une repondération vers les indices spectraux. Cependant, comme aucune déviation perceptive sur le plan horizontal n'a pu être observée au moment de l'insertion du bouchon, il est impossible de savoir si une quelconque adaptation a eu lieu sur le plan horizontal et si les résultats de l'étude ne reflètent pas seulement une adaptation à la modification des indices spectraux causée par le bouchon.

D'autres types de modification des indices binauraux ont été expérimentés. Une première étude (Held, 1955), utilisa un dispositif constitué d'appareils auditifs dont les microphones étaient déplacés (une rotation sur l'axe vertical), permettant de pivoter artificiellement l'espace auditif sur le plan horizontal. Les yeux bandés, les participants produisaient d'abord une erreur de 22° quand ils devaient localiser un haut-parleur cible. Après seulement 7 h passées dans leur environnement de travail habituel et équipés du dispositif, les participants présentaient une adaptation d'environ 10°.

Trente ans plus tard, Javer et Schwarz (1995) ont utilisé des appareils auditifs dans lesquels une ligne à retard était ajoutée. Introduire un retard dans une oreille est un moyen efficace de décaler les ITDs. L'ajout du retard provoque une déviation sur le plan horizontal de la direction perçue par l'auditeur et place les ITDs en contradiction avec les autres indices de localisation auditive, ainsi qu'avec les informations spatiales des autres modalités sensorielles. Les 12 participants de cette étude devaient porter ces appareils auditifs du matin au soir sur une période de quatre à six jours. Trois différentes durées de retard ont été testées, 342, 513 et 684 μs ⁹. Les performances de localisation auditive des participants étaient testées avant et après le port des appareils auditifs et quotidiennement durant la période d'adaptation aux appareils. La tâche de localisation auditive se déroulait en champ libre et consistait à ajuster la position d'un haut-parleur caché derrière un rideau, de façon à ce que les sons émis de cet haut-parleur soient perçus exactement en face de l'auditeur. Juste après insertion des appareils, la déviation perceptive moyenne était de 30° pour les participants testés avec un retard de 342 μs , de 66° pour ceux dont le retard était de 513 μs et 68° pour le seul participant testé avec un retard de 684 μs . Ce déplacement initial s'est en moyenne graduellement estompé durant les 3 à 5 jours d'exposition au retard. Certains participants montraient de larges progrès après seulement quelques heures. Au dernier jour d'expérience, d'importantes différences interindividuelles ont été observées quant au degré d'adaptation, d'une adaptation nulle à complète¹⁰. Les résultats du participant pour lequel un retard de 684 μs avait été introduit sont difficilement interprétables,

⁹ 684 μs correspond à peu près à l'ITD que procure un son provenant d'un angle de 90°.

¹⁰ La durée du retard imposé et le degré d'adaptation n'étaient pas corrélés.

car il présentait une déviation perceptive de 78° du côté opposé à la déviation initiale, lors des dernières tâches de localisation avec les appareils auditifs. Ceci est probablement dû à un mauvais fonctionnement des appareils auditifs ou à une confusion quelconque de la part du participant. La capacité d'adaptation à un retard supérieur à $600\ \mu\text{s}$ reste donc encore à prouver. Immédiatement après avoir retiré les appareils auditifs, un effet léger consécutif, très faible en comparaison de la déviation perçue initialement (5° en moyenne, absent chez certains participants) a été mesuré chez quelques participants, effet consécutif qui disparaissait après quelques minutes. Alors qu'il était largement admis à l'époque de cette publication que le modèle de Jeffress devait refléter l'encodage des ITDs chez l'humain, les auteurs ont souligné le fait que leurs résultats étaient en partie incompatibles avec ce modèle. Une recalibration des ITDs selon le modèle de Jeffress nécessiterait des modifications des dimensions des axones qui constituent les lignes à retard, ce qui paraît improbable en quelques heures d'adaptation et incompatible avec l'effet consécutif quasiment absent.

Les deux dernières études que nous venons de présenter (Held, 1955 ; Javer & Schwarz, 1995), sont les seuls travaux qui avaient pour but de tester une adaptation, dans des situations naturelles du quotidien, à une modification qui incluait les ITDs. Un décalage des ITDs peut survenir avec certaines maladies de l'oreille moyenne (Hartley & Moore, 2003 ; Lupo, Koka, Thornton, & Tollin, 2011), en particulier l'otite séro-muqueuse, qui se caractérise par des épanchements rétro-tympanique de mucus qui retardent la propagation du signal acoustique. Ce retard cause un décalage des ITDs pouvant facilement atteindre les $600\ \mu\text{s}$ (Thornton, Chevallier, Koka, Lupo, & Tollin, 2012). Le décalage des ITDs est également l'un des effets secondaires de l'installation d'appareils auditifs (Kuk & Korhonen, 2014). Les études de Held (1955), et en particulier Javer et Schwarz (1995), ont montré que l'adaptation à ce type de perturbation était possible. Cependant, les protocoles de test de localisation auditive de ces expériences restreignent la position des sources sonores à la direction droit devant. Ils ne permettent donc pas d'apprécier l'effet de la modification des ITDs sur l'ensemble du plan horizontal, ni de savoir si l'adaptation est homogène sur l'ensemble de l'espace auditif. Les tests mis en place dans ces études ne permettent pas d'émettre des hypothèses sur les stratégies adaptatives employées par le système auditif.

On ignore également si une persistance de l'adaptation existe, c'est-à-dire si l'apprentissage qui a eu lieu pendant la période d'adaptation est retenu après quelques jours passés sans le dispositif. Une limite technique de l'étude de Javer et Schwarz (1995), vient du fait que les appareils auditifs n'étaient pas moulés aux oreilles des participants et ne procuraient donc pas d'atténuation acoustique. Ceci implique qu'en plus du signal des appareils auditifs comprenant la ligne à retard, le signal acoustique naturel, sans décalage des ITDs, atteignait également les tympans des participants. Pour palier ce problème, les auteurs ont cherché à masquer le signal naturel en introduisant un gain compris entre 5 et 15 *dB* dans les appareils. La portion de signal acoustique naturel que pouvait entendre chaque participant était variable et incontrôlable, et a certainement joué sur l'importante variabilité des résultats.

1.2.5.2 *Perturbation des indices spectraux*

Nous avons vu que les indices spectraux reposent essentiellement sur le filtrage procuré par la forme des pavillons (cf. 1.1.1.2. **Les indices spectraux**). Modifier la forme des pavillons crée donc un nouveau filtrage, et la manipulation va perturber la localisation auditive. Afin de pouvoir utiliser à nouveau les indices spectraux, le système auditif devra apprendre les associations entre les nouvelles DTFs et les positions dans l'espace sonore. Une modification progressive des indices spectraux se produit tout au long de la vie avec l'évolution de la forme de l'oreille (Clifton et al., 1988), mais peut également survenir brusquement après un accident ou encore l'installation d'appareils auditifs (Byrne & Noble, 1998).

La première étude de l'adaptation à de nouveaux indices spectraux est celle de Hofman, Van Riswick et Van Opstal en 1998. Dans cette étude, des moulages en silicone (*silicone molds*) ont été placés dans les cavités pavillonnaires de quatre participants¹¹, qui les ont portés continûment pendant plusieurs semaines. L'insertion des moulages dégrada considérablement les performances de localisation des participants sur le plan vertical. Les auteurs ne rapportent pas d'altération de la performance sur le plan horizontal. Au cours de la période d'adaptation, la dégradation sur le plan vertical a diminué progressivement, jusqu'à une stabilisation des

¹¹Les trois auteurs et un sujet naïf.

performances à un niveau proche de la situation initiale sans moulages. La vitesse d'adaptation était variable entre les participants, de 19 à 39 jours. À la fin de l'expérience, les aptitudes de localisation des participants ont été testées immédiatement après qu'ils eurent retiré les moulages et aucun effet consécutif n'a été observé. La performance des participants lors de ce dernier test était en effet similaire à la performance initiale. C'est ce résultat qui était le plus stupéfiant. L'apprentissage de nouveaux indices spectraux n'avait pas remplacé les associations DTFs / espace sonore d'origine. Les auteurs concluent en faisant un parallèle entre ce type d'adaptation et l'apprentissage d'un nouveau langage.

Cette étude laisse cependant quelques questions en suspens. Est-ce qu'après à un retour à la situation normale, sans moulages, les indices spectraux sont retenus par le système auditif ? Est-ce que le système auditif a besoin de certains indices (visuels, tactiles, dynamiques) pour correctement localiser lors du passage d'une situation à une autre ? C'est-à-dire est-ce que quelques stimulations sont nécessaires quand les moulages sont retirés, ou bien la performance retourne-t-elle à son niveau d'origine dès le premier essai ?

Certaines études de la plasticité développementale ont suggéré que la vue jouait un rôle de guide pour la recalibration de l'espace auditif (cf. **1.2.1. Plasticité développementale**). Une étude récente a analysé l'amplitude et la vitesse d'adaptation à une modification des indices spectraux, à l'intérieur et en dehors du champ visuel (Carlile & Blackman, 2013). Les auteurs n'ont observé aucune différence d'adaptation entre ces deux régions de l'espace auditif. Cette étude, ainsi que la capacité maintes fois prouvée des non-voyants à développer et maintenir des aptitudes de localisation auditive supérieures à la normale (Doucet et al., 2005 ; Lessard, Paré, Lepore, & Lassonde, 1998 ; Röder et al., 1999), montrent que les informations visuelles ne sont pas primordiales pour la recalibration de l'espace auditif. Carlile et Blackman (2013) ont également examiné la persistance de l'adaptation. Les participants ont réinséré les moulages le temps d'un test de localisation qui a eu lieu après une semaine passée sans porter les moulages. Les résultats ont montré qu'environ 70 % de l'adaptation avait été retenue. La semaine passée sans moulages est courte comparée aux 30 à 60 jours pendant lesquels les moulages étaient

portés. Il serait intéressant de savoir si l'apprentissage peut être retenu pour de plus longues périodes.

Une autre étude du même groupe (Carlile, Balachandar, & Kelly, 2014), a comparé les impacts de différents types d'entraînement sur l'adaptation à des moulages en silicone. Pendant les 10 jours de port des moulages, les participants effectuaient des sessions d'entraînement quotidiennes d'une heure chacune. Les participants étaient séparés en quatre groupes en fonction du type d'entraînement qu'ils suivaient. Le groupe contrôle effectuait uniquement une tâche de localisation standard dans la pénombre et sans *feedback* ; un deuxième groupe recevait un *feedback* visuel (LED) après chaque réponse de localisation sonore ; pour un troisième groupe, après le *feedback* visuel, le son était joué de manière répétée depuis la position d'où il avait été émis avant la réponse, et l'auditeur devait tourner la tête jusqu'à ce qu'elle soit pointée en direction de la source sonore ; un dernier groupe suivait un entraînement identique au troisième, mais les lumières de la pièce étaient allumées, ce qui offrait un cadre visuel de référence. Les résultats ont montré que l'entraînement avec *feedback* visuel accélérait assez peu le processus d'adaptation, et que les entraînements dans lesquels une composante sensorimotrice était engagée (troisième et quatrième groupes) avaient l'impact le plus important sur la quantité et le taux d'adaptation. Ce résultat confirme ceux d'une étude sur l'apprentissage de nouveaux indices spectraux en environnement auditif virtuel (Parseihian & Katz, 2012), qui montrait également l'importance de cette composante sensorimotrice pour l'apprentissage

Des études ont utilisé le paradigme expérimental de modification des indices spectraux afin d'étudier certains mécanismes de localisation auditive (Hofman & Van Opstal, 2003 ; M. M. V. Van Wanrooij & Van Opstal, 2005). Ces études n'ont modifié les indices spectraux que dans une oreille et ont permis de montrer que la perception de l'élévation était en partie attribuable à des interactions binaurales, à partir desquelles les associations DTFs / espace sonore de chaque oreille sont pondérées par la perception de l'azimut.

Notons qu'aucune recherche en neuroimagerie n'a étudié les mécanismes de la plasticité en localisation auditive chez l'humain.

1.3 OBJECTIFS ET MÉTHODES

La présente recherche avait pour objectif d'étudier la plasticité du système auditif humain dans des tâches de localisation sonore, ainsi que les mécanismes d'encodage des indices de localisation auditive, tant sur le plan comportemental, qu'à travers les corrélats neuronaux du comportement observé. Dans chacun des quatre articles regroupés dans cette thèse, les mécanismes d'une potentielle plasticité étaient stimulés par la perturbation d'indices de localisation auditive. Les deux premiers articles sont consacrés à la modification de l'un des indices binauraux, les ITDs. Dans les deux derniers articles, ce sont les indices spectraux qui ont été modifiés.

Dans chaque expérience, la perturbation des indices acoustiques était effectuée par un dispositif intra-auriculaire qui permettait aux participants de s'adapter à la manipulation à travers leurs activités quotidiennes, en dehors du laboratoire. Bien qu'elle suppose de renoncer au contrôle de certains paramètres expérimentaux, une exposition continue à la manipulation dans des situations naturelles est une méthode efficace et écologique qui offre les meilleurs chances d'adaptation (cf. 1.2.3. **Importance du contexte comportemental** et 1.2.5. **Études chez l'humain**).

Pour modifier les ITDs, nous avons été les premiers à utiliser des bouchons d'oreille numériques. Ces dispositifs sont moulés aux conduits auditifs, permettant ainsi une étanchéité acoustique qui empêche la majeure partie du signal acoustique naturel d'atteindre le tympan. Le conduit auditif naturel est alors remplacé par un conduit numérique sous contrôle expérimental. Le moulage sur mesure procure une atténuation des ondes acoustiques parvenant au tympan de 25 à 50 dB suivant les fréquences. L'atténuation peut être mesurée alors que les bouchons sont installés dans les oreilles des participants par une comparaison du signal à l'entrée et à la sortie du bouchon. Les bouchons comprennent un microphone à l'extérieur, un processeur de signal numérique programmable et un transducteur (voir figure 1.6.A). Afin de ne pas affecter les ILDs, le gain des bouchons était identique pour les deux oreilles et il était ajusté de sorte que les bouchons procurent des niveaux de sonie naturels. La manipulation des ITDs était réalisée

par l'ajout d'un retard dans la chaîne de traitement du signal d'un des deux bouchons. La durée de retard imposée était toujours de $625 \mu s$, soit une durée proche de l'ITD maximale que peut procurer l'écartement de nos deux oreilles. Cette durée a été choisie afin que l'effet observé soit maximal.

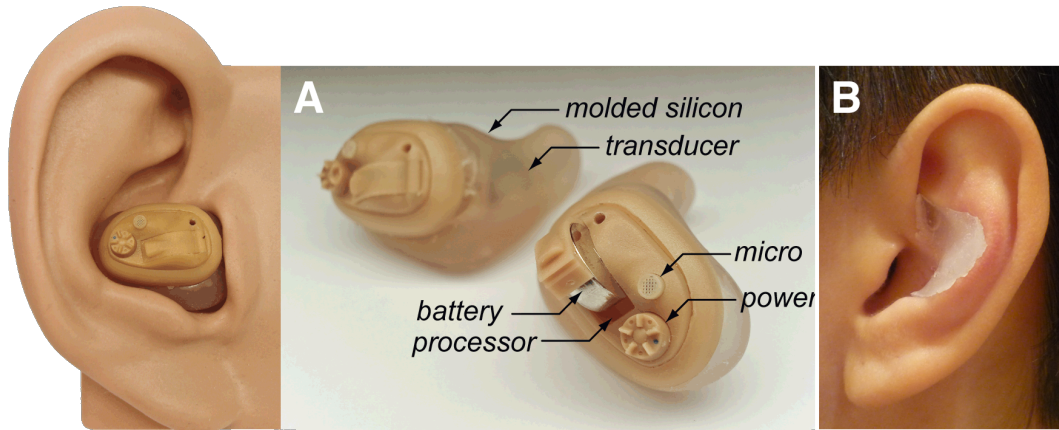


Figure 1.6 – Bouchons d'oreille numériques et moulage en silicone. A) Bouchons d'oreille numériques. B) Exemple de moulage en silicone inséré dans l'oreille gauche d'un participant.

Dans les deux dernières études, la modification des indices spectraux était effectuée en employant la même méthode que dans les précédentes études qui ont cherché à modifier cet indice. Des moulages en silicone n'obstruant pas le canal auditif étaient insérés dans la conque du pavillon (voir figure 1.6.B).

Tous les tests perceptifs de cette recherche se sont déroulés à l'université de Montréal, dans la chambre héli-anéchoïque du laboratoire international de recherche sur le cerveau, la musique et le son (BRAMS). Un dôme de 80 haut-parleurs était utilisé pour les tests de localisation auditive. L'utilisation de ce dôme permettait de présenter rapidement des stimuli auditifs provenant de sources sonores réparties dans l'ensemble de l'espace auditif. Les participants indiquaient leur perception de la provenance des sons en orientant la tête en direction de la source sonore. Un laser était attaché à leur tête, leur permettant d'avoir un retour sur la direction pointée. Cette technique a l'avantage d'être hautement écologique et de permettre une collecte précise et rapide des données par *head-tracking* (Carlile, Leong, & Hyams, 1997). Toutes les tâches étaient

contrôlées par des interfaces *Tucker-Davis Technologies* pilotées par des scripts *Matlab* tous écrits au cours de la recherche.

La technique de neuroimagerie choisie au cours de cette recherche pour étudier les corrélats neuronaux de la perception de l'espace auditif et de sa plasticité était l'imagerie par résonance magnétique fonctionnelle (fMRI, *Functional Magnetic Resonance Imaging*). Cette technique non-invasive repose sur l'enregistrement des variations hémodynamiques entraînées par l'activité cérébrale grâce à la capture du signal BOLD (*Blood-Oxygen-Level Dependent*). Il a été montré que le signal BOLD reflète étroitement les potentiels de champs locaux, et dans une certaine mesure le taux de décharge neuronal, dans le néocortex en général (Ekstrom, 2010), et plus précisément dans le cortex auditif primaire (Mukamel et al., 2005 ; Nir et al., 2007). L'utilisation de la fMRI permet une résolution spatiale supérieure aux autres techniques de neuroimagerie non-invasive. La résolution utilisée lors des études présentées dans cette thèse était de $1.5 \times 1.5 \times 2.5 \text{ mm}$. Le volume d'un voxel était alors de 5.625 mm^3 , soit environ cinq fois plus petit que le voxel de 3 mm^3 généralement utilisé. L'analyse des données fMRI était effectuée en utilisant *FMRISTAT* (Worsley et al., 2002) et des scripts *Matlab* et reposait principalement sur la construction de courbes d'accord pour chaque voxel. Les séances en fMRI se sont déroulées à l'Unité de Neuroimagerie Fonctionnelle (UNF) du Centre de recherche de l'Institut universitaire de gériatrie de Montréal (CRIUGM) et au Montreal Neurological Institute (MNI).

Il est parfois suggéré que la fMRI n'est pas la technique la plus appropriée pour étudier la perception auditive en raison du bruit intense émis par le scanner lors des acquisitions d'images. Cependant la technique de *sparse sampling* (Hall et al., 1999) permet de s'affranchir de cette contrainte par l'espacement temporel des acquisitions fonctionnelles et par leur placement à la fin de longs stimuli auditifs, qui ne sont alors pas masqués par le bruit du scanner. Une seconde contrainte de la fMRI est l'impossibilité d'utiliser des haut-parleurs pour diffuser les stimuli auditifs en raison du champ magnétique du scanner. L'utilisation d'écouteurs pour l'investigation de la perception de la position de sources sonores peut alors paraître problématique, mais l'espace auditif peut être très fidèlement reproduit par la présentation d'enregistrements binauraux individuels (Wightman & Kistler, 1989a, 1989b). De tels enregistrements sont effectués depuis

le canal auditif de chaque participant et contiennent alors tous les indices acoustiques propres à la morphologie du participant. Ces enregistrements permettent également d'extérioriser la perception des stimuli auditifs.

Le premier article de la thèse est comportemental seulement et avait pour principal objectif de savoir si l'adaptation à des ITDs décalées de $625\ \mu s$ était possible. L'utilisation des techniques susmentionnées a permis d'observer l'effet de la manipulation, ainsi que l'adaptation à cette manipulation, sur l'ensemble du plan horizontal frontal. Les résultats de l'étude montrent que le système auditif est capable de rapidement s'adapter à la modification et qu'aucun effet consécutif à l'adaptation n'est observable.

Le second article rapporte la première tentative d'observation de la plasticité en localisation auditive chez l'humain par une technique de neuroimagerie fonctionnelle. Sur le plan comportemental cette étude est la continuité de la première. Elle comporte des séances en fMRI qui ont eu lieu avant et après que la plasticité ait pris place. Ces séances en neuroimagerie nous ont permis d'identifier des corrélats neuronaux de l'adaptation aux ITDs décalés, ainsi que d'observer certains mécanismes d'encodage de l'espace acoustique sur le plan horizontal.

Le troisième article décrit une étude comportementale qui était destinée à explorer les mécanismes d'apprentissages de nouveaux indices spectraux, par l'étude de l'absence d'effet consécutif à une adaptation à des indices spectraux modifiés, et de la persistance de cette adaptation sur de longues périodes. L'étude a également permis de révéler des corrélations entre différences interindividuelles comportementales et propriétés acoustiques des indices spectraux.

L'étude en neuroimagerie présentée dans le quatrième article avait pour but d'explorer les mécanismes encore inconnus de l'encodage de l'élévation sonore dans le cortex auditif. L'adaptation à des indices spectraux modifiés nous a permis d'étudier les corrélats neuronaux de la perception de l'élévation, en comparant des données obtenues pour des stimuli physiquement identiques mais perçus différemment, au cours de séances fMRI qui avaient lieu avant et après l'adaptation.

Au total, cette recherche a nécessité la participation de 72 personnes qui ont chacune été testée pendant 10 à 15 jours, totalisant plus de 1000 heures de testing, auxquelles il faut ajouter 80 séances d'imagerie par résonance magnétique fonctionnelle.

CHAPITRE 2

ARTICLE 1 : RELEARNING SOUND LOCALIZATION WITH DIGITAL EARPLUGS

Régis Trapeau and Marc Schönwiesner

Publié dans *Canadian Acoustics*, 39(3), 116-117.

2.1 INTRODUCTION

The auditory system infers the location of sound sources from the processing of different acoustic cues. As the size of the head and the shape of the ears change over development, the association between acoustic cues and our expectation of external spatial position can not be fixed at birth, but has to be plastic. Recent studies on humans have shown that the auditory system is still capable of such plasticity during adulthood ([Hofman et al., 1998](#); [Javer & Schwarz, 1995](#); [M. M. V. Van Wanrooij & Van Opstal, 2005](#)). We aim to explore the principles that govern an adaptation to shifted Interaural Time Differences (ITDs), one of the two binaural cues for azimuthal perception with the Interaural Level Differences (ILDs).

We equipped six participants with binaural digital earplugs that allow us to delay the input to one ear, and thus disrupt the ITDs. Participants were asked to wear the plugs during all waking hours for 10 days and their ability to localize sounds on the horizontal plane was tested everyday in free field conditions.

2.2 METHODS

2.2.1 Participants

Two female and four male students aged 26–32 years, with no history of hearing disorder or neurological disease, participated as paid volunteers, after having given informed consent. The experimental procedures were approved by the local ethics committee.

2.2.2 Plugs

Each plug contains a programmable signal processor, microphone and transducer and was fitted to the ears of each participant to create an acoustic seal (attenuation around 25 dB). The gain of the plugs was adjusted to achieve normal loudness levels. The plugs could be programmed to delay the incoming sound by a desired duration.

2.2.3 Design and stimuli

Sound localization tests were controlled by a custom-designed MATLAB script (r2009a; MathWorks) and stimuli were generated using TDT System 3 (Tucker Davis Technology). The listener was seated in a hemi-anechoic room, in front of an array of 25 speakers (Orb Audio) mounted on a 180° arc placed at 90 cm of the listener's head on the horizontal plane, giving an azimuthal resolution of 7.5°. In a first sound localization task, the stimulus consisted of 250 ms pulsed-train (5 bursts of 25 ms) of low-pass filtered pink noise (center frequency = 2 kHz). This type of stimulus ensures that the ITDs contribute largely to the azimuthal perception (Middlebrooks & Green, 1991). In a second sound localization task, the stimulus was identical to the first one, except that its spectrum was randomized from trial to trial by roving the intensity in 1/6-octave bands by up to 40 dB. The spectral uncertainty of this stimulus was used to reduce the effectiveness of spectral cues (Kumpik et al., 2010; Wightman & Kistler, 1997). The overall level of both stimuli was 60 dB(A) at the position of the listener's head.

To allow a listener to indicate the perceived location a of stimulus, a laser pointer and a

head-tracker (Polhemus Fastrak) were attached on his head, both pointing toward a 0° in azimuth and elevation direction for a centred head position.

2.2.4 Procedure

During a sound localization run, each location was pseudorandomly presented five times, for a total of 125 trials per run. No feedback was given. At the beginning of a run, the listener was asked to seat and lean his neck on a neck rest, so that his head was centred and that the laser pointed the central speaker (0° in azimuth and elevation). This initial head position was recorded and the listener had to place his head back in this position (at less than 2 cm and 2°) before starting each trial. To start a new trial, the listener had to press the button of a stylus. If his head was correctly placed when the button was pressed, a stimulus was played from one of the 25 speakers. If the head was misplaced no sound was played and the listener was asked to place his head back to the initial position. After a stimulus was played, the listener had to direct his head (and the laser pointer) toward the speaker from which he perceived the sound originating and to press the stylus button to validate his answer. The azimuth of the pointed speaker was computed from the data given by the head-tracker.

Both sound localization measurements (fixed spectrum and random spectrum tasks) were first taken without plugs. Secondly, they were repeated with the plugs inserted and no delay added. Participants were then asked to continue wearing the plugs while engaging in daily activities, during all waking hours. The sound localization measurements were repeated the next day and repeated with a delay of 625 μ s added in the left plug. At this point of the experiment, no further modification were made to the plug's sound processing. These measurements were then repeated each day during a 10 to 12 days period. At the end of the experiment, measurements were taken with the plugs still inserted, and repeated immediately after removal of the plugs.

2.2.5 Analysis

The metric used to measure the localization accuracy of each listener was the mean signed error (MSE), a measure of the average discrepancy between listener's responses and targets

locations. Permutation tests were used to compare the MSEs between different tests of a given listener.

2.3 RESULTS

Insertion of the earplugs without delay affected the localization accuracy of three participants (results significantly different from those without plugs), even after 24 hours of wearing the plugs. Additionally, the delay added in the left earplug introduced a much smaller shift in the auditory space representation of these participants than the others. For these reasons, only the results of the remaining three participants (*P2*, *P3* and *P5*) are presented here. For those participants, no significant difference has been found between the results obtained during the task using a fixed spectrum stimulus and the task using a random spectrum stimulus. Therefore, only the results of the fixed spectrum stimulus task are detailed.

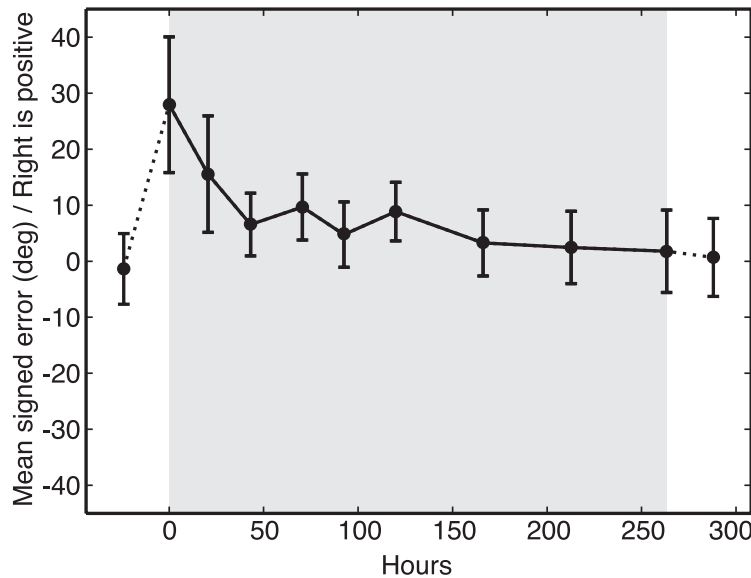


Figure 2.1 – Evolution of P3’s MSE. Hour zero corresponds to the measurement immediately after insertion of the plugs. First measurement was taken without plugs, last one, immediately after removal of the plugs.

The delay added in the left earplug shifted importantly the responses of *P2*, *P3* and *P5* in the right direction: MSE differences between the measurement without plugs and the first measurement with plugs and delay = 24.2° (*P2*), 29.3° (*P3*) and 21.4° (*P5*). After 48 hours, MSE

differences between without and with plugs decreased to: -0.95° ($P2$), 7.8° ($P3$) and 5.4° ($P5$). At the end of the experiment (10 to 12 days of wearing the plugs), the MSEs of $P2$, $P3$ and $P5$ were not significantly different to those measured before the experiment, without plugs. When measured immediately after removal of the plugs, the MSEs of those participants were still not significantly different to those measured before the experiment, therefore no aftereffect has been observed. Figure 2.1 shows the MSE evolution of $P3$ during the experiment. Figure 2.2 shows the raw data of the sound localization tests of $P3$ before, during and after wearing of the plugs.

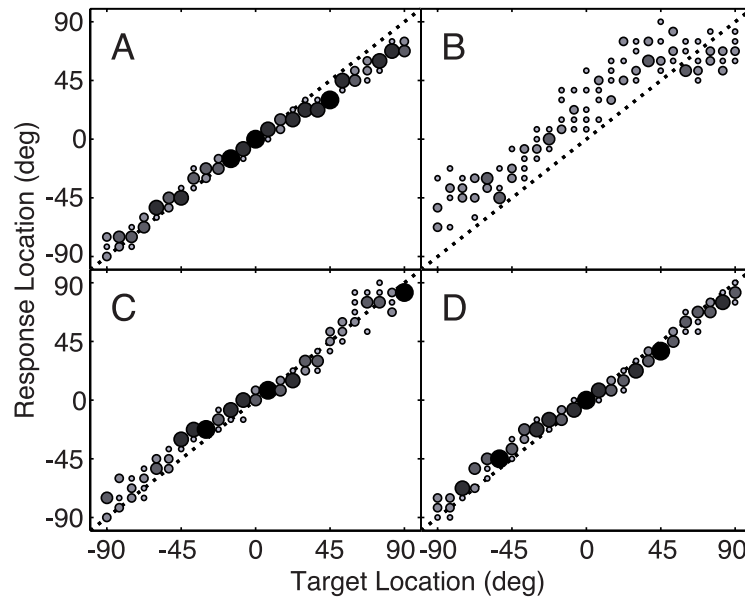


Figure 2.2 – Raw results of P3. A) without plugs. B) immediately after addition of a $625 \mu s$ delay in the left plug. C) on the last day of wearing the plugs. D) immediately after removal of the plugs. Grey value and radius of the dots increase with the number of similar responses for a specific target.

2.4 DISCUSSION AND CONCLUSIONS

The results show that the human auditory system is capable of a fast adaptation to shifted ITDs. The finding that the three participants showed no aftereffect in the opposite direction indicates that the adaptation could be attributed to a reweighting in the processing of the different spatial cues, more than to a recalibration of the ITDs. When wearing the plugs, the shifted ITDs are in contradiction with the other spatial cues and with visual feedback. A potential strategy of the auditory system could be to progressively put this biased information aside and

to start relying exclusively on the other cues. As comparable results have been measured when testing with fixed or random spectrum stimuli, the spectral cues do not seem to play a key role in the adaptation to altered ITDs. Thus, a reweighting in favour of ILDs may have been the optimal strategy for sound localization with shifted ITDs. Future experiments will aim to deepen our understanding of this adaptive process and to determine the cortical sites that are involved in these processes.

CHAPITRE 3

ARTICLE 2 : ADAPTATION TO SHIFTED INTERAURAL TIME DIFFERENCES CHANGES ENCODING OF SOUND LOCATION IN HUMAN AUDITORY CORTEX

Régis Trapeau and Marc Schönwiesner

Publié dans *NeuroImage* 118, 26-38.

Abstract

The auditory system infers the location of sound sources from the processing of different acoustic cues. These cues change during development and when assistive hearing devices are worn. Previous studies have found behavioral recalibration to modified localization cues in human adults, but very little is known about the neural correlates and mechanisms of this plasticity. We equipped participants with digital devices, worn in the ear canal that allowed us to delay sound input to one ear, and thus modify interaural time differences, a major cue for horizontal sound localization. Participants wore the digital earplugs continuously for nine days while engaged in day-to-day activities. Daily psychoacoustical testing showed rapid recalibration to the manipulation and confirmed that adults can adapt to shifted interaural time differences in their daily multisensory environment. High-resolution functional MRI scans performed before and after recalibration showed that recalibration was accompanied by changes in hemispheric lateralization of auditory cortex activity. These changes corresponded to a shift in spatial coding of sound direction comparable to the observed behavioral recalibration. Fitting the imaging results with a model of auditory spatial processing also revealed small shifts in voxel-wise spatial tuning within each hemisphere.

Keywords: auditory cortex, spatial hearing, hemifield code, plasticity, fMRI.

3.1 INTRODUCTION

Spatial hearing is an important function of the human auditory system. It guides attention (Broadbent, 1954; Scharf, 1998) and improves the detection, segregation, and recognition of sounds (Bregman, 1994; Dirks & Wilson, 1969; Roman et al., 2003). Sound directions are not represented on the auditory epithelium, but must be computed from monaural and binaural acoustic cues (Blauert, 1997; Middlebrooks & Green, 1991). The association between acoustic cues and the perceived location of a sound source depends on the shape of the listener's ears and head (Carlile, Martin, & McAnally, 2005; Langendijk & Bronkhorst, 2002; Middlebrooks, 1999). In order to conserve accurate sound localization when the shape of head and ears change during development, the auditory system must be able to modify the processing of acoustic cues or their association with sound locations (Clifton et al., 1988; Hartley & King, 2010). The general ability to recalibrate spatial hearing recalibration is also crucial to people with hearing aids or cochlear implants (Byrne & Noble, 1998; Carlile, 2014; Mendonca, 2014). These findings indicate that some natural capacity for adapting to changes in auditory spatial coding must exist.

Plasticity in spatial hearing has been demonstrated during adulthood in animals (King et al., 2011; Knudsen, 2002) and humans (Wright & Zhang, 2006). Vertical sound localization can be regained within weeks after altering spectral cues (Carlile & Blackman, 2013; Hofman et al., 1998; M. M. V. Van Wanrooij & Van Opstal, 2005). Horizontal sound localization based on interaural level differences can be regained within a few days after plugging one ear (Bauer et al., 1966; Kumpik et al., 2010). Only one study (Javer & Schwarz, 1995) manipulated interaural time differences (ITDs) and demonstrated that humans can adapt to shifted ITDs, but the mechanism remains unclear. To elucidate it, we shifted ITDs in relation to other auditory cues and vision in adult humans with programmable earplugs, and monitored participants with behavioral tests and fMRI while they regained normal sound localization.

Potential mechanisms can be derived from the known mechanisms of sound localization in the mammalian auditory system. Studies suggest that horizontal sound direction is represented by a population rate code. In the brainstem of small mammals, ITD tuning curves are wide, span

both hemifields, have maxima outside of the physiological ITD range, and are steepest around the midline (Brand et al., 2002; Lesica et al., 2010; McAlpine et al., 2001; Siveke et al., 2006). Data from cats (Middlebrooks et al., 1998; Stecker et al., 2005) and monkeys (Werner-Reiss & Groh, 2008) also support a population rate code. Studies in humans (Magezi & Krumbholz, 2010; Salminen, Tiitinen, Yrttiaho, & May, 2010; Salminen et al., 2009) also suggest a rate code of two neural populations, tuned to each hemifield (hemifield code: Salminen et al., 2012). These findings are in line with previous studies showing that human auditory cortex is involved in the processing of auditory space (Brunetti et al., 2005; Deouell, Heller, Malach, D’Esposito, & Knight, 2007; Krumbholz, Schönwiesner, Rübsamen, et al., 2005; Zatorre, Bouffard, Ahad, & Belin, 2002), and are also in line with results suggesting that a majority of auditory cortical neurons are tuned to sounds from the contralateral acoustic hemifield (Krumbholz, Schönwiesner, Cramon, et al., 2005; Palomäki et al., 2005; Pavani et al., 2002; D. L. Woods et al., 2009). Ferret studies showed that the auditory cortex also plays a role in experience-dependent recalibration of spatial hearing (Bajo et al., 2010; Nodal et al., 2010).

Based on previous findings on recalibration of spatial hearing with modified spatial cues, we hypothesized that manipulating ITDs would shift auditory spatial perception of our participants on the horizontal plane, and that this perceptual shift would progressively diminish as participants adapt. During such an adaptation, the perceived direction of sound sources must change. To quantify changes in sound direction perception, participants performed regular behavioral sound localization tests during the adaptation period. Based on the hemifield code and the previously observed preference for contralateral sounds, we hypothesized that behavioral adaptation would be accompanied by changes in the activation balance of the left and right auditory cortex. To test for such changes in hemispheric lateralization, we measured the amplitude and extent of fMRI responses in the left and right auditory cortex. We also aimed to determine whether adaptation was accompanied by detectable changes in directional tuning in the human auditory cortex. To characterize such changes in tuning, we computed high-resolution voxel-wise directional tuning curves before and after adaptation and fitted a computational model of sound direction coding to these curves.

3.2 MATERIALS AND METHODS

3.2.1 Participants

Twenty volunteers took part in the behavioral experiment after having provided informed consent. Six of them did not complete the experiment due to various reasons: three participants' earplugs stopped working during the experiment; two participants preferred to drop out of the experiment before the end; finally, during the scanning of one participant, the DSP device used for stimulus presentation blew a fuse. The 14 remaining participants (eight males, aged between 22 and 32 years) were right-handed (self-reported), had no history of hearing disorder or neurological disease, and had normal or corrected-to-normal vision. None of the participants had past or current psychiatric diseases, 7 out of 14 played an instrument (5/10 of the fMRI participants), but none was a professional musician. Participants had hearing thresholds of 15 *dB HL* or lower, for octave frequencies between 0.125 and 8 *kHz*. They had unblocked ear canals, as determined by a non-diagnostic otoscopy. Two participants experienced consistent front-back confusion when wearing the earplugs. We could not obtain sound localization measurements from these participants and thus excluded them from further analysis. Ten of the remaining twelve participants also took part in neuroimaging sessions. The experimental procedures conformed to the World Medical Association's Declaration of Helsinki and were approved by the local ethics committee.

3.2.2 Digital earplugs

We used programmable earplugs (Schönwiesner, Voix, & Pango, 2009), based on an in-ear technology platform (Sonomax, Montreal, QC, Canada). Silicon plugs were custom-modeled to each participant's ear canals, bilaterally. We measured the attenuation provided by the earplugs with a dual-microphone set-up, one inside the ear canal, one outside (Voix & Laville, 2009). We then equipped each earplug with a miniaturized microphone, transducer, and digital signal processor (Voyager, Sound Design Technologies, Burlington, ON, Canada) featuring one audio input (32 *kHz*, 20 bit ADC), one amplified audio output (16 *kHz*, 20 bit DAC) and a specialized

audio processing unit with a 2.048 MHz clock rate. The processing lag of each electronic insert was around 600 μ s. Such a delay, when applied simultaneously to both ears, is too short to create perceivable audio-visual disparities (Lewald & Guski, 2003). When turned off, the earplugs provided an attenuation between ~ 25 dB at and below 1 kHz and ~ 45 dB at 8 kHz. The overall gain, transfer function, and delay of each earplug, were controlled by a custom-made software and checked with a KEMAR manikin. The gain of the earplugs was adjusted to achieve normal loudness levels and a digital filter was used to flatten the transducer’s frequency transfer function. Figure 3.1 shows an example of the delay and attenuation that the earplugs provided.

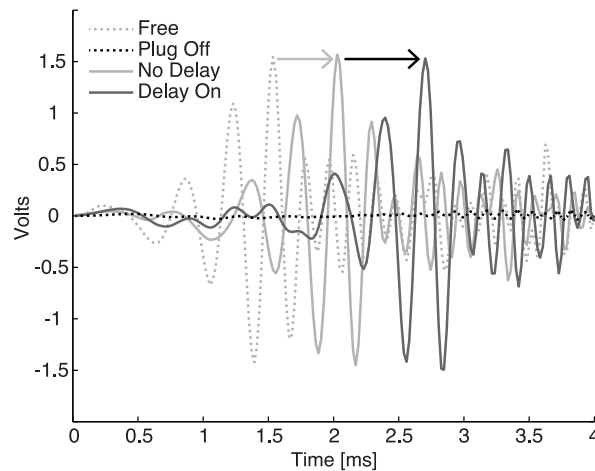


Figure 3.1 – Earplugs delay and attenuation. Example of recordings from one ear of a KEMAR mannequin of a 4 ms chirp stimulus presented from a loudspeaker placed in front of the mannequin’s head. The same stimulus was recorded when the ear was unobstructed (“free”); with an earplug fitted to one ear of the KEMAR mannequin, inserted but switched off (i.e., full attenuation, “plug off”); with the earplug switched on, but no delay added (“no delay”), the constant AD-DA delay of the DSP is visible (grey arrow); with the earplug switched on and time delay of 625 μ s added (black arrow, “delay on”).

3.2.3 Overall procedure

Participants wore the earplugs initially for two days without any modification of interaural time differences (day –2 to day 0, Figure 3.2). During this time, participants habituated to the earplugs, and we verified that the earplugs themselves (without delay) did not affect horizontal localization performance. A time delay of 625 μ s was then added to the signal processing chain of the left earplug for the next seven days (days 0 to day 7, Figure 3.2). Participants were informed that the earplugs would “slightly modify their sound perception”, but they did not know the

nature nor the effects of this modification. Participants were asked to wear the earplugs during all waking hours, but to remove them when sleeping or when they might be exposed to water. We tested participants' free-field sound localization performance each day; starting with an initial session before they were equipped with earplugs (day -2, Figure 3.2), and finishing with a session immediately after the earplugs were removed at the end of the experiment (day 7, Figure 3.2).

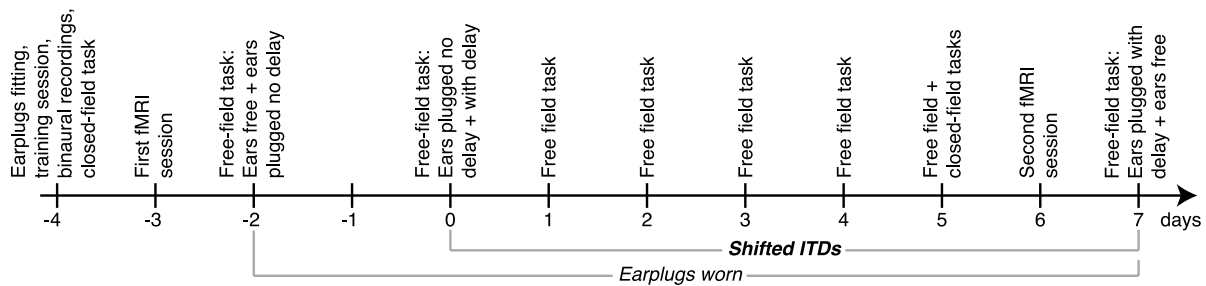


Figure 3.2 – Procedure timeline. During their first visit (day -4), participants were fitted with a pair of earplugs (not worn until day -2), completed a training session, and binaural recordings were taken. The first fMRI session always took place before the earplugs were worn (day -3). On day -2, participants performed two free-field localization tasks: one without earplugs and one with earplugs but without ITD modification. The latter was repeated on day 0, and then the delay was added to the left earplug. Participants immediately performed another free-field task with modified ITDs, and each day from day 1 to day 5. The second fMRI session took place on day 6. Day 7 was the last day with earplugs. Participants performed two free-field tasks, one with the earplugs, and immediately afterwards one without. A subset of participants also performed an additional closed-field sound localization task that determined their perceived midline with the fMRI stimuli, one day before each fMRI session (day -4 and day 5).

A subset of 10 participants underwent two identical sessions of fMRI scanning; the first session before the earplugs were worn (day -3, Figure 3.2), and the second session one day before removing the earplugs (day 6, Figure 3.2). To ensure that a shift in spatial hearing can be observed with the stimuli presented in the MRI scanner (see below), six of those participants performed an additional closed-field sound localization task that determined their perceived midline of the stimuli, one day before each fMRI session (day -4 and day 5, Figure 3.2).

To minimize procedural learning and stress during the experiment, participants completed a training session to familiarize themselves with all equipment and tasks (day -4, Figure 3.2).

3.2.4 Stimuli

The stimuli in the free-field sound localization task were 225 *ms* long trains of pulsed pink noise (5 bursts of 25 *ms*). The noise was digitally low-pass filtered with a cut-off frequency of 2 *kHz* using a 2nd order butterworth filter (12 *dB*/octave roll-off) to ensure that azimuthal sound direction is mainly conveyed by ITDs (Macpherson & Middlebrooks, 2002; Middlebrooks & Green, 1991). However, spectral cues and ILDs in the high frequencies are still usable at the resulting attenuations (−15 *dB* at 4 *kHz*, −30 *dB* at 8 *kHz*). The stimulus thus contained all localization cues to reflect participants’ auditory input in their daily environments while wearing the earplugs. The overall level of the stimuli was 60 *dB SPL* at the position of the participant’s head. In order to minimize location cues caused by differences in loudspeaker transfer functions, we used a Brüel & Kjær 1/2-inch probe microphone mounted on a robotic arm to measure these functions and designed inverse finite impulse response filters to equalize amplitude and frequency responses across loudspeakers.

In each trial during fMRI scanning, participants listened to an individual binaural recording of a 4 *s* long pulsed pink noise (same parameters as in the free-field task), presented from one of 17 different horizontal directions (−60° to +60° azimuth, 0° elevation). Because the digital earplugs were not MR-compatible, the binaural recordings were obtained from the microphones contained in the earplugs, and thus captured the participants head-related transfer functions while wearing the earplugs. Note that these recordings contained all localization cues, ILDs included. To simulate the ITD shift applied by delaying the left earplug, the left channel of the individual recording was delayed by 625 μ *s*. These binaural recordings, truncated at 225 *ms* and including the 625 μ *s* delay in the left channel, were also used in a closed-field sound localization task (see below).

Sound localization tests, binaural recordings, and stimulus presentations during fMRI sessions were controlled with custom MATLAB scripts (Mathworks, Natick, MA, USA). Stimuli and recordings were generated and processed digitally (97.7 *kHz* sampling rate, 24 bit amplitude

resolution) using TDT System 3 hardware (Tucker Davis Technologies, Alachua, FL, USA). The high sampling rate enabled us to reproduce microsecond ITDs without waveform interpolation.

3.2.5 Behavioral testing

Behavioral testing was conducted in a hemi-anechoic room. Participants were seated on a comfortable chair with a neckrest, located in the centre of a spherical array of loudspeakers (Orb Audio, New York, NY, USA) with a diameter of 1.8 m. An acoustically transparent black curtain was placed in front of the loudspeakers to avoid visual capture. The curtain also blocked the view of the surrounding space. Only a small label indicating the central location (0° in azimuth and elevation) was visible on the curtain. A laser pointer and an electromagnetic head-tracking sensor (Polhemus Fastrak, Colchester, VT, USA) were attached to a headband worn by the participant. The laser light reflected from the curtain served as visual feedback for the head position. Real-time head position and orientation were used to calculate azimuth and elevation of indicated directions.

3.2.5.1 Free-field sound localization task

Thirteen loudspeakers were used in the free-field sound localization task, covering a circle arc between -45° to $+45^\circ$ on the horizontal plane. The azimuthal angle between adjacent speakers was 7.5° . A subset of seven participants was tested with 14 additional loudspeaker, placed at elevations between -22.5° and $+22.5^\circ$, 6 on the azimuthal midline and 8 at azimuths of $\pm 22.5^\circ$ and $\pm 45^\circ$. Sound directions were presented in a pseudorandom order designed to minimize adaptation and assimilation (Cross, 1973), and were repeated five times during a session. After each sound presentation, participants indicated the perceived direction by turning their head towards the sound source and pressing a button on a hand-held button box. No feedback was given. At the beginning of a run, the participant was asked to sit and lean his neck on a neck rest, so that his head was centered in the loudspeaker arch and the head-mounted laser pointed at the central location. This initial head position was recorded and the participant had to return to this position with a tolerance of 2 cm in location and 2° in head angle before starting a trial. If the

head was correctly placed when the button was pressed, a stimulus was played from one of the tested speakers. If the head was misplaced, a brief warning pure tone was played from a speaker located above the participant’s head (Az: 0°, El: 82.5°). Horizontal localization accuracy was quantified as mean signed error (MSE), the average discrepancy between perceived and physical locations (Hartmann, 1983). Vertical localization accuracy was quantified as elevation gain (EG), the slope of the linear regression of perceived versus physical elevations (Hofman et al., 1998). A perfect localization run gives an EG of 1, while random elevation responses results in zero EG.

3.2.5.2 *Closed-field sound localization task*

A three-alternative forced-choice staircase procedure was used to measure the perceived midline of our participants. We presented individual binaural recordings from different horizontal locations (see 3.2.4. Stimuli) via Beyerdynamic DT 990 Pro open headphones (Beyerdynamic, Farmingdale, NY, USA). Participants indicated whether they perceived each stimulus left, right, or at the midline by turning the head in the respective direction. The task stopped after 10 “straight ahead” responses, the average azimuth of which served as an estimate of the perceived midline. Because the left channel of the binaural recordings was shifted by 625 μ s, initial perceived midlines were also shifted away from the physical midline.

For behavioral measures, standard error was obtained by bootstrapping the data 10000 times.

3.2.6 FMRI data acquisition

3.2.6.1 *MRI hardware and imaging*

Functional imaging was performed on a 3 Tesla MRI scanner (Magnetom TIM Trio, Siemens Healthcare, Erlangen, Germany), equipped with a 12 channel matrix head-coil (Siemens Healthcare, Erlangen, Germany), using an echo-planar imaging sequence and sparse sampling (Hall et al., 1999) to minimize the effect of scanner noise artifacts (gradient echo; repetition time (TR) = 8.4 s, acquisition time = 1 s, echo time = 36 ms; flip angle = 90°). Each functional volume comprised 13 slices with an in-plane resolution of 1.5 × 1.5 mm and a thickness of 2.5 mm

(field of view = 192 mm). The slices were oriented parallel to the average angle of left and right lateral sulci (measured on the structural scan) to fully cover the superior temporal plane in both hemispheres. As a result, the functional volumes included Heschl's gyrus, planum temporale, planum polare, and the superior temporal gyrus and sulcus. For each participant, a standard whole-brain T1-weighted structural scan (magnetization-prepared rapid gradient echo sequence) was obtained with a resolution of 1 mm³ for session one and 2 mm³ for session two.

3.2.6.2 Procedure

Because the earplugs were not MR-compatible, participants removed the earplugs just before entering the scanner room for the second session (the first session was conducted before earplugs were worn). During the fMRI measurements, the participants listened to the sound stimuli delivered through MR Confon Optime 1 headphones (MR Confon, Magdeburg, Germany), while watching a silent nature documentary. Participants were instructed to be mindful of the sound stimulation, but no task was performed in the scanner. Upon leaving the scanner room, the participants immediately reinserted the earplugs. Each session was divided into two runs of 152 acquisitions. Each acquisition was directly preceded by either a 4 s long stimulus (positioned in the TR of 8.4 s so that the maximum response is measured by the subsequent volume acquisition) or silence. Each of the 17 sound directions was presented 16 times, interleaved with 32 silent trials in total. Note that the stimuli were ITD-shifted by 625 μ s in both sessions. Spatial stimuli and silent trials were presented in a pseudorandom order, so that the angular difference between any two subsequent stimuli was greater than 37.5° and silent trials were not presented in direct succession. Stimulus sequences in both sessions were identical.

3.2.7 FMRI data analysis

The functional data were corrected for head motion by realigning each volume to the third volume of each run and were spatially smoothed with a 3 mm full-width half-maximum 3-dimensional Gaussian kernel. Each run was then linearly coregistered to the structural scan

of the first session. Preprocessing was done using the MINC software package (McConnel Brain Imaging Center, Montreal Neurological Institute, Montreal, Canada).

The analysis of the functional data was performed using fMRISTAT (Worsley et al., 2002) and custom MATLAB scripts. We used general linear models to estimate voxel-wise effects and t-values for contrasts of every sound direction versus the silent baseline. To increase signal-to-noise ratio, we combined any three adjacent directions into one condition, which resulted in 48 repetitions per condition. The combination of all sound directions formed a sound versus silence condition. T-maps were thresholded at $p < 0.05$, corrected for multiple comparisons using Gaussian Random Field Theory (Worsley et al., 1996). To assess the lateralization of the fMRI activation, we computed the normalized (across participants) number of significantly active voxels and the normalized activation strength (in percentage BOLD signal change), in each hemisphere for each condition. We calculated these measures independently for each hemisphere and computed lateralization ratios (left hemisphere / right hemisphere). Statistical analyses on these ratios were achieved using a two-way analysis of variance for repeated measures. For this analysis, we compensated for the partial dependence created by combining adjacent directions into one condition, by reducing the effective degrees of freedom for the factor “direction” from 15 to a conservative value of 7 (8 effectively independent directions – 1).

We computed voxel-wise directional tuning curves by plotting the response size in each condition for voxels that showed consistent tuning. We calculated the Pearson correlation between tuning curves extracted independently from the first and the second run of each session for voxels that responded significantly to at least one sound condition and then discarded voxels with correlation with a p-values higher than 0.1 and negative r-values. From these tuning curves we determined voxel-wise direction preference by two methods: a) by finding the direction that evoked the maximal response in a given voxel (MAX), and b) by computing the centre-of-gravity (COG) of the voxel tuning curve. We plotted the distributions of tuning preference from each of these methods and for each participant, hemisphere, and session. These distributions were normalized and averaged across participants.

For all neuroimaging data averaged across participants, standard error was obtained by bootstrapping the data 10000 times, using each participant's mean response.

3.2.8 Modeling

To aid the interpretation of the fMRI results, we constructed a model of auditory spatial processing expressed in terms of single-neuron spatial tuning curves. Depending on the values of its parameters, this model could either reflect the place code theory, the hemifield code theory, or a mix of both. We performed the same computations on the model data as on the fMRI data and used a genetic algorithm to find the parameters' values that minimized the difference between the model and the fMRI data. All parameters were under genetic algorithm control and their ranges were fixed as described below.

3.2.8.1 Model description

The model and parameter ranges were inspired by the models proposed by [Salminen et al. \(2009\)](#). Our model consisted of 10000 sound sensitive neurons distributed in both hemispheres. These neurons could be of 3 types: noise neurons, hemifield code neurons, and place code neurons. Noise neurons had random tuning curves with random overall amplitudes ranging between 0 and 1. The tuning curves of the hemifield code and place code neurons were modeled as Gaussians with peak amplitude of 1, and with means and standard deviations that depended on several parameters: the mean preferred azimuth of the hemifield code populations (*hemifield population center*, ranging from 45° to 135°); the jitter of the preferred azimuth across neurons (*hemifield population span*, from 3° to 40°); the standard deviation of the Gaussians that characterized the hemifield population tuning curves (*hemifield population σ* , from 32° to 96°); and the non-uniformity of the tuning width of place code neurons (*place code non-uniformity*, from 0 [uniform σ of 22.5°] to 1 [σ increased linearly from 1.1° at 0° azimuth to 87° at 180° azimuth]). Figure 3.3 illustrates how these parameters determine the shape and distribution of the tuning curves.

The proportion of noise neurons was set by the *noise ratio* parameter. The proportion

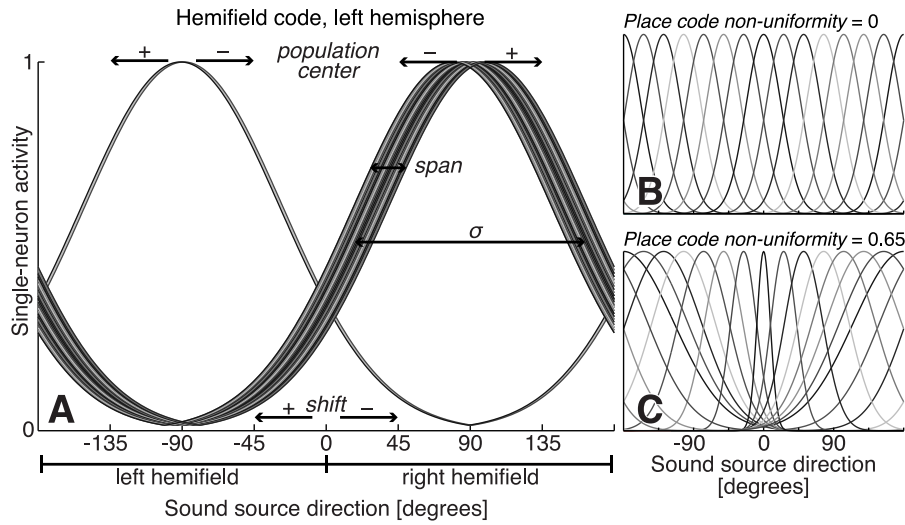


Figure 3.3 – Illustration of the model parameters. A) Tuning curves of hemifield code neurons in the left hemisphere. Each curve shows the level of one model neuron as a function of sound direction. In this example, the *hemifield population contralaterality* was set to 90 %. Consequently, most of the neurons are tuned to the right hemifield. The *hemifield population center* controls the direction at which the tuning curves are centered. Note that this value is always opposite for left-tuned and right-tuned populations, making their location symmetrical with respect to the midline. The *hemifield population span* is the jitter of the preferred azimuth across neurons. The *hemifield population σ* is the standard deviation of the Gaussians that characterize the hemifield population tuning curves. The *shift* parameter shifts the azimuthal stimulus locations at the input of the model and affects all types of neurons. B) Tuning curves of place code neurons with a *place code non-uniformity* set to 0 (uniform σ of 22.5°). C) Tuning curves of place code neurons with a *place code non-uniformity* set to 0.65 (σ increases linearly from 9° at 0° azimuth to 64° at a 180° azimuth).

of hemifield code neurons among the non-noise neurons was set by the *hemifield code ratio* parameter. Additional parameters were the proportion of contralaterally tuned neurons in the hemifield population (*hemifield population contralaterality*), and a parameter that would shift the azimuthal stimulus locations at the input of the model (*shift*).

3.2.8.2 Genetic algorithm

To find the parameter set that minimizes the differences between the model output and the observed fMRI data, a genetic algorithm was run for each session separately, on a population of 1000 candidate solutions (i.e., sets of parameters). Each candidate solution was created by randomly setting parameters to values within the ranges given above. For each generation, a set of solutions was selected by a mix of elitist strategy (Deb, 2001) and fitness proportionate

method (Holland, 1975) to create a new population of candidate solutions using a uniform crossover method (Syswerda, 1989). The probability of a mutation to occur during the creation of a new generation was set to 1 %.

The fitness of candidate solutions was evaluated by comparing the resulting model output with the observed fMRI data. For each evaluation, the modeled neurons were randomly distributed between 500 virtual voxels (10 to 30 neurons per voxel). The proportion of voxels in each virtual hemisphere was set to the average proportion of significantly active voxels in each hemisphere in the observed data. Normalized distributions of preferred directions were then computed from these virtual voxels in the same way as from the fMRI data and compared to the observed data. The fitness of a solution was computed as the mean likelihood of the model's preferred direction distribution (COG and MAX, see 3.2.7. FMRI data analysis) given the bootstrap distribution (10000 repetitions) of the observed preferred direction at each spatial condition. Due to the noise neurons, and because neurons were randomly distributed across virtual voxels, the fitness score varied slightly between evaluations of the same candidate solution. We therefore averaged fitness scores across four evaluations of a solution. The genetic algorithm terminated when the mean fitness of the current population of solutions reached an asymptote. To minimize the probability of converging in local maxima, we repeated this process 50 times with different initial parameters.

3.3 RESULTS

3.3.1 Behavioral results

We used digital earplugs to modify ITD for several days. The majority of participants adapted within 1 or 2 days to the modified ITDs. Participants showed no aftereffect upon removal of the earplugs, and retained the capacity to localize sounds with shifted ITDs for several days after earplug removal. The average time courses of the MSE and EG across participants are shown in Figure 3.4.

The first free-field sound localization test, without earplugs, showed that the participants had

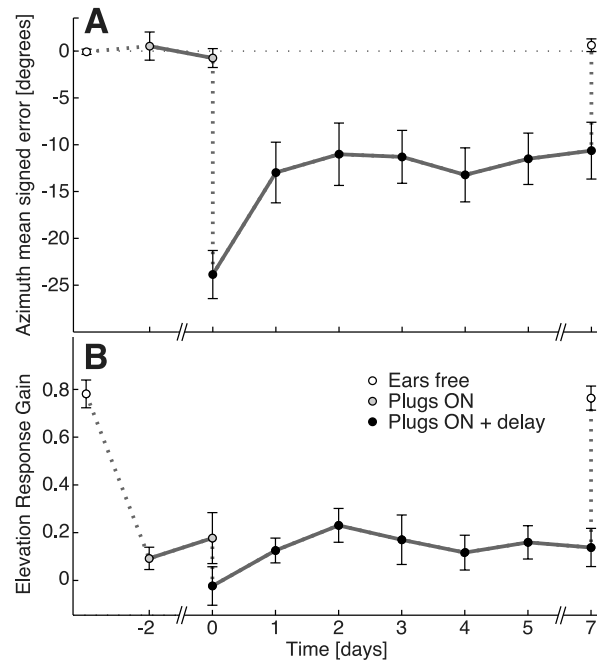


Figure 3.4 – Time course of sound localization performance. A) Time course of the mean signed error on the horizontal plane ($n = 12$). Participants errors were close to zero when tested without earplugs. The insertion of the earplugs without ITD shift did not shift their responses. The addition of the $625 \mu s$ delay in the left earplug shifted perceived sound directions by $\sim 25^\circ$ to the right. This shift diminished quickly to about 10° during the first two days, before plateauing for the next five days. After earplug removal, the average accuracy returned immediately back to baseline. B) Time course of the elevation gain ($n = 7$). Participants had high elevation gains without earplugs (~ 0.8). The insertion of the earplugs largely disrupted their elevation perception, decreasing the elevation gain almost to 0. Although participants showed a slight recalibration on the vertical plane, the elevation gain remained below 0.2. After earplug removal, the average elevation gain returned immediately back to baseline. White circles indicate performance without earplugs; grey circles indicate performance with earplugs but without ITD shift; black circles indicate performance with earplugs and shifted ITDs. Error bars indicate standard errors obtained by bootstrapping.

typical localization accuracy with horizontal localization errors close to zero ($MSE = -0.07^\circ \pm 0.28$ [mean $\pm$ SE], Figure 3.5.A), and high elevation gains ($EG = 0.78 \pm 0.06$, Figure 3.6.A). When the participants wore the earplugs without additional delay, the horizontal localization accuracy remained unshifted ($MSE = -0.74^\circ \pm 0.99$, Figure 3.4.A), but the azimuthal variance in their responses increased significantly (mean standard deviation without earplugs 3.78° , against 8.52° with earplugs; one-tailed permutation test: $p < 0.01$). With the earplugs, participants initially experienced an almost complete loss of elevation perception ($EG = 0.09 \pm 0.04$, Figure 3.4.B).

The addition of the delay in the left earplug of each participant shifted the perceived sound directions to the right ($MSE = -23.86^\circ \pm 2.69$, Figure 3.5.B) and tended to affect the

elevation gain (from $EG = 0.18 \pm 0.1$ to $EG = -0.02 \pm 0.08$, Figures 3.4.B and 3.6.B). After only 24 h, the perceptual horizontal shift of sound directions had diminished to about half of its initial value ($MSE = -12.97^\circ \pm 3.32$, Figure 3.4.A). After 48 h, the average localization accuracy across participants reaches a stable level of approximately -10° in MSE and 0.2 in EG. Nonetheless, we observed individual differences, especially on the horizontal plane. The majority of the participants (8 of 12) regained close-to-normal horizontal localization performance ($MSE = -3.87^\circ \pm 0.40$) during the last day of testing. The remaining four participants did not show a consistent improvement ($MSE = -22.39^\circ \pm 4.17$, on the last test with earplugs).

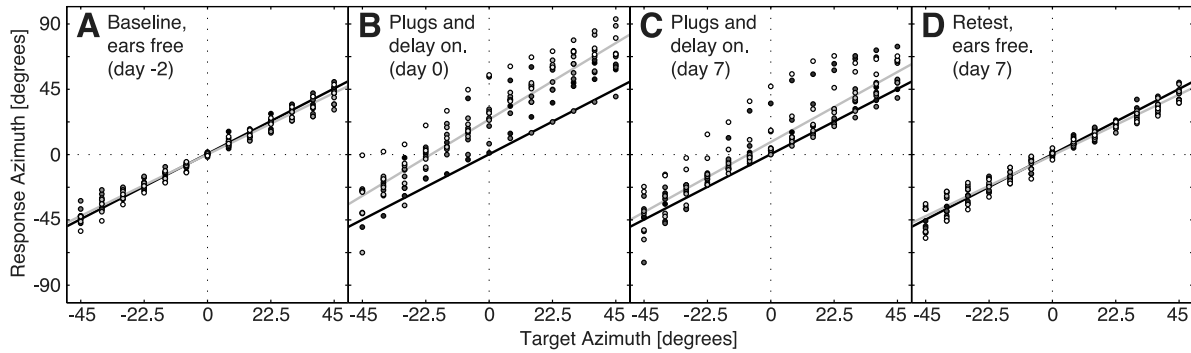


Figure 3.5 – Effect of the digital earplugs on horizontal sound localization performance. Each dot represents the average response of a participant at a specific target location. Grey values differentiate participants ($n = 12$). Perfect responses would fall on the back lines. A linear fit of the actual responses is shown by the grey lines. Positive angles are in the right hemifield. A) Before inserting the earplugs, all participants showed close-to-perfect horizontal sound localization accuracy. B) The insertion of the earplugs with delay in the left channel shifted the responses of the participants to the right (upwards in the plot). C) This perceptual shift decreased significantly after seven days with the plugs on. D) Performance was indistinguishable from baseline immediately after removal of the earplugs.

After the last localization test with the earplugs (Figures 3.5.C and 3.6.C), the participants removed the earplugs and immediately repeated the test. Localization accuracy in this final test was indistinguishable from baseline accuracy, i.e., there was no measurable aftereffect ($MSE = -0.62^\circ \pm 0.67$; $EG = 0.76 \pm 0.05$, Figures 3.4.A,B, 3.5.D and 3.6.D).

Five participants accepted to return three days after having removed the earplugs and to wear them again for a single localization test. These participants were still able to localize sounds with shifted ITDs ($MSE = -3.34^\circ \pm 1.39$, similar to $-3.92^\circ \pm 0.62$ on the last day with earplugs).

In order to test whether the recalibration increased the reliance on ILDs, we asked three

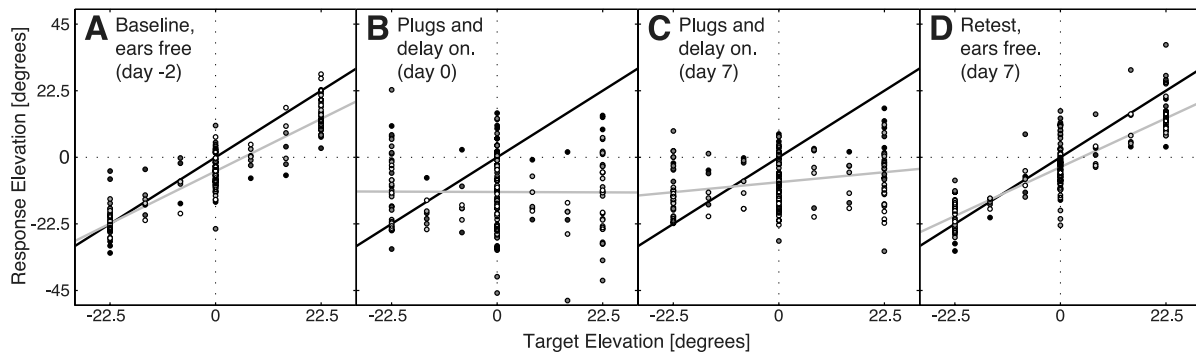


Figure 3.6 – Effect of the digital earplugs on vertical sound localization performance. Each dot represents the average response of a participant at a specific target location. Grey values differentiate participants ($n = 7$). Perfect responses would fall on the back lines. A linear fit of the actual responses is shown by the grey lines. Positive values are upward locations. A) Without earplugs, all participants had good elevation perception. B) The insertion of the earplugs largely disrupted the elevation perception of all participants (no increase of response elevation with target location). C) Participants regained some elevation perception after seven days with the plugs on, but were still far from baseline performance. D) Performance immediately after removal of the earplugs was similar to baseline.

participants who showed a large recalibration (from $\text{MSE} = -25.75^\circ \pm 2.63$ to $-2.98^\circ \pm 1.31$ during week-long recalibration period) to repeat the free-field localization test with stimuli low-pass filtered at 500 Hz to reduce the perceptual weight of ILDs. At our loudspeaker distance only minor head-shadowing occurs for frequencies below 500 Hz (Akeroyd, 2006; Kuhn, 1987). They localized these stimuli equally well ($\text{MSE} = -1.57^\circ \pm 1.61$) as the original stimuli ($\text{MSE} = -2.98^\circ \pm 1.31$).

The earplugs modified not only ITDs, but also spectral cues (head-related transfer functions, HRTFs) necessary for sound elevation perception. We investigated the relationship between the rate of recalibration of horizontal sound localization, measured as the change in azimuthal MSE during the first 48 h with shifted ITDs (day 0 to day 2), and several measures of elevation perception. We found a positive correlation between the horizontal recalibration rate and the EG obtained from the first test with earplugs but unshifted ITDs (Pearson $r = 0.77$, $n = 7$, $p = 0.04$, Figure 3.7.A). The correlation was due to differences in the amount of initial EG loss when inserting the earplugs, not due to preexisting vertical localization ability (correlation between the horizontal recalibration rate and EG without earplugs: $r = 0.003$, $n = 7$, $p = 0.99$). We also found a strong positive correlation between horizontal and vertical recalibration rates (Pearson

$r = 0.92$, $n = 7$, $p = 0.003$, Figure 3.7.B). We measured the rate of recalibration of vertical sound localization as change in EG during the first 48 h with earplugs and unshifted ITDs (day -2 to day 0).

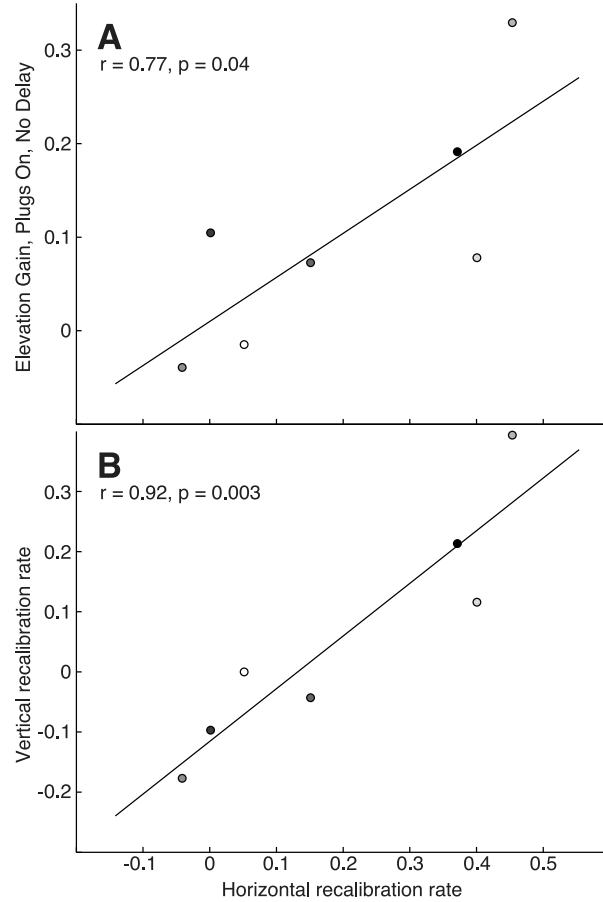


Figure 3.7 – Correlation between azimuth and elevation performance. Each dot represents a participant ($n = 7$) and grey values differentiate them. In both panels the x-axis corresponds to a measure of the rate of recalibration of horizontal sound localization (change in azimuthal MSE between days 0 and 2). In panel A, this variable is plotted against the initial amount of residual elevation perception with earplugs (EG obtained from the first test with earplugs but without delay). In panel B, it is plotted against the rate of recalibration of vertical sound localization (change in EG between day -2 and day 0).

3.3.2 Neuroimaging Results

We used high-resolution fMRI to investigate the neural correlates of adaptation to shifted ITDs. We found changes in the distribution of voxel-wise direction preferences and in the hemispheric lateralization of fMRI responses to spatial sounds during recalibration.

3.3.2.1 Directional tuning curves

We computed voxel-wise directional tuning curves. In each hemisphere, most of the tuning curves had similar profiles, with response maxima in the contralateral hemifield and large slopes crossing both hemifields (Figure 3.8). We did not observe noticeable differences between the average tuning curves of sessions one and two. In both sessions, average tuning curves in each hemisphere intersected close to, but consistently to the left of, the midline. The average percentage BOLD signal change across voxels and participants was 2.75 % at the maximum of the tuning curves and 2.17 % at the minimum. The effect of sound direction corresponded thus to about a 0.6 % BOLD signal change.

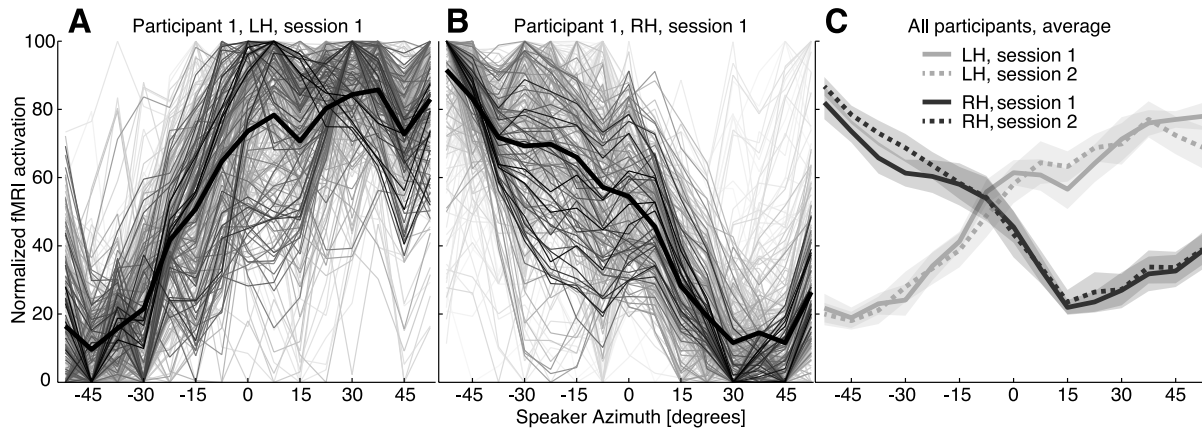


Figure 3.8 – Directional tuning curves. Directional tuning curves of all active voxels in the first fMRI session in the left (A) and right (B) hemisphere of one participant. Each curve represents the response magnitude recorded for one voxel for the different directions, normalized between 0 and 100. The darkness of the curves is proportional to the average responses before normalization. Mean tuning curves are given by the bold black lines. Most tuning curves had similar shapes, with maximal responses to sounds in the contralateral hemifield. C) Average tuning curves across all participants, for each hemisphere and each session, normalized across participants. Identical ITD-shifted recordings were presented in both fMRI sessions. Average tuning curves were consistent across participants. There was no noticeable difference between sessions. Grey lines represent data from the left hemisphere, black lines are from the right hemisphere. Solid lines indicate session one, dashed lines session two. Shaded areas represent the standard error obtained by bootstrapping the data across participants.

We then reduced voxel tuning curves to two measures of direction preference, the center of gravity and the location of the maximal response of the tuning curve. In both hemispheres and both sessions, the distribution of the COGs across participants was centered about 15° from the midline in the contralateral hemifield (Figure 3.9.A,B). In the right hemisphere, the height

of the distribution peak increased from session one to session two. A large proportion of the voxels responded the strongest to stimuli from the most lateral directions of the contralateral hemifield, other directions being much less represented (Figure 3.9.C,D).

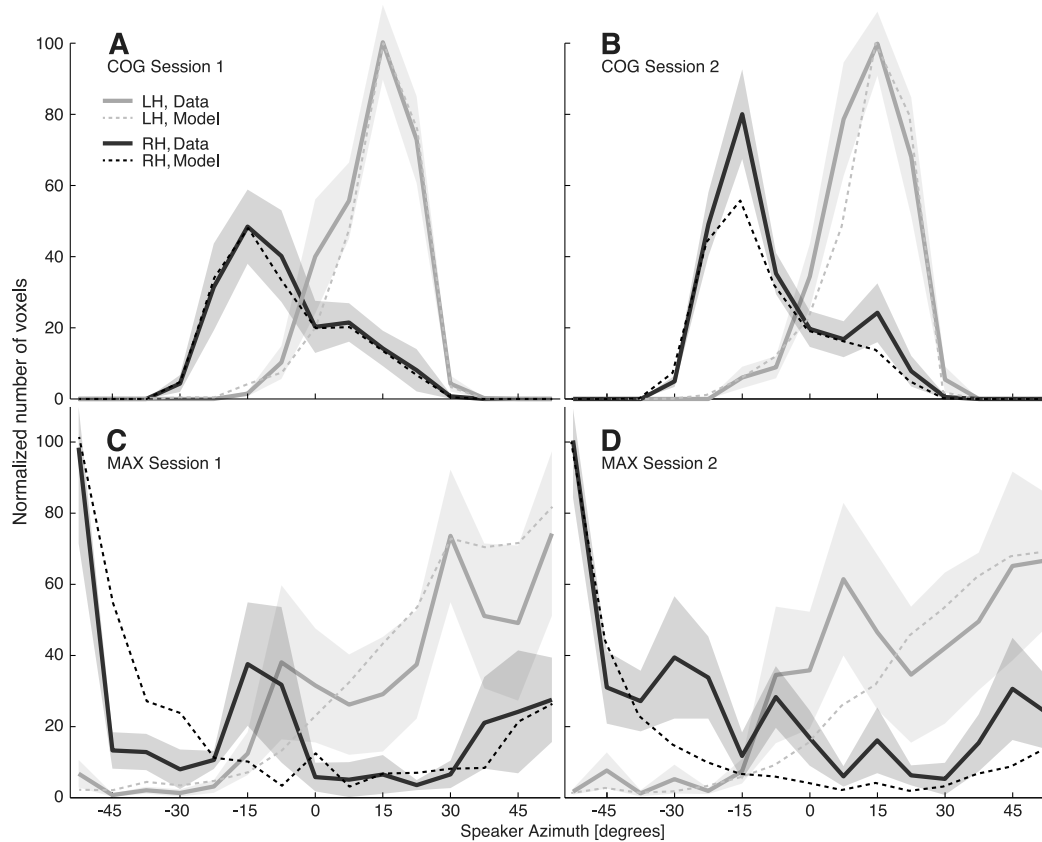


Figure 3.9 – Observed and modeled distributions of the center of gravity and maximum of directional tuning curves. Observed distributions are shown in solid lines, modeled distributions in thin dashed lines. Observed data were normalized across participants. A and B) The distribution of the tuning curves COGs in both hemispheres (right in black, left in grey) peaked about 15° from the midline in the contralateral hemifield for the first (A) and the second session (B). C and D) Most voxels responded maximally to the most lateral directions of the contralateral hemifield (line colors identical to A) in the first (C) and the second session (D). In session one, the contralateral bias was larger in the left (~ 90 %) than in the right hemisphere (~ 69 %). From session one to session two, contralaterality did not change in the left hemisphere, but increased to 75 % in the right hemisphere. Shaded areas represent the standard error obtained by bootstrapping the data. Dashed lines show the same measures, but calculated from our model with the parameters found by the genetic algorithm. The distribution of the centre-of-gravity and the maxima of modeled tuning curves closely approximate the observed curves.

In session one, nearly all voxels in the left hemisphere preferred contralateral stimuli, as their tuning curves had COGs and MAXs in the right hemifield (COG: 96 %; MAX: 85 %). In the right hemisphere, a slight contralateral bias was found in session one (COG: 73 %; MAX: 66 %).

In session two, these proportions remain stable in the left hemisphere and rose to 77 % and 72 %, for COG and MAX respectively, in the right hemisphere. Figure 3.10 shows the spatial distribution of COG across auditory cortex for each participant and the group average.

3.3.2.2 *Lateralization*

We examined the ratio of the number of voxels in the left and right hemispheres that responded significantly to each sound direction (Figure 3.11). This ratio increased for more rightward directions up to $\sim 30^\circ$ in both sessions. Results of a two-way analysis of variance for repeated measures performed on the ratios showed a significant main effect for sound direction ($F_{7,126} = 18.41$, $p < 0.0001$), and no interaction between direction and session ($F_{7,126} = 1.20$, $p = 0.318$). The same analysis performed on the ratio of the activation strength (in percentage BOLD signal change) of showed similar results, a significant effect for direction ($F_{7,126} = 5.04$, $p < 0.0001$), and no interaction between direction and session ($F < 1$).

In session one, we observed a large asymmetry of the distribution of the significant voxels, with a larger extent of activity in the left than the right hemisphere across all conditions (ratio > 1 for all directions). Smaller ratios were observed in the second session, with stimuli from the leftmost directions evoking stronger responses in the right than the left hemisphere (ratio < 1). The analysis of variance revealed an effect for the factor session in the ratio of the number of significant voxels ($F_{1,126} = 5.12$, $p = 0.049$), but not in the ratio of their mean responses ($F < 1$). This reflects the fact that during the first session most of our stimuli ($\sim 80\%$) were perceived in the right hemifield (mean perceived midline for the fMRI stimuli of $-45.58^\circ \pm 3.42$, for six participants). The day before the second fMRI session, the mean perceived midline had moved toward the midline ($-29.45^\circ \pm 3.88$), reducing the proportion of stimuli perceived in the right hemifield to $\sim 70\%$. An analysis of each hemisphere independently, revealed an effect for the factor session in the number of significant voxels in the right hemisphere ($F_{1,126} = 5.37$, $p = 0.045$), but not in the left hemisphere ($F < 1$). For both hemispheres, no effect of the factor session was found in the mean response of the significant voxels.

We modeled the distribution of ratios of the number of significantly active voxels in session

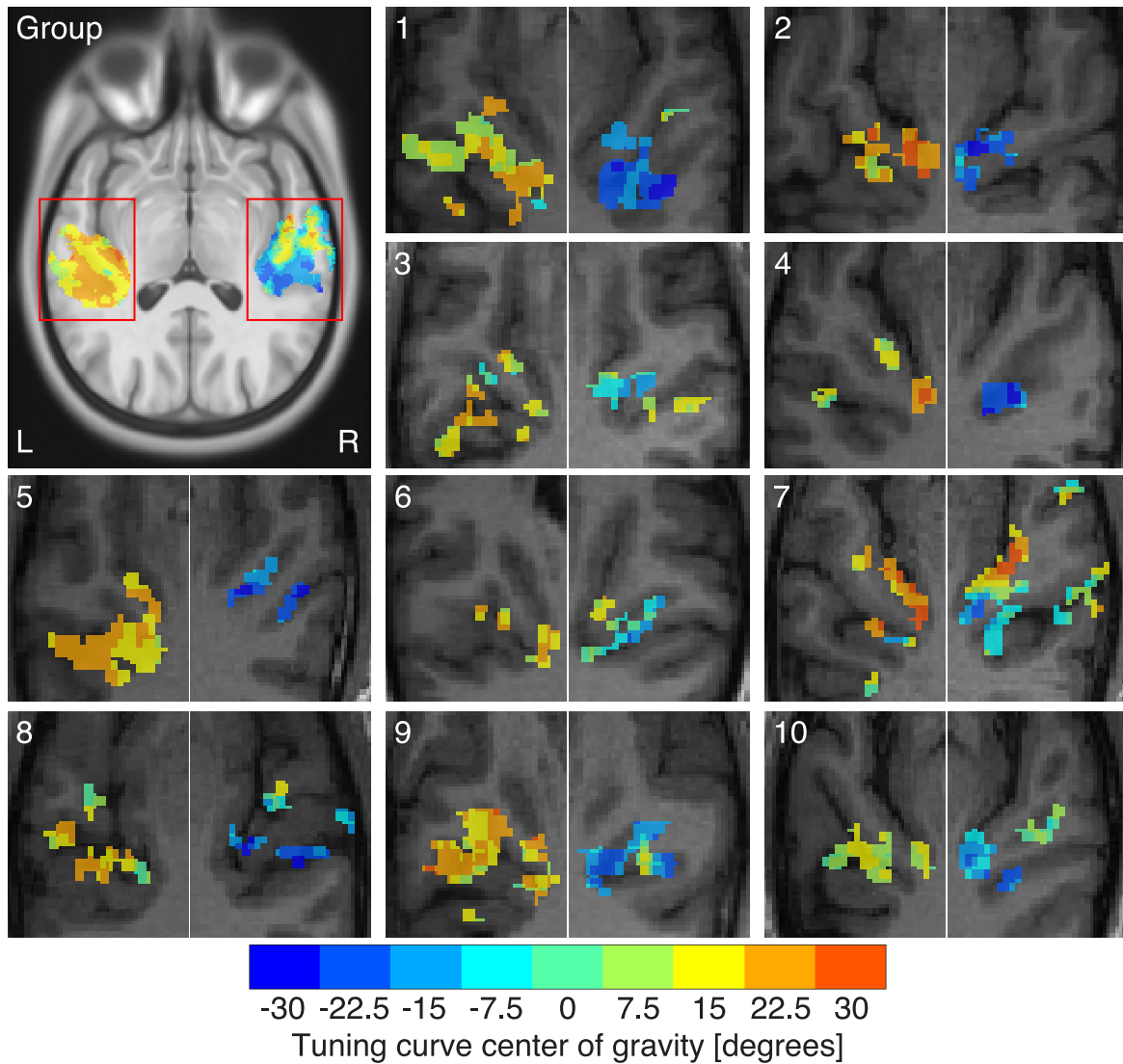


Figure 3.10 – Distribution of tuning curve center of gravity across auditory cortex. Results of both fMRI sessions are combined. Center of gravity of tuning curves in significantly activated voxels are color coded and superimposed on an axial section of each participant’s T1-weighted scan showing the left and right temporal lobes. Each slice was tilted 30° forward so that it is parallel to the Sylvian fissure. The fixed-effects group average is superimposed on a section of the ICBM-152 brain average. Warm colors represent centers of gravity located in the right hemifield, cold colors represent centers of gravity located in the left hemifield. The plots show the location and extent of the activity evoked by spatial sounds in individual brains. The contralaterality of the tuning curve center of gravity (most voxels in each hemisphere preferred sounds in the contralateral hemifield) is clearly visible. However, in 6 participants (3, 6, 7—10), the right hemisphere also preferred sounds from the ipsilateral hemifield, presumably because sounds were perceived mostly in the right hemifield before adaptation.

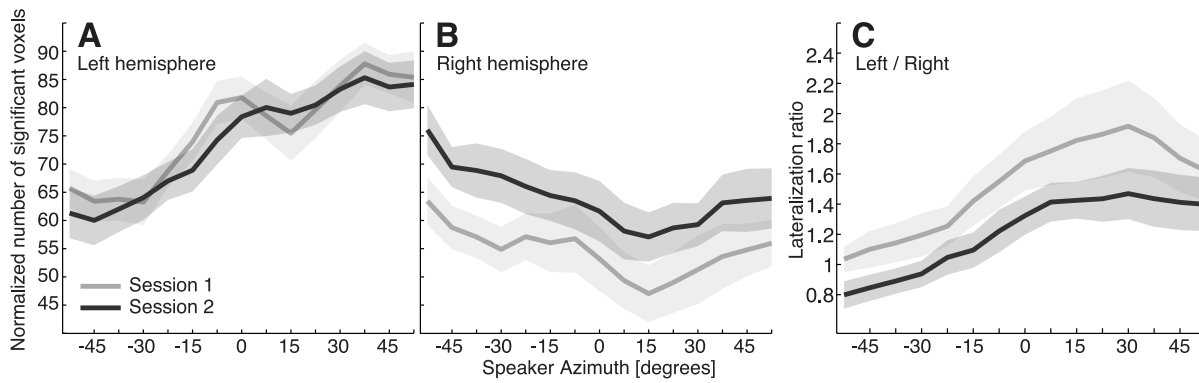


Figure 3.11 – Sound vs. silence contrasts in both hemispheres and lateralization ratio. A) The number of significant voxels in left hemisphere (normalized across participants) depended on azimuth, with rightward directions eliciting a more widespread response. There were no differences between sessions. B) The number of significant voxels in right hemisphere also depended on azimuth, but with leftward directions eliciting a more widespread response and an increase across directions from session one to two. C) We divided for each direction the number of significant voxels in the left hemisphere by that in the right hemisphere, resulting in a lateralization ratio. The lateralization ratio dropped significantly from session one to two for all directions. Results from the first fMRI session are shown in grey, and results from the second session in black. Shaded areas represent the standard error obtained by bootstrapping.

one with a second order gaussian ($R^2 = 0.99$) to predict the perceived directions from ratios obtained in session two. The predictions were consistently shifted leftward, in the same direction as the perceptual recalibration. The average shift between predictions and the actual directions presented was 35.7° across all directions, and 26.5° when considering only directions at which lateralization ratios varied with direction.

3.3.3 Modeling

The the genetic algorithm found model parameters that resulted in an excellent fit between model output and observed data (see Figure 3.9). In both sessions, the modeled centre of gravity and maximal response distributions that were less than one standard error away from the observed data. When fitting the model to the first session, the mean parameters found after 50 repetitions of the genetic algorithm were: a ratio of neurons responding randomly to sound locations (*noise ratio*) of 77 %; among the non-noise neurons, 64 % responding in a hemifield code fashion (*hemifield code ratio*); these neurons were centered around 94° in each hemisphere (*hemifield population center*) with a spread of 13° (*hemifield population span*) and the Gaussians

characterizing these neurons had a σ of 74° (*hemifield population σ*); 92 % of the hemifield code neurons were contralaterally tuned (*hemifield population contralaterality*); the σ non-uniformity of the place code neurons was of 0.42 (*place code non-uniformity*); the horizontal *shift* of the presented directions was 59° to the right. The parameters *hemifield population span* and *place code non-uniformity* had little influence on the quality of the fit (they varied randomly between solutions and fitness). The solution found for session two was similar to that for session one, except for two parameters: *hemifield population center* moved closer to the midline (from 94° to 89°), and *shift* moved also closer to the midline (from 59° to 52°). Manipulating the *shift* parameter helped us to interpret the pattern of COG distributions. Large changes in the *shift* parameter result in minor, symmetrical, changes of COG peak locations. This may explain why observed COG distributions were symmetrical around 0° , even though the stimuli were perceived as shifted to the right. The decrease of the *shift* value from session one to session two, was in same direction as the behavioral recalibration and would partly explain differences found in COG distributions between sessions. Decreasing *shift* parameter increases the peak of the modeled COG distribution in the right hemisphere, just as observed in the data. Figure 3.12 shows an example of the modeled tuning curves obtained with the parameters found for session one, as well as the average tuning curves of 100 solutions (50 from each session; see 3.2.8.2. Genetic algorithm). The modeled curves closely resemble the observed ones (compare with Figure 3.8).

3.4 DISCUSSION

3.4.1 Behavioral recalibration and implications of missing aftereffect

We showed that adult humans can relearn to locate sounds with shifted ITDs. Our behavioral results confirm and extend those of Javer and Schwarz (1995). We observed faster relearning of shifted ITDs compared to relearning of spectral cues (~ 4 weeks: Hofman et al., 1998) or adaptation to monaural earplugging (> 3 days: Bauer et al., 1966; no adaptation during several days: Florentine, 1976; ~ 1 week: Kumpik et al., 2010). All of our participants who regained close-to-normal horizontal sound localization, did so within 48 hours. The difference in adaptation

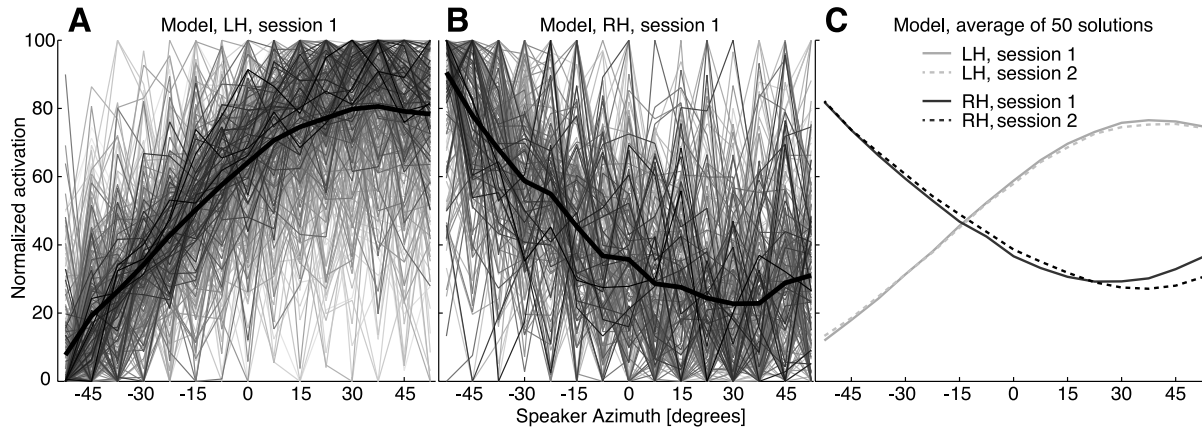


Figure 3.12 – Modeled directional tuning curves. Example of voxelwise directional tuning curves created by our model, in left (A) and right (B) virtual hemisphere. In this example, the mean parameters found after 50 repetitions of the genetic algorithm on session one were used. Each curve represents the response magnitude modeled for one virtual voxel for the different directions, normalized between 0 and 100. The darkness of the curves is proportional to the average responses before normalization. Mean tuning curves are given by the bold black lines. C) Average tuning curves of 50 different solutions (i.e., 50 different sets of parameters) found by the genetic algorithm, for each hemisphere and each session, normalized across solutions. Grey lines represent data from the left hemisphere, black lines are from the right hemisphere. Solid lines indicate session one, dashed lines session two.

speed for shifted ITDs compared to other sound localization cues disruptions, may indicate a different physiological mechanism of adaptation. Our participants adapted without explicit training by coping with the modified cue in their daily environment. Training may further increase adaptation speed (Bauer et al., 1966; Butler, 1987; Carlile et al., 2014; Kumpik et al., 2010; McPartland et al., 1997).

Participants did not mislocalize sounds in the opposite direction upon removal of the earplugs. Absent aftereffects (ITD: Javer & Schwarz, 1995; HRTF: Hofman et al., 1998; M. M. V. Van Wan-rooij & Van Opstal, 2005) or very small aftereffects (ILD: Kumpik et al., 2010) have also been reported when altering other auditory spatial cues in humans. A potential mechanism of adapting to an altered cue is to rely on other, intact, cues. Kumpik et al. (2010) showed that the adaptation to unilateral earplugging resulted in a reweighting toward monaural spectral cues provided by the unaffected ear. In our study, monaural spectral cues are unlikely to have contributed to spatial hearing, because spectral cues were altered in both ears by the earplugs, which protruded into the concha. Elevation perception was markedly reduced in our participants, even after recalibration, indicating that only a limited amount of relearning of the spectral cues

took place, and that these cues were not informative during horizontal recalibration. Moreover, we found previously that randomization of stimulus spectra (Wightman & Kistler, 1997) did not affect localization performance for our stimuli and apparatus (Trapeau & Schönwiesner, 2011). Our participants had access to veridical ILD cues when wearing the earplugs in their daily environments. During lab testing, ILD cues were also available, albeit strongly attenuated, and might thus have contributed to sound localization. To determine whether adaptive reweighting explained our results, we tested a subset of participants, who showed large recalibration, with 500 Hz low-pass filtered stimuli, in which ILD and spectral cues are minimal at our loudspeaker distance. The low-pass filter did not affect localization performance. Thus, adaptive reweighting towards ILDs cannot fully explain the observed recalibration.

The lack of an aftereffect does not imply that the ability to localize sounds with shifted ITDs disappears immediately upon removal of the plugs. Rather, our findings indicate that it persists through several days without the manipulation, and that both mappings (original and shifted ITDs) were available to the participants for at least three days after the exposure to shifted ITDs. We thus argue that some remapping of the correspondence between the localization cues and sound directions took place, even though it did not override the pre-existing mapping. Hofman et al. (1998) reported a similar lack of an aftereffect for modified HRTFs and speculated that the mechanism may be similar to the acquisition of a second language (which also interferes little with the representation of the native language).

3.4.2 Correlations between horizontal and vertical localization performance

We observed large inter-individual differences from full to little horizontal recalibration. Others have also observed large differences in sound localization relearning and training (Javer & Schwarz, 1995; Jeffress & Taylor, 1961; Shinn-Cunningham, Durlach, & Held, 1998; Wright & Fitzgerald, 2001). Part of our variability was explained by the reduction in HRTFs availability when the earplugs were first inserted, but not by pre-existing vertical localization accuracy (without earplugs). Spectral cues contribute little to horizontal sound localization for low-frequency stimuli, but recalibration in a real-world environment may still depend on the their

availability. HRTFs contribute to some extent to the clarity of an azimuthal percept (Flannery & Butler, 1981) and knowledge of the horizontal sound location helps monaural listeners to determine elevation (Butler, Humanski, & Musicant, 1990). Horizontal and vertical sound location and movement activate common cortical substrates (Pavani et al., 2002). In our data, the increased uncertainty in horizontal sound localization might be partly due to loss of elevation perception caused by the earplugs. The disruption of spectral cues might thus have hindered relearning of horizontal localization. We found that participants who adapted quickly to modified HRTFs also adapted quickly to shifted ITDs. Thus, either adaptation to one cue enables faster adaptation to the other, or adaptation operates on combined cues. Evidence for integrated coding of interaural time and level differences (Edmonds & Krumbholz, 2014) supports the latter view. Differences between persons, such as in the capacity for plasticity or in the daily sound exposure, may also account for some variability. Understanding this variability may enable customized training for people with assistive hearing devices.

3.4.3 Neural correlates of recalibration

We hypothesized that behavioral adaptation to shifted ITDs would be accompanied by changes in hemispheric lateralization or directional tuning in auditory cortex, based on auditory space encoding by two opponent neuronal populations (Brand et al., 2002; McAlpine et al., 2001; Middlebrooks et al., 1998; Siveke et al., 2006; Stecker et al., 2005; Werner-Reiss & Groh, 2008), which also appears to apply in humans (“hemifield code”, Salminen et al., 2012; Salminen, Tiitinen, Yrttiaho, & May, 2010; Salminen et al., 2009), and the preference for contralateral sounds previously observed in the human auditory cortex (Krumbholz, Schönwiesner, Cramon, et al., 2005; Palomäki et al., 2005; Pavani et al., 2002; D. L. Woods et al., 2009).

We computed voxel-wise tuning curves from the fMRI responses to the different sound directions. Most voxels had broad tuning curves with maxima at the lateral extremes of the tested range of directions and shallow slopes across that range. This result confirmed the proposed hemifield code in the human auditory cortex (Magezi & Krumbholz, 2010; Salminen, Tiitinen, Yrttiaho, & May, 2010; Salminen et al., 2009). Those EEG and MEG studies use a

stimulus-specific adaptation paradigm from which the encoding model is then inferred. The hemifield code model in humans is still under considerable discussion and there is relatively limited evidence for it at the moment, especially when comparing to the older literature that supports place code models (Bernstein, van de Par, & Trahiotis, 1999; Colburn, Isabelle, & Tollin, 1997; Par, Trahiotis, & Bernstein, 2001; Stern & Shear, 1996; reviewed in: Ahveninen, Kopčo, & Jääskeläinen, 2014). We provide new evidence for the hemifield code model in humans using a different imaging modality (fMRI) and a different experimental approach (extracting voxel-wise tuning curves for horizontal sound direction) compared to the previous MEG (Salminen, Tiitinen, Yrttiaho, & May, 2010; Salminen et al., 2009) and EEG (Magezi & Krumbholz, 2010) studies.

Most voxels in a cortical hemisphere responded best to sounds from the contralateral hemifield (maxima and centre of gravity of tuning curves were in the contralateral hemifield). Some previous human neuroimaging studies showed such a contralaterality (Krumbholz, Schönwiesner, Cramon, et al., 2005; Palomäki et al., 2005; Pavani et al., 2002; D. L. Woods et al., 2009), while others did not (Brunetti et al., 2005; Jäncke et al., 2002; Woldorff et al., 1999; Zimmer et al., 2006). Improved methods, such as higher spatial resolution, larger number of directions, and cross-validated responses, helped us to demonstrate this contralaterality more clearly than previously possible. High spatial resolution also appears to be necessary to record reliable BOLD signal changes related to ITD (Dechent, Schönwiesner, Voit, Gowland, & Krumbholz, 2007). Contralateral bias was larger in the left (~ 90 %) than in the right hemisphere (~ 69 %), likely because most stimuli were perceived in the right hemifield in the first session. Contralaterality increased in the right hemisphere during recalibration (to 75 %). Based on this observation and theoretical considerations from the hemifield code model, we hypothesized that behavioral recalibration might be accompanied by a change in the relative magnitude or extent of activation in left and right auditory cortex. Results indicate that this was the case: the ratio of active voxels in both hemispheres (lateralization ratio) changed between sessions for all directions; more voxels were significantly active in the right hemisphere after recalibration. Such a change in lateralization may have been caused by a change in the relative size of the two opponent populations or by a change in tuning. The results of our modeling analysis support a change in

tuning (see 3.4.4. Modeling). Although the extent of the activation changed between sessions, the mean response amplitude was very similar. Prior studies also reported changes in extent but not magnitude of activity in auditory (Jäncke, Shah, Posse, Grosse-Ryken, & Müller-Gärtner, 1998; Jäncke et al., 1999), visual (Goodyear & Menon, 1998; Marcar, Straessle, Girard, Loenneker, & Martin, 2004; Mohamed, Pinus, Faro, Patel, & Tracy, 2002), and motor cortex (Karni et al., 1995). Adjacent cortical areas may be recruited with only minor changes in the activity of the already active region. Alternatively, nonlinear coupling between stimulus change and neural or BOLD responses may increase the extent of the activation with negligible changes in magnitude. In such cases, the extent of the activity may be a more sensitive indicator of a change in brain processing than the amplitude. The lateralization ratio depended on the direction of the sound, and we used this dependence to estimate the corresponding shift in sound direction. This shift was approximately 25° , in the same direction and comparable to the perceptual shift ($\sim 20^\circ$) in the free-field between sessions.

There was no obvious change in average tuning curves between the sessions. However, our modeling results suggested that a small shift may have taken place.

3.4.4 Modeling

Starting from random parameter values, the genetic algorithm consistently found solutions that conformed to expectations based on the literature and closely matched our neuroimaging data. In the best-fitting solutions, the majority of direction-sensitive neurons (64 %) responded in a hemifield code fashion. Their tuning curves were broad Gaussians ($\sigma = 74^\circ$) with maxima around $+90^\circ$ and -90° . These values are close to the parameters of the hemifield code model proposed by Salminen et al. (2009) and place the steepest part of the tuning curves around the medial directions. Solutions also exhibited strong contralaterality (92 %).

The *shift* parameter shifted the presented sound directions at the input stage of the model and could be considered as resulting from subcortical processing of shifted ITDs. Manipulating the *shift* parameter in the model helped us to interpret the pattern of COG distributions. Large changes in the *shift* parameter result only in minor, and always symmetrical, changes of COG

peak locations. This may explain why observed COG distributions were symmetrical around 0° , even though the fMRI stimuli were perceived as shifted to the right. The value of the shift parameter in session one was in the same direction but somewhat larger (59°) than the perceptual shift measured in the perceived midline task (45°). In session two, the *shift* parameter was significantly smaller (52°). This decrease was in same direction as the behavioral recalibration. The change in the *shift* parameter may be evidence for a change in directional tuning, but we cannot determine the processing level at which it might have occurred. Subcortical neurons may have changed tuning driven by corticofugal projections. [Bajo et al. \(2010\)](#) demonstrated that removal of corticofugal projections abolishes experience-dependent recalibration of auditory spatial localization. Tuning may have also changed in the auditory cortex, such that the steepest part of the tuning curves migrated, as proposed by [Salminen, Aho, and Sams \(2013\)](#) to explain enhanced auditory spatial selectivity when attending to off-centre visual stimuli.

3.5 CONCLUSIONS

We used digital earplugs to modify ITDs during several days in a natural environment. We found that adult humans can adapt to this manipulation. This adaptation is accompanied by changes in hemispheric lateralization in auditory cortex observable with high-resolution fMRI. Neural-population modeling suggested that a small shift in direction tuning may also have taken place.

3.6 ACKNOWLEDGMENTS

This work was supported by the National Science and Engineering Council of Canada and the Québec Bio-Imaging Network. RT was supported by the NSERC-CREATE program in Auditory Cognitive Neuroscience. MS was supported by the Fonds de recherche du Québec - Santé.

CHAPITRE 4

ARTICLE 3 : FAST AND PERSISTENT ADAPTATION TO NEW SPECTRAL CUES FOR SOUND

LOCALIZATION SUGGESTS A MANY-TO-ONE MAPPING MECHANISM

Régis Trapeau, Valérie Aubrais and Marc Schönwiesner

Publié dans *The Journal of the Acoustical Society of America*. 140(2), 879-890.

Abstract

The adult human auditory system can adapt to changes in spectral cues for sound localization. This plasticity was demonstrated by changing the shape of the pinna with earmolds. Previous results indicate that participants regain localization accuracy after several weeks of adaptation and that the adapted state is retained for at least one week without earmolds. No aftereffect was observed after mold removal, but any aftereffect may be too short to be observed when responses are averaged over many trials. This work investigated the lack of aftereffect by analyzing single-trial responses and modifying visual, auditory, and tactile information during the localization task. Results showed that participants localized accurately immediately after mold removal, even at the first stimulus presentation. Knowledge of the stimulus spectrum, tactile information about the absence of the earmolds, and visual feedback were not necessary to localize accurately after adaptation. Part of the adaptation persisted for one month without molds. The results are consistent with the hypothesis of a many-to-one mapping of the spectral cues, in which several spectral profiles are simultaneously associated with one sound location. Additionally, participants with acoustically more informative spectral cues localized sounds more accurately, and larger acoustical disturbances by the molds reduced adaptation success.

PACS number: 43.66.Qp

4.1 INTRODUCTION

The human auditory system infers the location of a sound source from different acoustic cues. Interaural time (ITD) and level differences (ILDs), produced by the separation of the two ears, enable sound localization on the horizontal plane. These binaural cues are insufficient to uniquely determine the location of a sound source, because different locations can have the same ITD and ILD, such as locations along the vertical midline (Blauert, 1997). The auditory system disambiguates these locations using spectral cues generated by the direction-dependent filtering of the pinnae and the upper body (head related transfer functions: HRTFs; Morimoto & Ando, 1980; Wightman & Kistler, 1989b).

The morphology of the pinnae and the resultant spectral cues are specific to each individual (Middlebrooks, 1999), and change throughout the entire lifespan (Carlile et al., 2005; Otte, Agterberg, Van Wanrooij, Snik, & Van Opstal, 2013). To acquire and conserve accurate sound localization, the auditory system must be able to recalibrate the association between spectral cues and sound locations over time (Clifton et al., 1988; Hartley & King, 2010). Abrupt changes of spectral cues can also occur, for instance when assistive hearing devices are worn (Byrne & Noble, 1998).

It has been demonstrated that human adults can adapt to sudden and substantial changes in spectral cues (Carlile et al., 2014; Carlile & Blackman, 2013; Hofman et al., 1998; M. M. V. Van Wanrooij & Van Opstal, 2005; Majdak, Walder, & Laback, 2013). In such studies, spectral cues are modified by fitting earmolds to impair sound localization. After an adaptation period of several weeks during which the molds are worn continuously, participants regain significant localization accuracy. Training involving auditory and sensory-motor interactions accelerates the adaptation process (Carlile et al., 2014; Parseihian & Katz, 2012). Even after two months of exposure and adaptation to modified spectral cues, participants exhibit no aftereffect when removing the molds and accurately localize sounds with their native spectral cues. However, it is unknown whether some minimal exposure to the native spectral cues is needed to accurately localize after removing the molds, because many trials are typically averaged when measuring sound

localization performance, which may obscure short-lived effects. Even small indications of an aftereffect are theoretically important, because they may be used to decide between partially different models of the adaptation process. Hofman et al. (1998) suggested that adapting to a new pinna shapes resembles learning a new language, in that a second set of DTFs is stored without interfering much with the original one. Those authors also raise the issue that some cue might be necessary to disambiguate the between the stored sets, although that cue might come from the presented sound itself under certain conditions (Hofman & Van Opstal, 1998). If a switch is needed, it might also be cued by visual feedback, by prior knowledge of the stimulus spectrum, or by tactile sensation and the awareness that the molds were removed. The concept of sets of DTFs may also imply contextual effects when switching between them, just as, for instance, switching from one language to another unexpectedly will incur a perceptual penalty. In a second option, which we refer to as “many-to-one mapping”, new spectral profiles are associated with each sound direction. In this view, there are no distinct sets of DTFs and thus no switch nor any context effects. To decide among these options, we measured single-trial localization performance immediately after mold removal and manipulated visual feedback, tactile feedback, and stimulus spectrum.

The lack of aftereffect raised the question of the persistence of the newly learned spectral cues. Carlile and Blackman (2013) showed that new spectral-to-spatial associations learned over an adaptation period of 30 to 60 days are retained after one week without exposure to the modified cues. Here, we tested the persistence of the adaptation after a long period (one month) that far exceeded the adaptation period in our study (6 days).

Individuals differ in their ability to locate sound sources (Makous & Middlebrooks, 1990; Populin, 2008; Savel, 2009; Wightman & Kistler, 1989a). A portion of these differences may be due to acoustical differences in spectral cues caused by different ear shapes, which may lead to differences in the amount of information provided by a set of spectral cues for discriminating directions. Previous studies have failed to demonstrate a relation between acoustic factors and performance (Andéol, Savel, & Guillaume, 2014; Andéol, Macpherson, & Sabin, 2013; Majdak, Baumgartner, & Laback, 2014), possibly because they used a suboptimal metric to quantify the

efficacy of individual spectral cues. In the present study we develop a new metric to investigate links between acoustic factors and performance.

4.2 METHODS

4.2.1 Participants

Forty-two participants (14 males, mean age 26 ± 5 years), took part in the experiment after having provided informed consent. Thirty took part in the main experiment and twelve took part in control experiments (see Sec. 4.2.6. **Control experiments**). All participants had hearing thresholds of 15 *dB* hearing level (HL) or lower at octave-spaced frequencies between 0.125 and 8 *kHz*. Participants had no history of hearing disorder or neurological disease, and had normal or corrected-to-normal vision. They had unblocked ear canals, as determined by a non-diagnostic otoscopy. The participants received a monetary compensation for their participation. The experimental procedures conformed to the World Medical Association’s Declaration of Helsinki and were approved by the local ethics committee.

4.2.2 Earmolds

Silicone molds were applied to the conchae of the participants to modify ear shapes and distort spectral cues. Participants were informed that the ear molds would “slightly modify their sound perception”, but did not know the specific effects of this modification. The molds were created by applying a fast-curing medical-grade silicon (SkinTite, Smooth-On, Macungie, PA, USA) in a tapered layer of about 3 to 7 *mm* thickness on the cyma conchae and the cavum conchae of the external ear, keeping the ear canal unobstructed.

4.2.3 Overall procedure

The experiment was designed to disrupt participants’ spectral cues using earmolds, to measure the acoustical and behavioral effects of the molds, to follow participants’ adaptation to the new listening situation, and to test whether this adaptation persisted after mold removal.

Additionally, we wanted to explore the mechanisms underlying the absence of aftereffects. Participants were randomly assigned to one of three conditions: *control* ($n = 15$), *visual* ($n = 8$), or *spectral* ($n = 7$). A subset of 12 participants was assigned to a *headphone* condition. These groups differed only in the type of aftereffect test; apart from this test, the experimental protocols were identical across participants. The visual condition was designed to test whether visual feedback aids reverting to the original spectral cues after mold removal. The spectral condition was designed to test whether knowledge of the stimulus spectrum is necessary to revert to the original spectral cues after mold removal. The headphone condition was designed to test whether the tactile sensation of not wearing molds triggers the switch to the original spectral cues.

To minimize procedural learning and stress during the experiment, participants completed at least one free-field localization run beforehand to familiarize themselves with the task and all equipment. Participants performed the task until they were comfortable with it. No feedback was given. Once this familiarization was complete, participants performed one localization run which served as baseline. During their first lab visit, participants assigned to the visual or spectral conditions completed an additional baseline localization run in the free-field, specific to their condition group.

After the free-field baseline tests were completed, HRTFs were acquired for all participants. Individual binaural recordings used as stimuli for the subset of participants assigned to the headphone condition were also acquired at this stage. The participants assigned to the headphone condition also performed a baseline localization task with headphones.

Earmolds were then fitted to the participants' ears and participants immediately repeated the standard free-field localization task. HRTFs were acquired again, this time with the molds on. From that point on, participants wore the earmolds for six consecutive days.

Starting on the day after the earmolds were fitted (day 1 in Figure 4.6), the participants performed a daily routine of free-field localization tasks and training sessions, which continued for six days. On the day the molds were removed, all participants completed a free-field localization run with the molds on, then removed their molds and immediately performed

the free-field localization task specific to their condition group. Participants involved in the headphone condition performed an additional headphone localization run while still wearing the molds.

To explore the persistence of the adaptation to disrupted spectral cues, participants returned to the lab twice, one week and one month after mold removal. During these two visits, molds were briefly re-inserted and participants completed a free-field localization run while wearing the molds.

Behavioral testing, training, and binaural recordings were conducted in a hemi-anechoic room ($2.5 \times 5.5 \times 2.5$ m). Participants were seated on a comfortable chair with a neck rest, located in the centre of a spherical array of 80 loudspeakers (Orb Audio, New York, NY, USA) with a diameter of 1.8 m. An acoustically transparent black curtain was placed in front of the loudspeakers to avoid visual capture. Only a small label indicating the central location (0° in azimuth and elevation) was visible on the curtain. A laser pointer and an electromagnetic head-tracking sensor (Polhemus Fastrak, Colchester, VT, USA) were attached to a headband worn by the participant. The laser light reflected from the curtain served as visual feedback for the head position. Real-time head position and orientation were used to calculate azimuth and elevation of indicated directions.

4.2.4 Stimuli

The stimuli in the free-field sound localization task were 225 ms long trains of pulsed pink noise (five equally-spaced bursts of 25 ms duration). The overall level of the stimuli was 60 dB sound pressure level (SPL) at the position of the participant's head. A shorter version (125 ms, 3 bursts) of this stimulus was used during the training sessions. Stimuli were created afresh on each trial on TDT System 3 hardware (Tucker Davis Technologies, Alachua, FL, USA), at 50 kHz sampling rate and 24 bit amplitude resolution, and controlled by custom MATLAB (Mathworks, Natick, MA, USA) scripts. In order to minimize location cues caused by differences in loudspeaker transfer functions, we used a Brüel & Kjær $\frac{1}{2}$ -inch probe microphone mounted

on a robotic arm to measure these functions and designed inverse finite impulse response filters to equalize amplitude and frequency responses across loudspeakers.

Stimuli in the spectral condition were 225 *ms* excerpts of the stimuli described in Zatorre, Bouffard, and Belin (2004). Forty-five environmental sounds and random mixtures of 8, 15, 30 or all of these sounds were presented for a total of 181 different stimuli. Each stimulus had a unique spectrum. Stimuli were presented in a pseudorandom order such that no stimulus was repeated within any sequence of ten trials.

Stimuli in the headphones condition were individual binaural recordings (ears free) of the free field stimulus described above. The recordings were played through DT 990 Pro headphones (Beyerdynamic, Heilbronn, Germany).

4.2.5 Localization tasks

Thirty-seven loudspeakers were used in the free-field sound localization tasks, covering directions from -45° to $+45^\circ$ in azimuth and elevation. Thirteen loudspeakers were equally distributed (azimuthal angle between adjacent speakers of 7.5°) on a horizontal arc (0° elevation), and 12 loudspeakers were equally distributed on a vertical arc (0° azimuth). Twelve additional loudspeakers were symmetrically distributed in the quadrants created by the two arcs (angle between speakers of 22.5°). Stimulus presentation and response procedures were identical to the ones described in Trapeau and Schönwiesner (2015). Sound directions were presented in a pseudorandom order and were repeated five times during a run. After each sound presentation, participants indicated the perceived direction by turning their head towards the sound source and pressing a button on a hand-held button box. At the beginning of a run, participants were asked to sit comfortably and rest their neck on a neck rest, so that the head was centered in the loudspeaker array and the head-mounted laser pointed to the central location. This initial head position was recorded and the participant had to return to this position with a tolerance of 2 *cm* in location and 2° in head angle before each trial. If the head was correctly placed when the button was pressed, a stimulus was played from one of the speakers. If the head was

misplaced, a brief warning tone (sinusoid of 150 Hz) was played from a speaker located above the participant's head (Az: 0°, El: 82.5°).

The localization procedure in the spectral condition was identical, but a sound with a unique spectrum was presented in each trial. The localization procedure in the headphones condition was also identical, except that binaural recordings of the stimuli were presented through headphones. In the visual condition, a visual feedback was provided after each response by briefly blinking a light-emitting diode (LED) at the location of the sound source to indicate the correct response.

4.2.6 Control experiments

Because head-turning in the frontal field is a relatively natural response we hypothesized that it would involve little to no procedural learning. We estimated the effect of procedural learning by running a separate experiment with six participants who followed the same procedure as in the main experiment, but always with free ears. The results showed that the localization performance of these participants remained stable throughout the six training days. We thus concluded that no significant procedural learning occurred during the adaptation period of the main experiment. Results of this complementary experiment are shown along with the results of the main experiment in Figure 4.6.

During the main experiment, front-back confusions were not recorded. Instead, participants were asked informally whether any of the stimuli appeared to come from the back. Participants never reported to have perceived a stimulus coming from the back with free ears and rarely with molds. The majority of the participants did not experience any front-back confusions at all. In the event of a stimulus heard coming from the back, the participants were asked to transpose their perception to the frontal hemisphere.

To quantify the ratio of front-back confusions that occurred during the main experiment and to assess the effect of transposing a rear percept to the front, we ran a control experiment with six additional participants. These participants followed the same familiarization tests as in the main experiment and completed two sound localization tests, one with ears free and one with molds.

Participants were asked to report each trial in which the stimulus was perceived as coming from behind, and the experimenter recorded these occurrences. No front-back confusions were reported with free ears and only two out of the six participants reported some trials as perceived from the back when wearing the molds. The front-back confusion rates of these two participants were 2.7 %, and 8.1 %, respectively. The average front-back confusion rate with molds was thus 1.8 % in these six participants. For comparison, [Carlile and Blackman \(2013\)](#) and [Carlile et al. \(2014\)](#) reported a confusion rate of over 10 % for listeners wearing molds, and a rate of 1 %–2 % with free ears. [Makous and Middlebrooks \(1990\)](#) reported a rate of 6 % for listeners with free ears. The lower confusion rates in our study likely result from differences in the experimental setup and procedures (see Sec. 4.4. Discussion).

To assess the effect of transposing a rear percept to the front, we analyzed the results of the participant who exhibited a front-back confusion rate of 8.1 % with molds, by comparing his results before and after removing the trials in which front-back confusions occurred. Removing the front-back confusion trials left the results essentially unchanged [before vs after removing front-back confusion trials, vertical performance: root mean squared error (RMSE) = 23.71 vs. 24.01, standard deviation (SD) = 7.01 vs. 6.61, elevation gain (EG) = 0.3 vs. 0.29; horizontal performance: RMSE = 9.99 vs. 10.09, SD = 3.17 vs. 3.06]. We concluded that front-back confusions did not impact the presented results of the main experiment.

4.2.7 Training

The training sessions were 15 *min* long and the procedure was inspired by the one described by [Parseihian and Katz \(2012\)](#). A continuous train of noise bursts (see 4.2.4. Stimuli) was presented from a random speaker and the delay between stimuli depended on the angular difference between the sound direction and the participant's head direction; the smaller the difference, the faster the rate. Participants were instructed to identify the sound source location as fast as possible by pointing their head in the perceived direction of the sound source. Once the participants held their head in the correct direction for 500 *ms*, the stimulus was considered

found and changed location to a random loudspeaker at least 45° away from the previous one. Participants were instructed to find as many locations as possible during the run.

4.2.8 Statistical analysis

Vertical and horizontal localization accuracies were quantified by computing the root mean square of the discrepancies between perceived and physical locations (RMSE, [Hartmann, 1983](#); [Savel, 2009](#)). The vertical and horizontal variance of individual responses was quantified by computing the standard deviation (SD) of response angles for each target direction and then taking the mean across directions. Vertical localization performance was also quantified by the elevation gain (EG), defined as the slope of the linear regression of perceived versus physical elevations ([Hofman et al., 1998](#)). Perfect localization corresponds to an EG of 1, while random elevation responses result in an EG of 0. Throughout the analysis, paired comparisons were statistically assessed using Wilcoxon signed-rank tests. Relations between measures were determined using Spearman correlation coefficients, p-values smaller than 0.05 were regarded as significant.

To analyze trial-by-trial accuracy during the localization test that immediately followed mold removal, we plotted the trial-by-trial error in the localization test immediately after mold removal. Baseline localization accuracy, however, varies across directions ([Carlile et al., 1997](#); [Makous & Middlebrooks, 1990](#)). To account for such differences, we calculated the mean and standard deviation of the localization error for each direction independently in the baseline test, and then standardized the trial-by-trial errors in the aftereffect test with these values. A plot of the standardized trial-by-trial errors thus shows the severity (in terms of how many standard deviations away) of the localization error in the aftereffect test was for a given trial with respect to typical performance for the same direction in the baseline test. We then tested the hypothesis of a decreasing localization error over the first ten trials using a Mann-Kendall trend test.

4.2.9 Binaural recordings

Individual HRTFs were recorded with and without molds, and their direction-dependent components, the directional transfer functions (DTFs), were extracted (Middlebrooks, 1999). Chirps of 5 *ms* duration were presented 50 times from each of the 80 loudspeakers and recorded with binaural in-ear microphones. Repetitions were averaged for each location. ER7-14C probe microphones (Etymotic Research, Elk Grove, Ill) were used for the first 12 participants. For the remaining 18 participants we used miniature microphones (Sonion 66AF31), placed 2 *mm* inside of the entrance to the ear canal. Such recordings from the canal entrance do not contain the non-directional ear canal transfer function. However, non-directional contributions to the HRTF were removed from the recordings (see 4.2.10. Directional transfer functions) and thus the change of microphone and placement did not significantly influence the recordings and results. Participants were asked not to move during the recordings and keep the reflection of the head-mounted laser on the fixation mark. Head alignment was verified immediately after each recording by computing the ITD for sounds presented from the horizontal plane. If a deviation from the midline was detected ($> 1^\circ$) the recording was repeated. Individual binaural recordings of the stimuli used in the free-field sound localization task were also acquired for the 12 participants who took part in the headphone condition.

4.2.10 Directional transfer functions

DTFs were extracted from the binaural recordings without and with molds using the method proposed by Middlebrooks (1999). The recordings at each location were filtered by a bank of triangular band-pass filters equally spaced at 0.0286 octaves (2 % frequency difference). To remove non-directional portions of the transfer functions (including microphone and ear canal transfer functions) each measured transfer function was divided by the grand average (RMS) of the transfer functions of a spatially uniform sample of 48 directions (equally spaced 22.5° apart in azimuth and elevation, i.e., including the rear field). DTFs for recorded directions located on the median plane, i.e., 13 elevations from -45° to $+45^\circ$, were then analyzed further.

Figures 4.1.A and 4.1.B show the DTFs obtained for the left ear of one participant, without and with mold respectively.

For each DTF set obtained without molds, we computed the autocorrelation matrix in the median plane (Figure 4.1.C), i.e., the matrix of correlation coefficients between the DTFs of each elevation against each other (Hofman & Van Opstal, 1998). From this autocorrelation matrix we computed a measure [“vertical spectral information” (VSI)] of how well different elevations can potentially be discriminated by a listener using a given set of DTFs. VSI measures the average dissimilarity of DTFs across elevations and is defined as one minus the mean of the coefficients contained in the autocorrelation matrix (excluding the main diagonal). A DTF set of identical transfer function for all elevations will result in a VSI of 0, whereas highly different transfer functions will result in a high VSI. The maximal VSI in our recordings was 1.07, which is very close to the empirical maximum for our number of elevations of 1.072 found by computing the VSI of 10000 randomly generated DTF sets of 13 elevations and then taking the maximum VSI of this sample.

To aid comparison with previous studies, we also calculated the spectral strength (which quantifies the average amount of spectral detail in each DTF), as defined by Andéol et al. (2013, 2014), and derived from Middlebrooks (1999). The spectral strength is the variance of each DTF, averaged across elevations. The drawback of this measure is that identical DTFs across elevations may give a high spectral strength value even though the listener would be unable to discriminate elevations.

To determine the dissimilarity between the DTFs obtained with and without mold, we first computed the correlation matrix between all DTFs measured with free ears vs. all DTFs measured with molded ears (Figure 4.1.D) and calculated the root-mean-square distance between this correlation matrix and the autocorrelation matrix of the DTFs measured with free ears. This measure is referred to as the “VSI dissimilarity” throughout the text.

We computed VSI, VSI dissimilarity and spectral strength, separately for the left and right ear. When examining the relation between these acoustical metrics and behavior, left and right ear values were averaged.

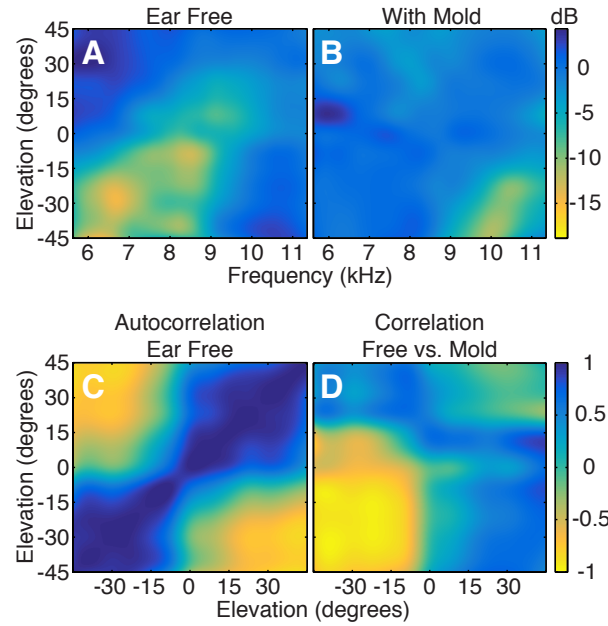


Figure 4.1 – Illustration of the acoustic effects of an earmold on the DTFs of one participant. A) DTFs of the free left ear of one participant, computed in the 5657–11314 *Hz* frequency band. The color code gives the output spectral amplitude for each frequency and elevation. Each row of these values thus corresponds to the spectral profile of the DTF at that elevation. Values in between the 13 measured elevations and 36 frequencies were linearly interpolated for display. B) DTFs of the same ear fitted with a mold. C) Autocorrelation matrix obtained from the DTFs measured on the free ear. These matrices contain values varying between -1 and 1 (correlation coefficients), are symmetric, and values along the main diagonal are always unity. The mean of the matrix, excluding the diagonal, is the VSI (0.83 in this case). D) Correlation matrix between the two DTFs sets, ear free vs. with mold. The VSI dissimilarity between free ear and ear with mold is the euclidean distance between the matrices in C and D (1.08 in this case).

4.3 RESULTS

4.3.1 Relation between VSI (Vertical Spectral Information) and ears free performance

We computed VSI and spectral strength from the DTF sets obtained without molds in five octave bands between 4 kHz and 16 kHz ($4\text{--}8\text{ kHz}$, $4.8\text{--}9.5\text{ kHz}$, $5.7\text{--}11.3\text{ kHz}$, $6.7\text{--}13.5\text{ kHz}$, $8\text{--}16\text{ kHz}$). VSIs varied significantly among frequency bands (Kruskal-Wallis test, $p = 10^{-7}$), being highest in the $5.7\text{--}11.3\text{ kHz}$ band (Figure 4.2). The previously used spectral strength measure (Andéol et al., 2013, 2014; Middlebrooks, 1999), however, did not vary significantly among frequency bands. When computing VSI in non-overlapping $1/2$ -octave bands ($4\text{--}5.7\text{ kHz}$, $5.7\text{--}8\text{ kHz}$, $8\text{--}11.3\text{ kHz}$, $11.3\text{--}16\text{ kHz}$), it correlated negatively with vertical RMSE in the $5.7\text{--}8\text{ kHz}$ frequency band (Figure 4.5.A; $R = -0.53$, Bonferroni corrected $p = 0.0126$). In this frequency band,

the relations between the VSI and other elevation metrics were in the predicted directions but not significant (EG: $R = 0.28$, $p = 0.13756$; SD-El: $R = -0.18$, $p = 0.34952$). No relation was found between the VSI and performance in the others $1/2$ -octave frequency bands. Spectral strength did not correlate with behavioral data in any of the frequency bands tested.

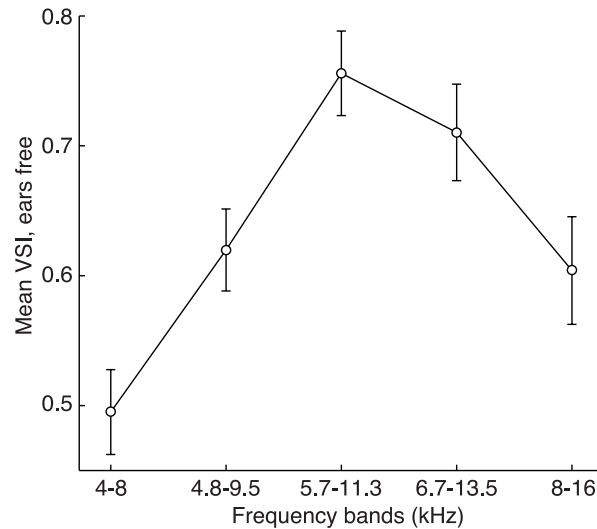


Figure 4.2 – Mean VSI across octave bands. Mean VSI from the DTF sets obtained without molds in five octave bands between 4 kHz and 16 kHz. VSI varies significantly among frequency bands. Maximum VSI is located in the 5.7–11.3 kHz band.

4.3.2 Acoustical effect of the molds

Modification of the relief of the concha reduced the notch situated in the 8–11 kHz frequency band (Figures 4.3.A, 4.3.B, 4.3.C). A comparable pattern was observed in previous studies (Carlile et al., 2014; Carlile & Blackman, 2013; Hofman et al., 1998; M. M. V. Van Wanrooij & Van Opstal, 2005).

To verify the consistency of spectral changes across participants, we computed a map of the proportion of participants for which the molds induced marked changes in spectral amplitude at each elevation and frequency (Figure 4.3.D). “Marked changes” were absolute differences between the DTFs measured with and without molds that were above a given threshold. This threshold was a participant-specific measure of spectral differences across elevations with free ears, that was computed as the mean absolute difference between all combinations of the RMS

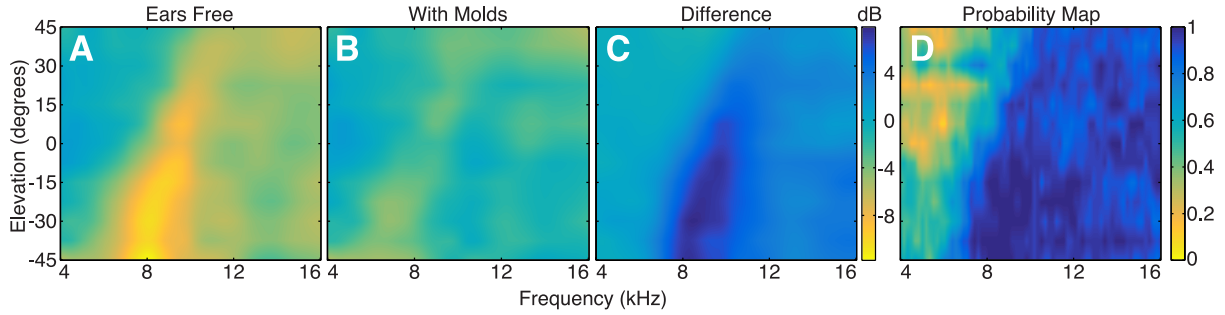


Figure 4.3 – Mean acoustic effect of an earmold on the DTFs. A) Mean DTFs of all participants with free ears, computed between 4 and 16 kHz. B) Mean DTFs of all participants with molds. C) Gain difference of A and B. D) The probability map shows the proportion of participants for which the molds induced marked changes in spectral amplitude at each elevation and frequency bin.

of the DTF (in dB) at each elevation (average across participants: $2.02 \text{ dB} \pm 0.14$). Using this threshold, we created, for each participant, a binary map of marked changes for each elevation and frequency bin (i.e., values above the threshold were set to one; all other values were set to zero). The average of these maps across participants shows the proportion of participants for which the molds induced above-threshold changes in spectral amplitude at each elevation and frequency. This proportion was $62 \% \pm 3$ (mean across participants $\pm$ standard error) in the 4–8 kHz frequency band, and $93 \% \pm 1$ in the 8–16 kHz band.

We checked that the addition of the molds to our participants’ ears did not simply decrease acoustical differences between vertical sound locations and thus reduce the amount of spectral information available for vertical localization. Such a reduction may have led to physiologically implausible DTFs. The VSI computed with and without molds between 5.7 and 11.3 kHz (the octave frequency band with highest VSI) showed that the molds did not reduce the acoustical differences between elevations (ears free = 0.76 ± 0.02 ; with molds = 0.81 ± 0.03). The left and right ear VSI without molds were correlated, as predicted ($R = 0.47$, $p = 0.00875$). This correlation still existed after fitting the molds ($R = 0.68$, $p = 0.00003$). The VSI with free ears and with molds of all participants are shown in Figure 4.4.A.

To ensure that the spectral modification induced by the molds was in a physiologically plausible range, i.e., no larger than the variability due to different ear shapes between individuals, we computed the VSI dissimilarity (in the 5.7–11.3 kHz frequency band and for each ear) between

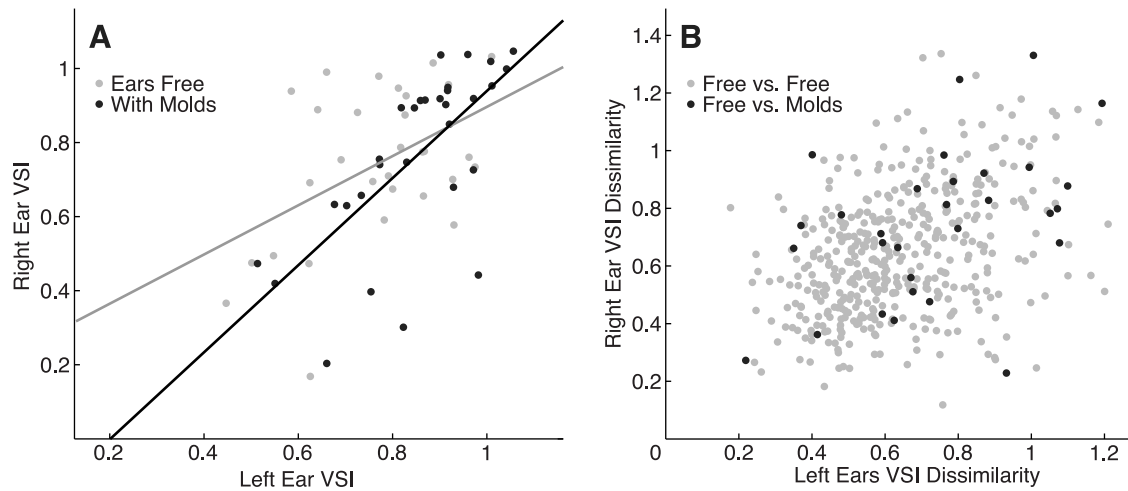


Figure 4.4 – Effects of the molds on the VSI (5.7–11.3 kHz frequency range). A) Left and right ear VSI. Each dot represent one participant. Grey dots are VSI values of free ears, black dots are VSI values with molds. Regression lines show the correlations between VSI obtained for the left and right ears. VSI ranges with and without molds largely overlap, indicating that the molds did not decrease VSI. The molds also conserved natural similarity between left and right ear VSI B) VSI dissimilarity. Grey dots show the VSI dissimilarity of free left and right ears between all combinations of two participants. Each black dot shows the VSI dissimilarity between free ears and molds for one participant. Convex hulls of both dot clouds largely overlap, indicating that the molds did not induce physiologically implausible acoustic changes.

every possible pair of participants and compared it with the VSI dissimilarities between free and molded DTFs of each participant (Figure 4.4.B). The overlap between both distributions confirmed that spectral modifications induced by the molds were of similar magnitude as the typical difference of DTFs between individuals.

4.3.3 Behavioral effects of the molds

The insertion of the molds degraded the performance on both the vertical and the horizontal plane (Figure 4.6, day 0; one-tailed Wilcoxon signed rank tests of performance ears free vs. molds; vertical plane, ears free vs. molds: RMSE = 10.74 ± 0.4 [mean $\pm$ standard error] vs. 24.54 ± 1.5 , $p = 9 \times 10^{-7}$; SD = 4.84 ± 0.23 vs. 8.43 ± 0.75 , $p = 6 \times 10^{-5}$; EG = 0.8 ± 0.03 vs. 0.29 ± 0.05 , $p = 10^{-6}$; horizontal plane: RMSE = 5.98 ± 0.26 vs. 11.53 ± 1.62 , $p = 2 \times 10^{-5}$; SD = 2.67 ± 0.13 vs. 4.89 ± 0.71 , $p = 7 \times 10^{-5}$). Molds reduced horizontal localization accuracy at all elevations, but most strongly at low elevations (10.17° at -45°) and less so at high elevations (3.06° at $+45^\circ$). Reduced horizontal accuracy was not explained by changes in binaural cues

used for horizontal sound localization, because the molds had no effect on these cues. In the tested range of azimuths (-45° to $+45^\circ$), ILDs and ITDs vary linearly across the horizontal plane. We computed linear regressions of each participant's ILDs and ITDs and compared the slopes and the intercepts with and without molds. Wilcoxon signed rank tests confirmed that ILD and ITD were unchanged by the molds (ILDs linear regression slopes: 0.34 ± 0.1 vs. 0.33 ± 0.1 , $p = 0.21$; ILDs linear regression intercepts: $0.30 \text{ dB} \pm 0.30$ vs. $0.21 \text{ dB} \pm 0.25$, $p = 0.84$; ITDs linear regression slopes: 8.10 ± 0.08 vs. 8.08 ± 0.09 , $p = 0.76$; ITDs linear regression intercepts: $8.99 \mu\text{s} \pm 7.19$ vs. $6.69 \mu\text{s} \pm 6.28$, $p = 0.82$).

Across individuals, the amount of degradation in vertical and horizontal performance was correlated (RMSE: $R = 0.71$, $p = 2 \times 10^{-5}$; SD: $R = 0.85$, $p = 6 \times 10^{-7}$). The regression between the vertical and horizontal degradation indicates that the minimum increase in vertical RMSE that caused an increase in horizontal RMSE larger than one SD measured with free ears was 12° .

4.3.4 Relation between behavioral effects of the molds and VSI dissimilarity

The insertion of the molds produced large behavioral effects. To determine whether the behavioral effects of the molds were related to the acoustical dissimilarity between DTFs with and without molds, we calculated the relation between these effects and the VSI dissimilarity. We found a correlation between the difference in initial elevation performance with molds (first test with molds compared to free ear baseline) and VSI dissimilarity in the 5.7–11.3 kHz frequency band (Figures 4.5.B, 4.5.C; RMSE: $R = 0.52$, $p = 0.00359$; EG: $R = -0.53$, $p = 0.00275$). As expected, larger spectral changes caused by the molds led to larger localization errors and smaller EG. Behavior and VSI dissimilarity still correlated after six days of wearing the molds (Figures 4.5.D, 4.5.E; Last test with molds – ears free baseline, RMSE: $R = 0.51$, $p = 0.00477$; EG: $R = -0.39$, $p = 0.03550$). Even if the relation went in the predicted direction, no significant correlation was found between the vertical SD and the VSI dissimilarity. To test whether the relation between acoustical (VSI) dissimilarity and localization accuracy found in the vertical plane also holds in the horizontal plane, we computed a measure analogous to the VSI across horizontal sound directions. However, this measure did not correlate with the difference in

horizontal localization accuracy. Differences between the left and right ear DTF may also influence horizontal sound localization (Hofman & Van Opstal, 2003), however, we found no correlation between changes in binaural DTF dissimilarity due to the molds and reduction of horizontal localization performance.

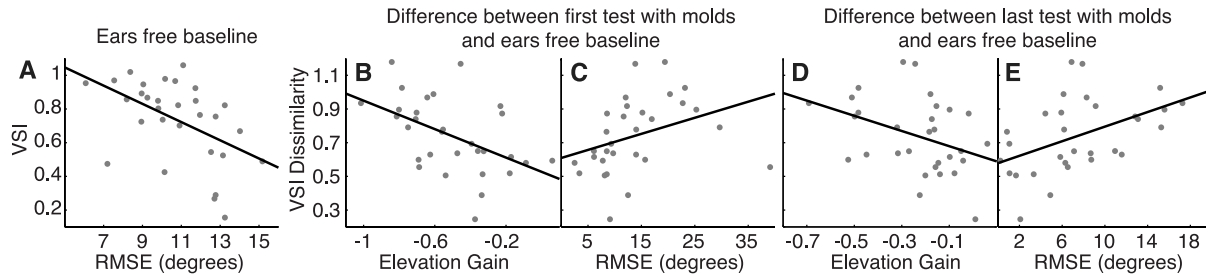


Figure 4.5 – Correlations between behavioral results and acoustical metrics. Each dot represents a participant ($n = 30$). A) Correlation between the vertical localization error (RMSE) and the acoustic information content (VSI) with free ears in the 5.7–8 kHz frequency band. B–E) Correlation between VSI dissimilarity (free vs. mold) in the 5.7–11.3 kHz band and vertical behavioral measures: EG (B) and RMSE (C) differences between the first test with molds and the free ear baseline; EG (D) and RMSE (E) differences between the last test with molds and the free ear baseline.

4.3.5 Adaptation

The combination of daily multisensory experience and sensory-motor training at the lab for six days resulted in improved sound localization performance with the molds. The performance on the last day with molds were significantly better than on the first day with molds for all metrics except vertical SD (one-tailed Wilcoxon signed rank tests of first vs. last test with molds; vertical plane: RMSE = 24.54 ± 1.5 vs. 18.58 ± 0.95 , $p = 10^{-6}$; SD = 8.43 ± 0.75 vs. 7.98 ± 0.58 , $p = 0.1995$; EG = 0.29 ± 0.05 vs. 0.55 ± 0.04 , $p = 2 \times 10^{-6}$; horizontal plane, RMSE = 11.53 ± 1.62 vs. 8.49 ± 0.95 , $p = 0.0171$; SD = 4.89 ± 0.71 vs. 3.56 ± 0.31 , $p = 0.0458$). The development of RMSE, SD, and EG from day 0 to day 6 is shown in Figure 4.6.

To compare differences in adaptation across participants independent of differing acoustical disruption caused initially by the molds, we divided the amount of reduction in localization error (RMSE with molds on day 0 vs. on day 6) by the amount of initial increase in error caused by the molds (RMSE on day 0 with molds vs. without molds). Individuals showed a continuum of adaptation success from no adaptation to full recovery of localization performance and did

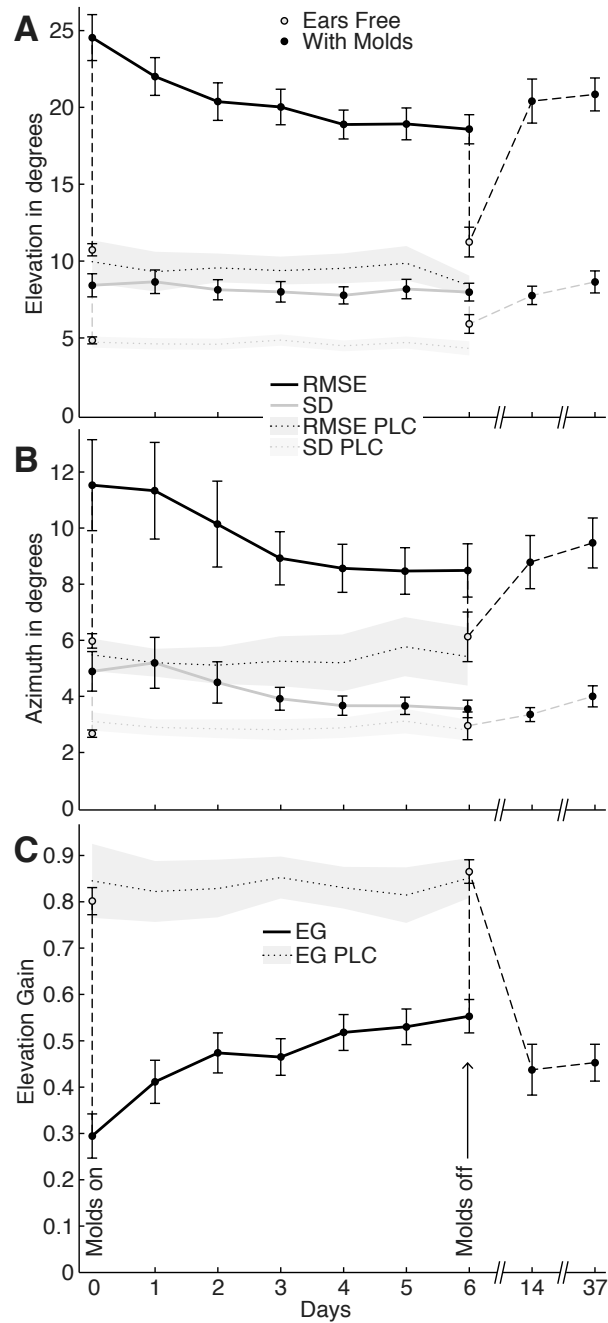


Figure 4.6 – Time course of sound localization performance. A) Time course of localization error (RMSE) and response variability (SD) on the vertical plane, and (B) on the horizontal plane. The insertion of the molds significantly increased RMSE and SD on both planes. During the following 6 days, localization error decreased and elevation gain increased. Immediately after mold removal, performance returned to baseline level (grey dot on day 6). C) Time course of the elevation gain. Participants had high elevation gains without molds (~ 0.8 , grey dot on day 0). The insertion of the molds decreased the elevation gain to ~ 0.3 . Elevation gain increased during adaptation to 0.55 on day 6 (black dots). Immediately after mold removal, elevation gain returned to baseline level (grey dot on day 6). Grey circles indicate performance without molds, black circles indicate performance with molds. Error bars indicate standard error. Data points represent the mean across all participants ($n = 30$), except for the test with free ears on day 6, which included only control group participants ($n = 15$) but not participants in the visual and headphone groups. Dotted lines are the results of the control experiment that aimed to quantify procedural learning during the experiment (PLC: procedural learning controls, $n = 6$). The localization performance (vertical and horizontal RMSE and SD, EG) of these participants remained stable throughout the six training days. No improvements due to procedural learning were found in any of the measures.

not fall into discernable groups, such as low and high performers. Individual differences in the increase in elevation gain and in the reduction in localization error with adaptation are also apparent in Figures 4.5.D and 4.5.E. The amount of adaptation was independent of age and gender.

We examined the relation between the amount of vertical and horizontal adaptation, i.e., the correlation between the difference in performance along the vertical and horizontal planes from day 0 to 6. Vertical and horizontal adaptation correlated strongly (RMSE: $R = 0.54$, $p = 0.00223$, SD: $R = 0.51$, $p = 0.00486$).

4.3.6 Aftereffect

To test whether the adaptation to the molds induced an aftereffect, the participants removed the molds after the last free-field localization run with molds (day 6 in Figure 4.6), and immediately performed a free-field localization task specific to their condition group. For each group, we tested whether the performance obtained during this test was reduced compared to the baseline performance. Immediately after removing the molds, sound localization performance returned to near baseline values. In the control group, there was a small effect in the vertical SD, suggesting increased response variability on the vertical plane when compared to baseline ($SD = 4.89 \pm 0.37$ vs. 5.91 ± 0.61 , $p = 0.0453$). The visual, spectral, and headphone conditions had no effect on localization performance.

Participants may re-adapt to their original ears during the first few sound presentations after mold removal, thus too fast to observe in the average accuracy across an entire localization run. To test this hypothesis, we analyzed trial-by-trial accuracy during the localization test that immediately followed mold removal (Figure 4.7). No consistent changes in accuracy were apparent during the test; specifically, performance in the first few trials was not obviously differ from that in later trials for any of the groups. We also tested for a trend of decreasing localization error during the first few trials. No such trends were found in any of the conditions (Mann-Kendall trend test, all $p > 0.3$).

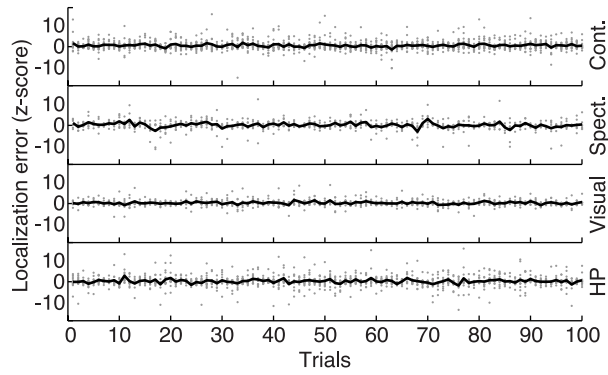


Figure 4.7 – Trial-by-trial localization error in the aftereffect tests. Black lines display the mean localization error across participants. Each grey dot represents the error of one participant in a given trial. Errors were standardized by the direction-specific localization error in the respective baseline tests. Results in each condition are represented (cont. = control, spect. = spectral, HP = headphone). No consistent changes in localization accuracy were apparent during the test in any of the conditions.

4.3.7 Persistence tests

Vertical performance at the 1-week and 1-month post-tests (Figure. 4.6, days 14 and 37) was not as impaired as performance at the first test with molds (day 0). On both planes, 1-week post-test performance resembled the performance achieved three days after insertion of the molds (day 3, 70 % of the adaptation retained on the vertical plane, 90 % on the horizontal plane), while 1-month post-test performance was closest to the performance two days after insertion of the molds (day 2, 62 % of the adaptation retained on the vertical plane, 68 % on the horizontal plane). The trial-by-trial accuracy of both post-tests showed the same pattern as the one described in Sec. 4.3.6. **Aftereffect:** no changes in accuracy were apparent during the run, specifically, performance in the first few trials was not different from that in later trials.

4.4 DISCUSSION

We modified spectral cues for sound localization in adult listeners over several days to better understand brain mechanisms of adaptation to modified sensory input. Fitted silicone earmolds initially disrupted sound localization, but participants regained significant localization accuracy by continuously wearing the molds and participating in daily training sessions. Localization accuracy returned to baseline immediately after removing the molds. Visual feedback, unknown

sound spectrum, and tactile feedback of earmolds had no impact on this lack of an aftereffect. Retests with molds a week or a month after removal showed that participants retained the ability to localize sound with the modified cues.

4.4.1 Acoustic factors of behavioral performance

Horizontal and vertical sound localization abilities without molds were comparable to previous reports of two-dimensional human sound localization performance (Hofman et al., 1998; Makous & Middlebrooks, 1990; Middlebrooks, 1997; S. R. Oldfield & Parker, 1984b; Trapeau & Schönwiesner, 2015). Front back confusions were very rare and had no impact on the presented results (see Sec. 4.2.6. Control experiments). The majority of the participants did not experience any front-back confusions at all. This is likely due to several factors: First, molds cause much less front-back confusions for sound stimuli located in the visual field, than for stimuli located outside of the visual field (Carlile & Blackman, 2013). Most importantly, while we controlled the position and orientation of participants' heads, they were able to do head micro-movements, which are known to resolve front-back confusions (Perrett & Noble, 1997a; Wightman & Kistler, 1999). When spectral cues are absent, dynamic ITD is a very salient cue for front-back discrimination (Macpherson, 2013).

Individuals differed in vertical localization accuracy, and we aimed to determine to what extent these differences were explained by acoustic differences in directional transfer functions caused by different ear shapes. To quantify the efficacy of the spectral cues for vertical localization provided by a given ear, we defined a new measure (VSI) that reflects how well different elevations can potentially be discriminated by the direction-dependent filtering of that ear. VSI varied significantly across 1-octave frequency bands and was highest between 5.7 and 11.3 kHz, which indicates that spectral information (DTF shape) in this frequency band was most useful to discriminate between elevations. Using a very different method, the same band was found to be critical for elevation judgments by selectively removing frequency bands in a virtual auditory display (Langendijk & Bronkhorst, 2002). The correspondence between these behavioral results and our acoustical results strongly suggests that the main spectral cues for vertical sound

localization lie in this frequency band. The 5.7–11.3 kHz frequency band tends to contain a prominent spectral notch (Middlebrooks, 1997; Raykar, Duraiswami, & Yegnanarayana, 2005). Spectral notches are known to carry elevation information. In the cat, notch-sensitive neurons in the dorsal cochlear nucleus and their projections to type O units of the central nucleus of the inferior colliculus are believed to be involved in early-stage processing of spectral cues (Davis et al., 2003; Imig et al., 2000; E. D. Young et al., 1992). The importance of this frequency band in human elevation perception may thus be due to the presence of prominent notches.

VSI also varied between individuals, and this variation was correlated with individual vertical localization accuracy in the 5.7–8 kHz frequency band. Individual differences in vertical localization abilities were previously reported (Populin, 2008; Wightman & Kistler, 1989a), but no relation between this variability and acoustic factors has been reported (Andéol et al., 2013, 2014; Majdak et al., 2014). Studies investigating this relation used spectral strength (Andéol et al., 2013) as a metric to quantify the quality of the spectral cues in each individual. Spectral strength is a measure of the overall saliency of the spectral cues, but does not provide information on the efficacy of the spectral cues to discriminate between elevations. Spectral strength did not vary across frequency bands, nor did it correlate with behavior. VSI correlated with vertical localization error in the 5.7–8 kHz frequency band, in which it explained ~ 25 % of the behavioral variance. As suggested by previous studies, non-acoustic factors such as attention, perceptual and motor abilities, and neural processes involved in spectral cue analysis, might also contribute to individual differences in vertical sound localization.

Using a measure of VSI dissimilarity, we showed that the amount of acoustical change in spectral cues caused by the molds correlated with the loss in elevation performance that followed the insertion of the earmolds. This correlation was maintained at the end of the adaptation period, showing that the speed and amount of adaptation to modified spectral cues depends on the severity of the acoustic disruption. VSI dissimilarity may be useful in estimating to what extent different assistive hearing devices disrupt sound localization while patients adapt to the device. We computed VSI dissimilarity on elevations restricted to the median plane, but to explore ear-specific effects of molds (or hearing devices) DTFs should be extracted from

elevations off the midline where VSI dissimilarity may be somewhat larger due to the orientation of the pinna axis.

4.4.2 Effect of the molds on horizontal sound localization

Spectral cues have a minor influence on the perception of horizontal sound location (Macpherson & Middlebrooks, 2002). Some unilaterally deaf patients, however, successfully use spectral cues for azimuth perception (Slattery & Middlebrooks, 1994; M. M. Van Wanrooij & Van Opstal, 2004), and it has been shown that spectral cues contribute to some extent to the accuracy of the azimuthal percept in normal-hearing participants (Musicant & Butler, 1984; Razavi, O'Neill, & Paige, 2005; M. M. Van Wanrooij & Van Opstal, 2007). In the present study, earmolds decreased performance on both the vertical and the horizontal plane, even though binaural cues remained intact. The reduction in horizontal performance strongly correlated with the reduction in vertical performance, but not with the amount of acoustical change in spectral cues caused by the molds, nor with changes in binaural DTF dissimilarity due to the molds. The outcome of the processing of spectral cues (for instance the elevation percept) must thus play a role in fine azimuthal sound localization. The notion of integrated processing of sound localization cues is supported by recent evidence for integrated processing of interaural time and level differences (Edmonds & Krumbholz, 2014). In addition, horizontal and vertical sound movement activates common cortical substrates (Pavani et al., 2002). M. M. V. Van Wanrooij and Van Opstal (2005) proposed that elevation perception is partly attributable to binaural interactions, in which the spectral-to-spatial mapping provided by each ear's spectral cues is weighted by the azimuthal perception. Incorrect elevation estimates might perhaps in turn affect the precision of the azimuthal perception. This was the case in our participants: larger disruption in elevation performance tended to result in larger disruption of horizontal performance.

4.4.3 Individual differences in adaptation to modified spectral cues

Participants' localization accuracy on the horizontal and vertical planes improved to varying degrees during the six days of exposure to modified spectral cues. Individual differences

ranged from full adaptation (accuracy indistinguishable from performance with free ears) to no adaptation (accuracy on day 6 indistinguishable from performance on day 0 after mold insertion). Participants did not fall into discernable groups, such as low and high performers, but rather formed a continuum of different adaptation rates between these extreme values. Large individual differences in adaptation to modified localization cues were previously reported (spectral cues: [Carlile & Blackman, 2013](#); [Hofman et al., 1998](#); [M. M. V. Van Wanrooij & Van Opstal, 2005](#); ITDs: [Javer & Schwarz, 1995](#); [Trapeau & Schönwiesner, 2015](#); supernormal cues: [Shinn-Cunningham et al., 1998](#)). Changes in sound localization induced by compressed vision also appear to differ substantially between individuals ([Zwiers, Van Opstal, & Paige, 2003](#)).

It is unclear at the moment whether these differences are attributable to individual differences in the capacity to adapt, and if so, which factors contribute to such differences. There are also differences in the acoustical effect of the earmolds. The initial disruption of the localization cues varied across participants in previous studies and in the present study. In the present study, the acoustical dissimilarity between DTFs with and without molds still explained a portion (26 %) of the differences in localization accuracy after six days of adaptation and training. Participants whose DTFs were disturbed more strongly showed less or slower improvement of localization accuracy over time. We used a fixed volume of silicone for each mold, which may have contributed to variations in the acoustical effect of the molds. Using a constant fraction of the concha volume as mold volume ([Carlile et al., 2014](#)) may reduce such variations. Finally, differences in lifestyle between participants may lead to differences in the amount and variety of daily sound exposure, in the amount of multisensory interactions, and so forth. Such differences may also account for some of the observed variability in adaptation. A better understanding of this variability may require a more tightly controlled DTF manipulation and adaptation environment than we were able to achieve in the present study.

4.4.4 Aftereffect

Most studies of adaptation to disrupted sound localization cues in humans reported no evidence of aftereffects upon removal of the manipulation (spectral cues: [Carlile & Blackman,](#)

2013; Hofman et al., 1998; M. M. V. Van Wanrooij & Van Opstal, 2005; ITDs: Javier & Schwarz, 1995; Trapeau & Schönwiesner, 2015). One study reported a very small aftereffect (monaural occlusion: Kumpik et al., 2010). Similarly, ferrets reared with a unilateral earplug that disturbed ILD information develop sound localization abilities comparable to controls (King et al., 2000) and showed no aftereffect after removal of the earplug. Barn owls also regain normal sound localization performance after several weeks of adapting to a unilateral earplug, but in stark contrast to the results in mammals, the birds mislocalized sounds in the opposite direction after removal of the plug (Knudsen, Esterly, & Knudsen, 1984). The aftereffect gradually diminished over several days. This difference between species might be explained by the different encoding mechanisms of horizontal sound direction (Ashida & Carr, 2011). The manner in which the Jeffress-type place code appears to be implemented in birds may be more susceptible to aftereffects. Localization in the visual system also operates on place codes and consistent aftereffects were observed following adaptation to prism glasses (Hay & Pick Jr., 1966; Singer & Day, 1966).

In agreement with these ideas and previous results, no aftereffect was observed in the present study. We initially hypothesized that participants were able to re-accommodate to their original spectral cues so quickly upon mold removal that the aftereffect is unobservable when averaging across all trials of our localization task. Contrary to this hypothesis, localization accuracy in the initial trials was normal, including the first sound exposure after mold removal. This was true even when unfamiliar sounds with unknown spectra were presented in these trials and when the original DTFs were simulated before mold removal and the participants thus did not know that stimuli corresponded to the acoustical situation without the molds. These results indicate that the new spectral-to-spatial mapping learned during adaptation did not override the pre-existing one, and that both mappings are available to the participant simultaneously. A possible conceptual mechanism for such an effect is a many-to-one mapping, in which several spectral profiles (DTFs) may become associated with one spatial location. This situation would require no switch between sets of DTFs and would thus not result in an aftereffect. Our measurements of trial-by-trial performance in the two localization tests one week and one month

after mold removal are also consistent with this notion. There was no evidence of performance improvements during the first few trials, i.e., participants did not require any exposure to the altered spectral cues to localize sounds. Such a mapping could for instance be implemented in the spectro-temporal correlation model proposed by [Hofman and Van Opstal \(1998\)](#) under the assumption that newly-learned DTFs are added as input to the “spectral correlator” stage.

4.4.5 Persistence

The perceptual capability of localizing sound with modified DTFs persisted for at least one month after removal of the molds. We re-measured localization accuracy one week and one month after removal of the molds and found that performance was significantly better than when the molds were first inserted. On the horizontal and vertical plane, more than 60 % of the adaptation was retained after one month without exposure to the modified spectral cues. This result is remarkable considering the short adaptation period of six days. Retention of adaptation to modified spectral cues after one week was previously demonstrated with much longer adaptation periods of 30 to 60 days ([Carlile & Blackman, 2013](#)). Even with the shorter adaptation period, our results at the one-week retest are comparable to those of [Carlile and Blackman \(2013\)](#): about 70 % of the adaptation was retained after one-week without molds in both studies.

A portion of the slight reduction in accuracy might have been caused by participants becoming less familiar with the testing procedure over the one- and four-week period. However, participants needed very little initial training with the task to achieve consistent performance, likely because head-turning in the frontal field is a natural response to sound and thus involved little to no procedural learning. Nevertheless, baseline tests with free ear were not performed during the persistence tests and we can thus not rule out an effect of procedural learning.

Cue persistence may be the result of a trade-off between the conflicting requirements of stability and adaptability of the mechanism. Except for accidents and modern interventions like assistive hearing devices, spectral cues change slowly, which perhaps explains the slow learning rates that we and others have observed for these cues. Such slow rates may cause some inertia in

the unlearning of spectral cues as well. The decay of cue memory may be appropriate to estimate the time constant of cue adaptability, because it might be less influenced by perceptual training and individual differences in sound environment than the learning of new cues. From our data, we estimated that this time constant is about 6 months, assuming exponential decay.

CHAPITRE 5

ARTICLE 4 : THE ENCODING OF SOUND SOURCE ELEVATION IN THE HUMAN AUDITORY

CORTEX

Régis Trapeau and Marc Schönwiesner

Publié dans *The Journal of Neuroscience*, 2530-17.

Abstract

Spatial hearing is a crucial capacity of the auditory system. While the encoding of horizontal sound direction has been extensively studied, very little is known about the representation of vertical sound direction in the auditory cortex. Using high-resolution functional magnetic resonance imaging, we measured voxel-wise sound elevation tuning curves in human auditory cortex and show that sound elevation is represented by broad tuning functions preferring lower elevations as well as secondary narrow tuning functions preferring individual elevation directions. We changed the ear shape of participants (male and female) with silicone molds for several days. This manipulation reduced or abolished the ability to discriminate sound elevation and flattened cortical tuning curves. Tuning curves recovered their original shape as participants adapted to the modified ears and regained elevation perception over time. These findings suggest that the elevation tuning observed in low-level auditory cortex did not arise from the physical features of the stimuli, but is contingent on experience with spectral cues and covaries with the change in perception. One explanation for this observation may be that the tuning in low-level auditory cortex underlies the subjective perception of sound elevation.

5.1 SIGNIFICANCE STATEMENT

This study addresses two fundamental questions about the brain representation of sensory stimuli: how the vertical spatial axis of auditory space is represented in the auditory cortex, and

whether low-level sensory cortex represents physical stimulus features or subjective perceptual attributes. Using high-resolution fMRI we show that vertical sound direction is represented by broad tuning functions preferring lower elevations as well as secondary narrow tuning functions preferring individual elevation directions. In addition, we demonstrate that the shape of these tuning functions is contingent on experience with spectral cues and covaries with the change in perception, which may indicate that the tuning functions in low-level auditory cortex underlie the perceived elevation of a sound source.

5.2 INTRODUCTION

Spatial hearing is the capacity of the auditory system to infer the location of a sound source from the complex acoustic signals that reach both ears. This ability is crucial for an efficient interaction with the environment, because it guides attention (Broadbent, 1954; Scharf, 1998) and improves the detection, segregation, and recognition of sounds (Bregman, 1994; Dirks & Wilson, 1969; Roman et al., 2003). Sound localization is achieved by extracting spatial cues from the acoustic signal that arise from the position and shape of the two ears. The separation of the two ears produces interaural time (ITD) and level differences (ILD), which enable sound localization on the horizontal plane. The direction-dependent filtering of the pinnae and the upper body generates spectral cues, which allow sound localization on the vertical plane as well as disambiguating sounds that originate from the front and the back of the listener (Blauert, 1997; Wightman & Kistler, 1989a).

The cortical encoding of the location of sound sources has been mostly studied on the horizontal plane. There is a large body of evidence from both animal and human models, that horizontal sound direction is represented in the auditory cortex by a rate code and two opponent neural populations, each one broadly tuned to the contralateral hemifield (Magezi & Krumbholz, 2010; Recanzone, Guard, Phan, & Su, 2000; Salminen et al., 2009; Stecker et al., 2005; Werner-Reiss & Groh, 2008). Note, however, that single-unit recordings in animals often show a greater diversity of spatial sensitivity than can be captured by far-field measures such as EEG and MEG

or by data pooled over many single-unit recordings (Bizley et al., 2007; Middlebrooks & Bremen, 2013; Mрсic-Flogel et al., 2005; T. M. Woods, Lopez, Long, Rahman, & Recanzone, 2006; Xu et al., 1998). These studies have shown sound elevation sensitivity in neurons in primary and higher-level auditory cortical areas, but it is still unclear how the auditory cortex represents elevation. One fMRI study in humans (Zhang, Zhang, Hu, & Zhang, 2015) presented stimuli at different elevations and computed an up vs. down contrast, but did not construct tuning curves. Their results suggested a non-topographic, distributed, representation of elevation in auditory cortex. Other neuroimaging attempts to study elevation coding were limited by poor behavioral discrimination of the different elevations presented through headphones (Fujiki, Riederer, Jousmäki, Mäkelä, & Hari, 2002; Lewald, Riederer, Lentz, & Meister, 2008).

We aimed to reveal the encoding of sound elevation in the human auditory cortex by extracting voxel-wise elevation tuning curves and to identify the relationship between these tuning curves and elevation perception by manipulating the acoustic cues for elevation perception. In a first fMRI session, sound stimuli recorded through the participants' unmodified ears were presented. These stimuli included the participants' native spectral cues and allowed us to measure encoding of sound elevation under natural conditions. In two subsequent sessions, stimuli carrying modified spectral cues were presented. Hofman et al. (1998) showed that adult humans can adapt to modified spectral cues. Audio-sensory-motor training accelerates this adaptation (Carlile et al., 2014; Parseihian & Katz, 2012; Trapeau, Aubrais, & Schönwiesner, 2016). We modified spectral cues by fitting silicone earmolds to the outer ears of our participants. Participants wore these earmolds for one week and received daily training to allow them to adapt to the modified cues. Stimuli recorded through the modified ears were presented in two identical fMRI sessions, one before and one after this adaptation. Any change in tuning between the two sessions is therefore due to the recovery of elevation perception achieved during adaptation, rather than physical features of the spectral cues, which allows us to directly link auditory cortex tuning to elevation perception.

5.3 METHODS

5.3.1 Participants

Sixteen volunteers took part in the experiment after having provided informed consent. The sample size was based on data from a previous study (Trapeau et al., 2016). One participant did not complete the experiment due to a computer error during the first scanning session. The 15 remaining participants (9 male, 6 female) were between 22 and 41 years of age (26 years on average $\pm$ 4.7 standard deviation). They were right-handed as assessed by a questionnaire adapted from the Edinburgh Handedness Inventory (R. C. Oldfield, 1971), had no history of hearing disorder or neurological disease, and had normal or corrected-to-normal vision. Participants had hearing thresholds of 15 *dB HL* or lower, for octave frequencies between 0.125 and 8 *kHz*. They had unblocked ear canals, as determined by a non-diagnostic otoscopy. The experimental procedures conformed to the World Medical Association’s Declaration of Helsinki and were approved by McGill University’s Research Ethics Board.

5.3.2 Procedure overview

Each participant completed three fMRI sessions. In each session, participants listened passively to individual binaural recordings of sounds emanating from different elevations. Stimuli played in fMRI session 1 were recorded from the participants’ unmodified ears and thus included their native spectral cues. Stimuli played in session 2 were recorded from participants’ ears with added silicone earmolds. These earmolds were applied to the conchae to modify spectral cues and consequently disrupt elevation perception. Session 3 was identical to session 2 but took place after participants had worn the earmolds for one week. To track performance and accelerate the adaptation process, participants performed daily sound localization tests and training sessions in the free-field and with earphones.

5.3.3 Apparatus

The molds were created by applying a fast-curing medical-grade silicon (SkinTite, Smooth-On, Macungie, PA, USA) in a tapered layer of about 3 to 7 *mm* thickness on the cymba conchae and the cavum conchae of the external ear, keeping the ear canal unobstructed (Figure 5.1.A).

Behavioral testing, training, and binaural recordings were conducted in a hemi-anechoic room ($2.5 \times 5.5 \times 2.5$ *m*). Participants were seated in a comfortable chair with a neck rest, located in the centre of a spherical array of loudspeakers (Orb Audio, New York, NY, USA) with radius of 90 *cm* (Figure 5.1.B). A LED was mounted at the center of each loudspeaker. In order to minimize location cues caused by differences in loudspeaker transfer functions, we measured these functions using a robotic arm that pointed a Brüel & Kjær $\frac{1}{2}$ -inch probe microphone at each loudspeaker from the center of the sphere and designed inverse finite impulse response filters to equalize amplitude and frequency responses across loudspeakers.

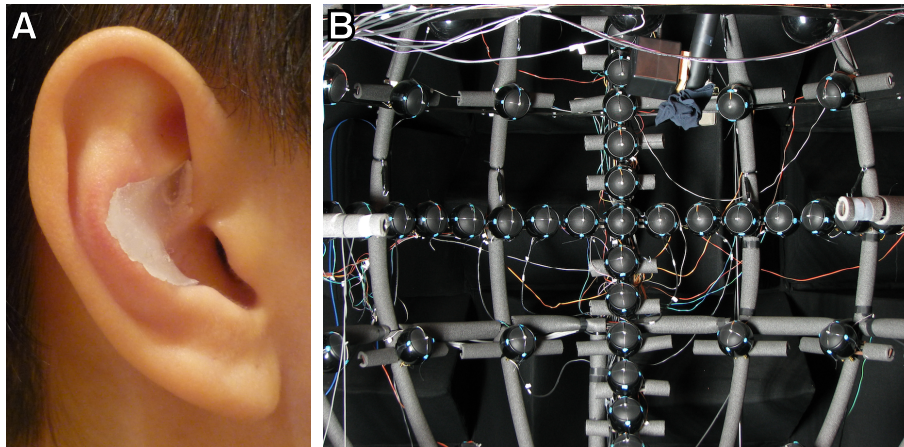


Figure 5.1 – Silicone mold and array of loudspeakers. A) Example of a silicone mold in a participant’s right ear. B) Spherical array of loudspeakers used for behavioral testing.

An acoustically transparent black curtain was placed in front of the loudspeakers to avoid visual capture. A small label indicating the central location (0° in azimuth and elevation) was visible on the curtain and served as a fixation mark.

A laser pointer and an electromagnetic head-tracking sensor (Polhemus Fastrak, Colchester, VT, USA) were attached to a headband worn by the participant. The laser light reflected from the

curtain served as visual feedback for the head position. Real-time head position and orientation were used to calculate azimuth and elevation of pointed directions.

Individual binaural recordings were taken with miniature microphones (Sonion 66AF31, Roskilde, Denmark). Sound stimuli of both the closed-field behavioral tasks and fMRI sessions were delivered through the same pair of S14 insert earphones (Sensimetrics, Maiden, MA, USA). The same pair of earphones was used to present stimuli in the closed-field behavioral task. Frequency response equalization of the insert earphones was achieved using the custom designed filters provided by Sensimetrics. No equalization was applied to the miniature microphones because their exact transfer function depends partly on the position in the ear canal and is difficult to correct. However, the probe microphones were audiological devices and had a sufficiently flat transfer function. The slight spectral coloration was not elevation dependent and did not hinder elevation perception.

Sound localization tests, binaural recordings, and stimulus presentations during fMRI sessions were controlled with custom MATLAB scripts (Mathworks, Natick, MA, USA). Stimuli and recordings were generated and processed digitally (48.8 kHz sampling rate, 24 bit amplitude resolution) using TDT System 3 hardware (Tucker Davis Technologies, Alachua, FL, USA).

5.3.4 Detailed procedure

Earmolds were fitted to the participants' ears during their first visit. Fitting was done in silence, and the molds were removed immediately after the silicone had cured (4–5 min). Individual binaural recordings were then acquired using miniature microphones placed 2 mm inside of the entrance to the blocked ear canal. Participants were asked not to move during the recordings and keep the dot of the head-mounted laser on the fixation mark. Recordings were first acquired from free ears. Without removing the microphones, the earmolds were then inserted for a second recording and removed immediately afterwards.

To minimize procedural learning during the experiment and to make sure that participants would have a good elevation perception of our stimuli before their first scan, participants completed three practice sessions. Each practice session took place on a different day, was done

without molds, and consisted of a sound localization run, a training run, and a localization run again for both free-field and earphone stimuli. In a previous study using the same equipment, no procedural learning was found for this localization task (Trapeau et al., 2016).

The first fMRI scanning session was then performed, in which binaural recordings from free ears were presented. The second scanning session was performed on the next day. Participants wore the molds during this session and binaural recordings taken with molds were presented. Earmolds were inserted in silence right before the participants entered the scanner and were removed immediately after the end of the scanning sequence.

From the next day onwards, participants wore the molds continuously. They were informed that the earmolds would “slightly modify their sound perception”, but did not know the specific effect of this modification. Once the molds were inserted, participants immediately performed two localization runs (one with earphones, the other in the free field), followed by two training runs (with earphones and free field). On each of the following seven days, participants completed a sound localization run, a training run, and a localization run again for both free-field and earphone stimuli.

Participants were scanned for a third time after seven days wearing the molds. The third scanning session was identical to the second one. Participants continued wearing the molds until they performed a last localization run with the molds, less than 24 *h* after the last scanning session.

5.3.5 Stimuli

All stimuli used in this experiment, including fMRI stimuli, consisted of broadband pulsed pink noise (25 *ms* pulses interleaved with 25 *ms* silence). Stimuli were presented at 55 dB SPL in all conditions to optimize elevation perception (Vliegen & Van Opstal, 2004).

During the free-field sound localization task and training, the pulsed pink noise was presented directly through the loudspeakers. During the closed-field sound localization task and training, as well as during fMRI sessions, stimuli were presented via the same pair of MRI-compatible earphones (Sensimetrics S14). Stimuli presented via earphones were individual

binaural recordings of the pulsed pink noise presented from loudspeakers at various locations. The locations of the stimuli presented in the different tasks are displayed in Figure 5.2.

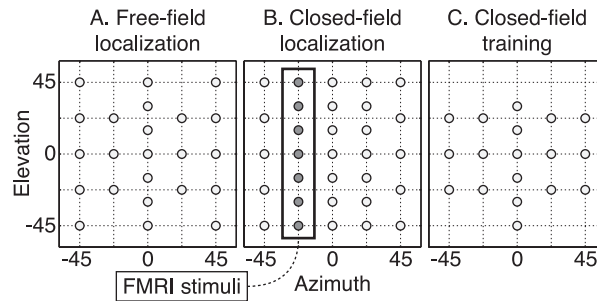


Figure 5.2 – Stimuli locations in the different tasks. A) Speaker arrangement during the free-field localization task. B) Stimulus locations during the closed-field localization task. C) Stimulus locations during the closed-field training.

Seven elevations were presented during the fMRI sessions (-45° to $+45^\circ$ in 15° steps). Because elevation perception is more accurate for azimuths away from the median plane (Carlile et al., 1997; Makous & Middlebrooks, 1990; Wightman & Kistler, 1989a), and for sounds in the left hemifield (Burke, Letsos, & Butler, 1994), our fMRI stimuli were recorded at 22.5° azimuth in the left hemifield (Figure 5.2.B). To ensure identical binaural cues at each elevation, we matched ITD across elevations by shifting the left and right channels to match the ITD measured at 0° elevation (maximum ITD correction: $\sim 80 \mu s$), and we matched ILD across elevations by normalizing broadband RMS amplitude of the left and right channel to the amplitudes measured at 0° elevation (maximum ILD correction: $\sim 2 dB$). The overall intensity at each elevation and for each set of recordings (free vs. molds) was also normalized to the same RMS amplitude to remove potential intensity cues (maximum overall intensity correction: $\sim 2 dB$). The same treatment was applied to the recordings with earmolds. To check that earmolds did not introduce physiologically implausible distortions in the stimulus spectra we plotted fMRI stimulus spectra and difference spectra for both the free ears and earmold conditions (Figure 5.3.A). Earmolds mostly changed the spectra for frequencies above 6 kHz, while lower frequencies were practically unaffected. Removing the common, non-directional, portion of this limited set of transfer functions helps to compare spectral differences across elevations

(Figure 5.3.B). See Trapeau et al. (2016) for a more detailed analysis of the acoustic properties of the earmolds.

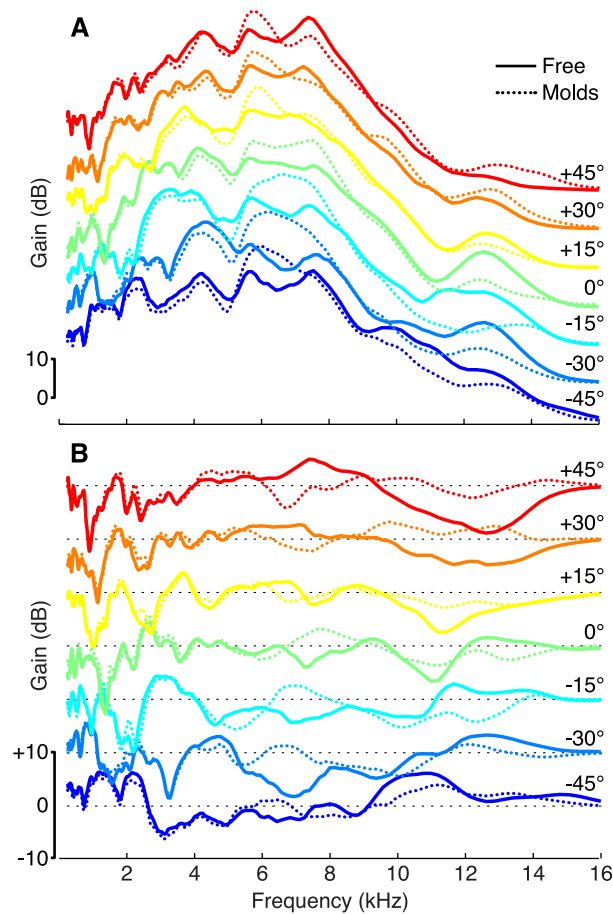


Figure 5.3 – Spectra and difference spectra of fMRI stimuli. Solid lines represent the free ears condition, dotted lines represent the earmolds condition. A) Spectra of the fMRI stimuli at each elevation for one randomly picked participant. Spectra have been computed from *KEMAR* recordings of the stimuli delivered through the pair of Sensimetrics S14 earphones used during the experiment. B) Difference spectra at each elevation for the same participant. Each transfer function was divided by the average of the transfer functions across elevations. Black dashed lines represent a 0 dB gain. Earmolds mostly changed the spectra for frequencies above 6 kHz, while lower frequencies were practically unaffected.

The duration of the stimulus used in both sound localization tasks (free-field and closed-field), as well as in the closed-field training task, was 225 ms, i.e. 5 pulses. It was 125 ms (3 pulses) in the free-field training task and 4350 ms (87 pulses) during the fMRI sessions.

5.3.6 Localization tasks

5.3.6.1 *Free-field localization task*

The free-field localization task was similar to the one described by [Trapeau et al. \(2016\)](#). In each trial a 225 ms pulsed pink noise was presented from one of 23 loudspeakers covering directions from -45° to $+45^\circ$ in azimuth and elevation (see Figure 5.2.A for speaker locations). Each sound direction was repeated five times during a run, in pseudorandom order. Participants responded by turning their head and thus pointing the head-mounted laser toward the perceived location. At the beginning of a run, participants sat so that the head was centered in the loudspeaker array and the laser dot pointed to the central location. This initial head position was recorded and the participant had to return to this position with a tolerance of 2 cm in head location and 2° in pointed angle before each trial. If the head was correctly placed when the button was pressed, a stimulus was played from one of the speakers. If the head was misplaced, a brief warning tone was played from a speaker located above the participant's head (Az: 0° , El: 82.5°).

5.3.6.2 *Closed-field localization task*

The procedure in the closed-field localization task was the same as in free-field, except that stimuli were presented via MRI-compatible earphones instead of loudspeakers. Stimuli were individual binaural recordings of the free-field stimulus presented from 31 different locations (see Figure 5.2.B for stimulus locations).

5.3.6.3 *Statistical analysis*

Vertical localization performance was quantified by the elevation gain (EG), defined as the slope of a linear regression line of perceived versus physical elevations ([Hofman et al., 1998](#)). Perfect localization corresponds to an EG of 1, while random elevation responses result in an EG of 0.

Behavioral EG was calculated on all locations except when reporting the closed-field EG

corresponding to a scanning session, where it was calculated using only the locations presented in the scanner, i.e. -45° to $+45^\circ$ elevations at 22.5° azimuth in the left hemifield (Figure 5.2.B). No behavioral EG was measured in the scanner; these measurements were taken before or after the scanning sessions, as indicated in the timeline in Figure 5.4.D. The closed-field EG reported for the first fMRI session was thus computed from the data collected during practice day 3 in Figure 5.4.D (using trials with stimuli at 22.5° azimuth). The closed-field EG reported for the second fMRI session (in which binaural recordings with molds were presented) was computed from the data collected during adaptation day 0, and the closed-field EG reported for the third fMRI session was computed as weighted average of data collected during adaptation days 7 and 8 (because comparable data was available from before and after the session). The average was weighted by the time separating the two localization tests from the scanning session. Throughout the analysis, paired comparisons were statistically assessed using Wilcoxon signed-rank tests, p-values smaller than 0.05 were regarded as significant.

5.3.7 Training tasks

Both training tasks involved auditory and sensory-motor interactions, which has been shown to accelerate the adaptation process when compared to tasks that comprise visual feedback only (Carlile et al., 2014).

5.3.7.1 Free-field training task

The free-field training sessions were 10 minutes long and the procedure was inspired by Parseihian and Katz (2012). A continuous train of stimuli (each stimulus was 125 ms and consisted of 3 pulses of pink noise) was presented from a random location in the frontal field (between $\pm 45^\circ$ elevation and $\pm 90^\circ$ azimuth, 61 loudspeakers). The interstimulus interval (ISI) depended on the angular distance between the sound direction and the participant's head direction (measured continuously); the smaller the difference, the shorter the ISI – an effect similar to a Geiger counter. Participants were instructed to point their head toward the sound source location as fast as possible (no button needed to be pressed), which allowed active

exploration of the auditory space and aids adaptation (Carlile et al., 2014; Parseihian & Katz, 2012). Once the participants kept their head in an area of 6° radius around the target location for 500 ms, the stimulus was considered found and the location was switched to a random new loudspeaker with the constraint that the new location was at least 45° away from the previous one. Participants were instructed to find as many locations as possible during the training run.

5.3.7.2 Closed-field training task

The closed-field training consisted of paired auditory and visual stimuli. The auditory stimuli were identical to the ones used in the closed-field localization task (one stimulus consisted of 5 pulses of pink noise and had a total duration of 225 ms), and were also presented through MRI-compatible earphones. A trial consisted in the presentation of five such stimuli (25 pulses in total) from a given location, each stimulus being separated by a 275 ms ISI. Trials were separated by 1 s pauses without any stimulation. Each trial corresponded to a pseudorandom location in the frontal field (Figure 5.2.C). The visual stimuli were LEDs blinking in synchrony with the acoustic stimuli and at the location corresponding to the sound source in the binaural recording in each stimulus presentation. Because binaural recordings were taken while participant's head pointed at the central location, participants were asked to keep their head (and the laser dot) pointed to this position during stimulus presentation. The task was automatically paused when the head position or orientation was incorrect. To introduce sensory-motor interactions, participants were asked to shift their gaze toward the blinking LED.

To ensure that participants attended the sound stimuli, they were asked to detect decreases in sound intensity that occurred infrequently (28 %) and randomly in the third, fourth, or fifth noise burst (Recanzone, 1998). Participants indicated a detection by pressing a button, which paused the task. To add a sensory-motor component to the task, participants pointed their head toward the position of the detected deviant sound (no sound was played during head motion). Once their head direction was held for 500 ms in a 6° radius area around the target location, the target was considered found and feedback was given to the participant by briefly blinking the

LED at the target location. The task was restarted when the head position returned at the central location.

During one training run, each location was presented seven times and a decrease in intensity occurred twice per location. One run took about 8 minutes to complete. The amount of decrease in intensity was adjusted by a 1-up 2-down staircase procedure so that participants achieved a hit rate of 70.71 %. The initial intensity decrement was 10 *dB*, and participants typically ended between 2 and 4 *dB*.

5.3.8 FMRI data acquisition

5.3.8.1 *Imaging protocol*

Imaging was performed on a 3 *T* MRI scanner (Trio, Siemens Healthcare, Erlangen, Germany), equipped with a 12 channel matrix head-coil. For each participant and each session, a high-resolution ($1 \times 1 \times 1 \text{ mm}$) whole-brain T1-weighted structural scan (magnetization-prepared rapid gradient echo sequence) was obtained for anatomical registration. Functional imaging was performed using an echo-planar imaging sequence and sparse sampling (Hall et al., 1999) to minimize the effect of scanner noise artifacts (gradient echo; repetition time (TR) = 8.4 s, acquisition time = 1 s, echo time = 36 ms; flip angle = 90°). Each functional volume comprised 13 slices with an in-plane resolution of $1.5 \times 1.5 \text{ mm}$ and a thickness of 2.5 mm (field of view = 192 mm). The voxel volume was thus only 5.625 mm^3 , about a fifth of a more commonly used $3 \times 3 \times 3 \text{ mm}$ voxel (27 mm^3). The slices were oriented parallel to the average angle of left and right lateral sulci (measured on the structural scan) to fully cover the superior temporal plane in both hemispheres. As a result, the functional volumes included Heschl's gyrus, planum temporale, planum polare, and the superior temporal gyrus and sulcus.

5.3.8.2 Procedure

During the fMRI measurements, the participants listened to the sound stimuli while watching a silent nature documentary. Participants were instructed to be mindful of the sound stimulation, but no task was performed in the scanner.

Each session was divided into two runs of 154 acquisitions including one initial dummy acquisition, for a total of 308 acquisitions per session and per subject. Each acquisition was directly preceded by either a 4350 *ms* long stimulus (positioned in the TR of 8.4 *s* so that the maximum response is measured by the subsequent volume acquisition) or silence. In a session, each of the 7 sound elevations was presented 34 times, interleaved with 68 silent trials in total. Spatial stimuli and silent trials were presented in a pseudorandom order, so that the angular difference between any two subsequent stimuli was greater than 30° and silent trials were not presented in direct succession. Stimulus sequences in the three scanning sessions were identical for each participant, but differed across participants.

5.3.9 FMRI data analysis

The functional data were preprocessed using the MINC software package (McConnel Brain Imaging Center, Montreal Neurological Institute, Montreal, Canada). Preprocessing included head motion correction and spatial smoothing with a 3 *mm* full-width half-maximum 3-dimensional Gaussian kernel. Each run was linearly registered to the structural scan of the first session and to the ICBM-152 template for group averaging. Group averaging was only used for visualization of the approximate location of active voxels at the group level, and to estimate tuning functions for regions outside of the auditory cortex. General linear model estimates of responses to sound stimuli vs. silence were computed using fMRISTAT (Worsley et al., 2002). A mask of sound responsive voxels (all sounds vs. silence and each elevation vs. silence) in each session was computed for each participant and all further analysis was restricted to voxels in the conjunction of the masks from the three sessions. Voxel significance was assessed by thresholding T-maps at $p < 0.05$, corrected for multiple comparisons using Gaussian Random Field Theory (Worsley

et al., 1996). Because our stimuli were presented in the left hemifield and due to the mostly contralateral representation of auditory space (Krubholz, Schönwiesner, Cramon, et al., 2005; Palomäki et al., 2005; Pavani et al., 2002; Trapeau & Schönwiesner, 2015; D. L. Woods et al., 2009), we analyzed data only in the right hemisphere.

We computed voxel-wise elevation tuning curves by plotting the response size in each elevation vs. silence contrast. Because we were interested in the shape rather than the overall magnitude of the tuning curves, tuning curves were standardized by subtracting the mean and dividing by the standard deviation (z-scores). We determined the elevation preference of each voxel by computing three measures: the centre of gravity of the voxel tuning curve (COG), the maximum of the tuning curve and the slope of the linear regression of standardized activation vs. standardized elevation (analogous to the behavioral EG, we refer to this slope as the tuning curve EG). The COG is a simple noise-robust non-parametric method (Schönwiesner, Dechent, Voit, Petkov, & Krubholz, 2015), computed as follows:

$$COG = \frac{1}{n} \sum_{i=1}^n \frac{r_i}{n \cdot \bar{r}} \cdot e_i, \quad (5.1)$$

where n is the number of presented elevations (7 in the current experiment), e_i is the elevation value, and r_i is the respective response amplitudes in percent signal change. $\bar{r}$ is the average response amplitude across elevations. The maximum of the tuning curve is less robust to noise, but avoids compression at the highest and lowest elevations. It is therefore more appropriate when plotting sharp tuning functions across elevations and, after verifying that both methods yield similar results, the maximum was used in the analysis of sharp tuning. To avoid selection bias in the mean tuning curves, we selected voxels according to their COG, maximum or EG in one of the two functional runs and extracted tuning curves from the other run (and inversely; cross-validation).

For all neuroimaging data averaged across participants, standard error was obtained by bootstrapping the data 10000 times, using each participant's mean response.

FMRI measures a hemodynamic response, which fortunately is very tightly coupled to neural

responses. The BOLD signal correlates directly with the sustained local field potential (and to a lesser degree with spiking activity) as recorded in-vivo with microelectrodes during fMRI measurements (Ekstrom, 2010; Logothetis, Pauls, Augath, Trinath, & Oeltermann, 2001). This correlation has also been shown in low-level auditory cortex (Mukamel et al., 2005; Nir et al., 2007) and holds for awake behaving primates viewing visual stimuli (Goense & Logothetis, 2008). Note that sustained local field potentials reflect mostly intracortical processing. These data are the basis for inferences on neural tuning functions from fMRI recordings.

5.4 RESULTS

5.4.1 Behavioral results

Participants had normal elevation perception with a mean elevation gain (EG) of 0.75 ± 0.05 (mean $\pm$ standard error) in the first free-field test. Figures 5.4.A–C shows behavioral results from a representative participant. The three practice sessions slightly increased free-field EG (0.81 ± 0.04). The practice sessions were especially important for the closed-field condition, because participants required training to perform as well in the less natural closed-field as in the free-field. The average closed-field EG in the first practice session was 0.42 ± 0.06 , which improved to free-field performance by the third session (0.81 ± 0.05). The poor results of the first session were mainly due to reduced externalization of the stimuli (as reported by the participants). The first practice session drastically helped to achieve an externalized perception of the stimuli and practically all of the increase in performance happened in the first practice session. Performance was stable from the second to the third practice session (Figure 5.4.D). According to participant reports, stimuli were mainly heard as coming from the earphones before the first training and became clearly externalized during training. The large and fast increase in performance is likely due to the acquisition of an externalized percept during the first training.

The insertion of the earmolds produced a marked reduction in the free-field EG (0.38 ± 0.05 ; one-tailed Wilcoxon signed-rank test: $p = 3 \times 10^{-5}$). The same reduced EG (0.34 ± 0.06)

was observed in the closed-field when presenting individual binaural recordings taken with earmolds. After seven days of wearing the molds continuously and undergoing short daily training sessions, the mean EG recovered to 0.6 ± 0.06 ($p = 4 \times 10^{-4}$) in the free-field and to 0.52 ± 0.05 ($p = 1.5 \times 10^{-4}$) in the closed field (Figure 5.4.D). Individual differences in adaptation were large and ranged from full recovery to no adaptation.

Because in the fMRI scanner, sounds were presented via earphones from a subset of locations, we also calculated EG values in the closed field based only on those locations (seven vertical directions at 22.5° azimuth in the left hemifield). These EG values were comparable to the free- and closed-field values in the first session (0.86 ± 0.06) and showed a similar reduction (0.42 ± 0.07 ; $p = 2 \times 10^{-4}$) and recovery (0.56 ± 0.07 ; $p = 0.0177$) in the second and third scanning session (red diamonds in Figure 5.4.D).

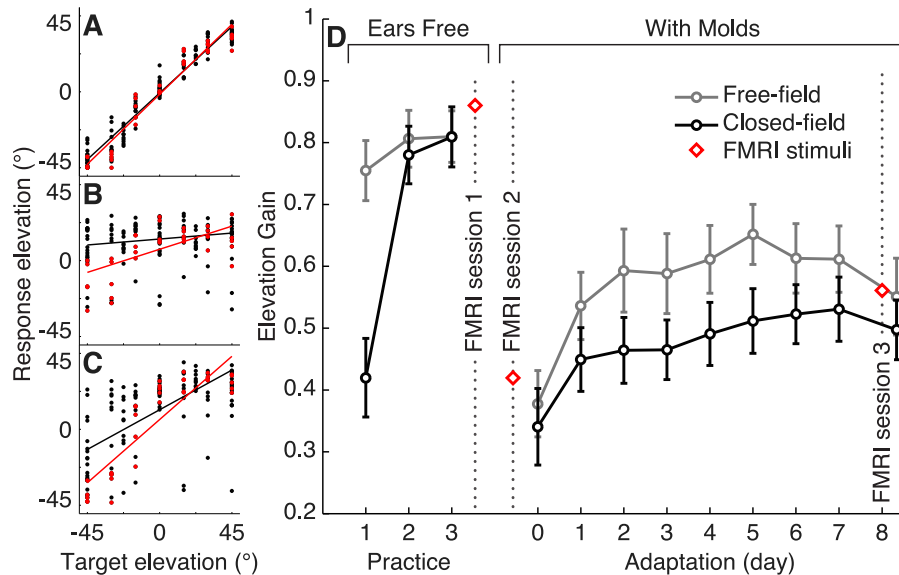


Figure 5.4 – Sound localization performance. Left panels: closed-field sound localization performance of a representative participant. Each dot represents one response (perceived elevation vs. target elevation). Red dots highlight targets at 22.5° azimuth in the left hemifield (locations presented in the scanner). Black and red lines are linear fits of the responses for all locations and for locations at 22.5° azimuth in the left hemifield, respectively. The EG is the slope of these regression lines. A) Baseline performance with free ears (practice day 3). B) First performance with molds (pre-adaptation). C) Last performance with molds (post-adaptation). D) Time course of EG averaged across all participants ($n = 15$), shown for the practice period (ears free), and for the adaptation period (with molds). Dashed lines represent the free-field EG, solid lines represent the closed-field EG, error bars indicate standard error. EG was calculated from all target locations. Red diamond symbols represent the closed-field EG calculated using only the locations corresponding to the fMRI stimuli, i.e. seven elevations from -45° to $+45^\circ$ at 22.5° azimuth in the left hemifield.

5.4.2 Elevation tuning curves with free ears

Sound stimuli evoked significant activation in the auditory cortex of all participants in the three scanning sessions. We computed elevation tuning curves for all sound responsive voxels and observed signal changes that co-varied with elevation. The average tuning curve across voxels and participants showed a decrease of activation level with increasing elevation (Figure 5.5.A). This decrease was observed in all participants, although the shape and slope of the mean tuning curves differed between participants. The same response pattern was found in both hemispheres, but as expected, the activation was stronger and more widespread in the right hemisphere, contralateral to the sound sources, and we thus only report data from that hemisphere.

The average tuning curve indicated that the majority of voxels may represent increasing elevation as an approximately linear decrease in activity. To determine whether there were any subpopulations of voxels with different tuning, such as narrow tuning curves selective for different sound directions, we plotted the mean tuning curve across all voxels that preferred a given elevation (quantified as the COG and the maximum of the tuning curve) or showed similar EG. Tuning curves were cross-validated to avoid selection bias. All resulting tuning curves had broad negative slopes (Figures 5.5.B–M). Cross-validation showed that single-voxel elevation preferences to mid- or high-elevations were not repeatable across runs and thus due to noise. This is also supported by the distributions of the COG and EG across all sound responsive voxels and participants, which exhibited only one distinct population, centered (mode) at -15.12° COG and -0.84 EG (Figures 5.5.N,O). We also applied a dimensionality-reduction technique that excels at clustering high-dimensional data (t-SNE, [Van der Maaten & Hinton, 2008](#)). While this procedure accurately separated tuning curves of voxels tuned to the left and right hemifield from a previous study on horizontal sound directions, it did not identify distinct subgroups in the present elevation tuning curves.

The shape of the voxel-wise elevation tuning curves could be related to elevation perception, or to acoustical properties of the stimuli that change with elevation, or to a mixture of both. To

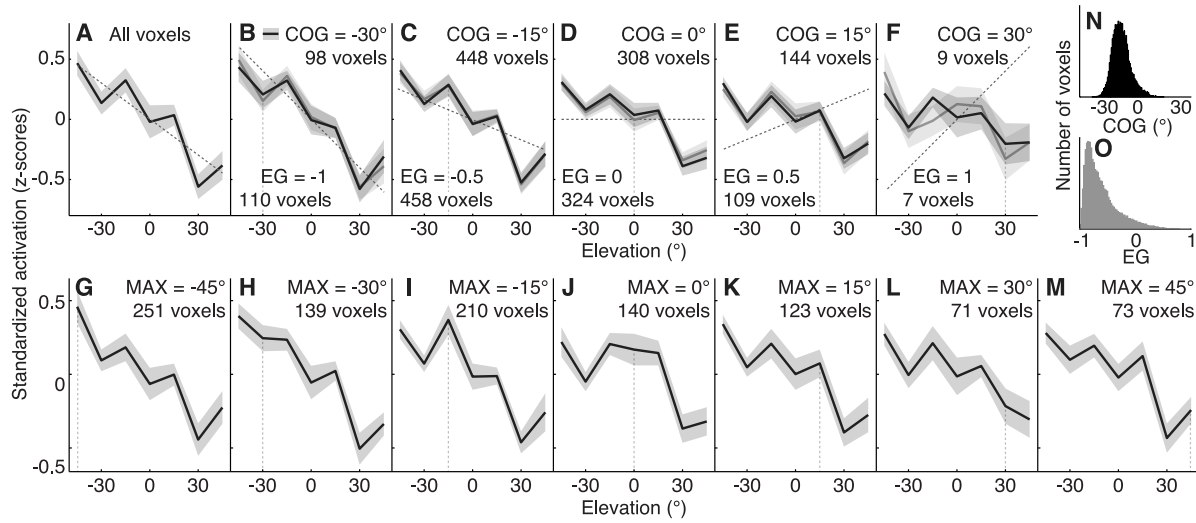


Figure 5.5 – Mean elevation tuning curve of all sound-responsive voxels in the right hemisphere. A–M) Mean elevation tuning curves. Shaded areas indicate the standard error across participants ($n = 15$, except panel F: $n = 14$). A) Mean elevation tuning curve of all sound-responsive voxels (solid black line) and a linear regression of measured activity against stimulus direction (dashed line). The slope of the regression line corresponds to the elevation gain (EG) of the mean tuning curve. B–M) Mean tuning curves (cross-validated) of voxels binned according to their tuning curve COG, EG or maximum. The mean number of voxels across participants selected by cross-validation for COG, EG and maximum is listed in each plot. In each bin, corresponding tuning curves were found in all participants, except panel F (COG = 30°; EG = 1), in which corresponding tuning curves were found in 14/15 participants. In panels B–F, mean tuning curves of voxels by COG are shown in black and mean tuning curves of voxels selected by EG are shown in grey. In panels G–M, mean tuning curves are shown in black. Cross-validated activity did not peak at selected COGs, maxima (vertical dotted lines) or follow the selected EGs (angled dotted lines), but rather decreased with increasing elevation in all voxels. N) Distribution of tuning curve COG. O) Distribution of tuning curve EG.

address this question we first quantified systematic changes with elevation of spectral amplitude in different frequency bands of the stimuli presented in the MRI scanner. We computed a linear regression coefficient between elevation and the mean spectral amplitude per octave band, which results in a frequency-dependent “acoustic elevation gain” (Figure 5.6.A). We found strong positive correlations in octave bands between 4 and 8 kHz. In these bands, spectral amplitude covaried linearly with elevation. The earmolds slightly shifted these positive correlations toward lower frequencies, but did not reduce them. Negative covariations at low frequencies (< 1 kHz) probably reflect elevation cues from torso and head and was unaffected by the earmolds.

Such acoustic changes might have contributed to the observed voxel tuning functions. To address this important caveat we conducted a control experiment. Two pairs of participants

performed an additional fMRI session in which we swapped the binaural recordings for each pair of participants. This allowed us to compare functional responses of a same group of participants to a same set of stimuli, in two different conditions: when stimuli matched participants spectral cues and led to full elevation perception (own ears); when stimuli were swapped between participants and elevation perception was reduced (swapped ears). To avoid any learning of the swapped spectral cues prior to the scanning session, behavioral performance with swapped stimuli was assessed after the scanning session. All control participants showed reduced elevation perception with swapped stimuli ($EG = 0.38 \pm 0.04$). Their performance was higher for locations corresponding to stimuli presented in the scanner ($EG = 0.52 \pm 0.12$), than for other locations ($EG = 0.34 \pm 0.04$). One participant in particular had a very large EG of 0.82 for the locations presented in the scanner compared to other locations ($EG = 0.38$) and was excluded from the analysis. The mean EG for the remaining participants at the locations presented in the scanner was 0.42 ± 0.11 . With stimuli corresponding to their own ears, elevation tuning curves of the control participants, clearly displayed a linear decrease with increasing elevation (Figure 5.6.B). With swapped stimuli, the shape of the tuning curves did not exhibit this pattern (Figure 5.6.C). Mean tuning curves were no longer monotonic and were much flatter (tuning curve EG with own ears = -0.56 ± 0.03 ; tuning curve EG with swapped ears = -0.27 ± 0.07).

The better behavioral performance at locations previously heard in the scanner suggests that some adaptation occurred during the 45 minutes of passive listening in the scanner. We analyzed the time course of the tuning curve shape during the scanning session by extracting average tuning curves across auditory cortex in a moving time window of about 10 minutes (7 trials per elevation). Tuning curve EG increased over the course of the experiment from essentially zero at the beginning to around 0.3 for the final stimulus presentations (Figure 5.7). This suggests that the remaining amount of tuning curve EG observed with the swapped recordings is due to adaptation during the measurement, and that the tuning curve EG is close to zero when elevation perception is entirely absent (i.e. at the beginning of the scanning session).

Because the prevalent shape of tuning in auditory cortex appeared to be a linear decrease with increasing elevation, we modeled a linear effect of elevation to highlight areas of auditory cortex

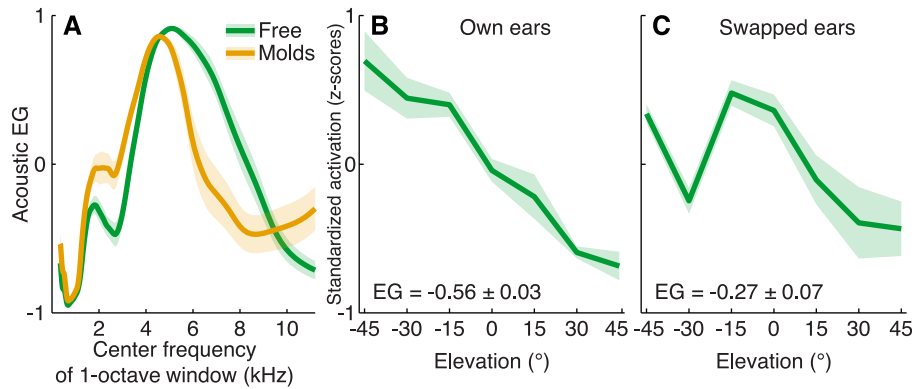


Figure 5.6 – Acoustic elevation gain and mean elevation tuning curves in control experiment. A) Frequency-dependent acoustic elevation gain. Each point represents the mean slope across participants ($n = 15$, shaded areas indicate the standard error across participants) of the linear regression between standardized elevations and the standardized spectral amplitudes of the difference spectrum at each elevation computed in a 1-octave window. The acoustic EG was computed separately for recordings with free ears (green line) and earmolds (orange line). Strong positive correlations are visible between 4 and 8 kHz. B–C) Mean elevation tuning curves of all sound-responsive voxels in the right hemisphere, averaged across the three control participants. B) For stimuli corresponding to the participants own ears, the mean tuning curve displayed a linear decrease with increasing elevation. C) Mean tuning curves were no longer monotonic and much flatter when stimuli corresponding to someone else’s ears were presented.

exhibiting that tuning. We selected voxels that exhibited a significantly positive or negative effect in this contrast in at least one of the three sessions. At this stage of the analysis, one participant had to be excluded because of a malfunction of the earphones during the second scanning session. Ten out of 14 participants displayed significant linear covariation of single-voxel activity and sound elevation (Figure 5.8).

Activation decreased with increasing elevation in all significantly active voxels in this contrast. We found no voxels whose activity increased with elevation. The mean tuning of voxels in this contrast had the same shape as the one observed in all sound-responsive voxels, but was less noisy. Interindividual variability of the mapping of elevation preference in the auditory cortex was large, which was expected by the great anatomical variability of auditory cortices across individuals (Rademacher et al., 2001) and the functional variability reported for other types of mappings (for instance Griffiths & Hall, 2012; Humphries, Liebenthal, & Binder, 2010; Trapeau & Schönwiesner, 2015). There is a known relation between frequency and elevation such that sounds at higher elevations tend to contain more power at high frequencies (Blauert, 1997; Parise, Knorre, & Ernst, 2014). We tested whether elevation preference correlated with frequency

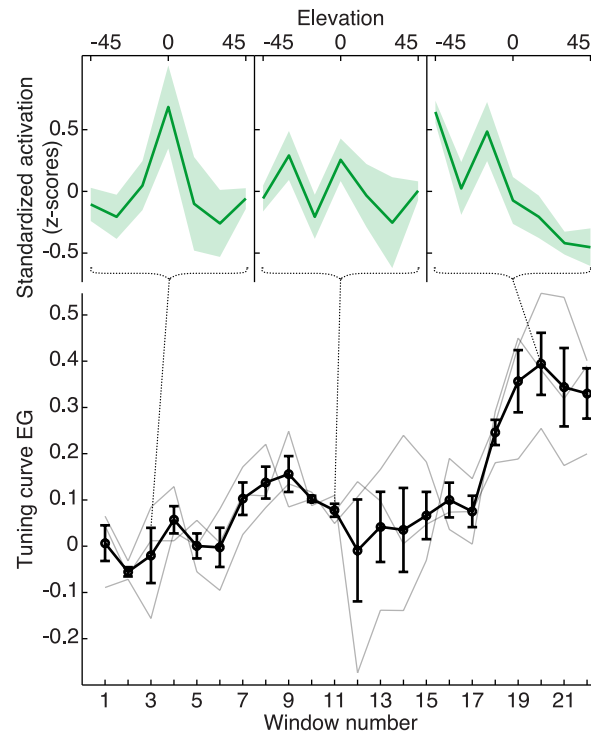


Figure 5.7 – Evolution of the mean elevation tuning curve and of the tuning curve EG with swapped stimuli. Mean tuning curves were extracted from the functional data of the control experiment in a moving time window of about 10 minutes (7 trials per elevation), yielding 22 windows for the 45-minute scanning session. The top panels show the mean tuning curves at window number 3, 11, and 20. The lower panel shows the time course of the individual (grey lines) and mean (bold black line) tuning curve EG during the session. Tuning curve EG increased over the course of the experiment from essentially zero at the beginning to around 0.3 for the final stimulus presentations.

preference by comparing individual elevation preference maps and a published average map for frequency preference (Schönwiesner et al., 2015). The mean correlation between individual maps of elevation COG and the average map of frequency COG did not differ significantly from zero ($p = 0.56$, Wilcoxon signed rank test of Spearman correlation coefficients across participants). In addition, the mean frequency preference of voxels that showed a significant covariation with elevation (Figure 5.8, white outline) did not differ significantly from the frequency preference of other sound-responsive voxels (sounds-vs-silence contrast, colored voxels in Figure 5.8 excluding white outline; $p = 0.12$, Wilcoxon signed rank test of mean frequency tuning of the two groups of voxels across participants).

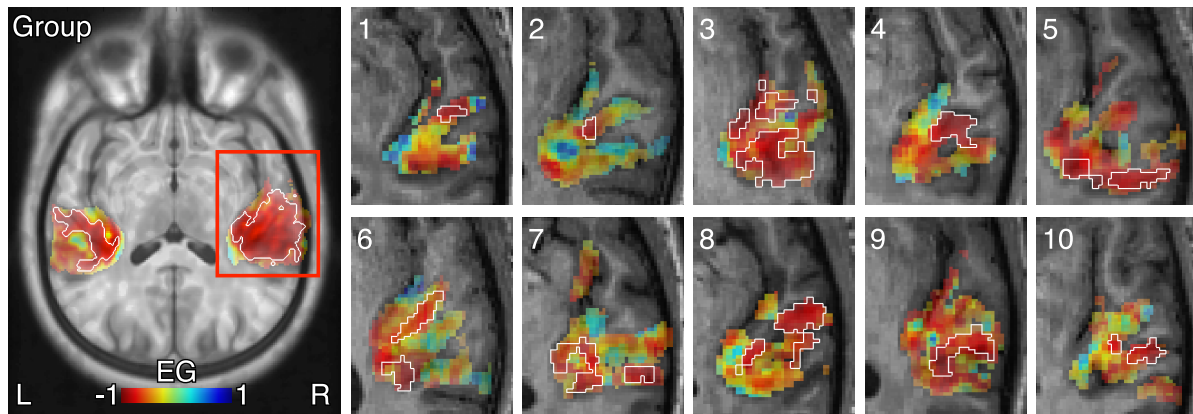


Figure 5.8 – Distribution of tuning curve elevation gain across auditory cortex. Fixed-effects group average data (labeled Group) and individual data (labeled 1–10) are shown. Slices are oriented to run parallel to the lateral sulcus (see red line on insert schematic brain). The color code indicates tuning curve EG of sound-responsive voxels. The white outlines delineate voxels that showed a significant covariation with elevation in at least one of the three sessions.

5.4.3 Effect of the molds and adaptation on elevation tuning

The above results demonstrate a relationship between physical sound position on the vertical axis and auditory cortex activity. Our manipulation of spectral cues and elevation perception allowed us to go one step further and to relate the shape of these tuning curves directly to elevation perception.

If voxel activity in low-level auditory cortex relates to the perception of sound source elevation, then we predict that our measure of elevation perception (behavioral EG) should correlate with the equivalent brain measure (tuning curve EG) across all participants and sessions. Behavioral EG correlated significantly with tuning curve EG (in voxels that showed a significant covariation with elevation – the contrast used in the previous section) across individuals and sessions (Figure 5.9.A; spearman correlation coefficient: $R = -0.44$, $p = 0.019$), indicating that changes in tuning curves covary with changes in reported stimulus elevation across sessions. We did not find significant correlations between these measures within sessions or between behavioral adaptation and individual differences in neural tuning, perhaps due to insufficient within-session variance or statistical power.

On the basis of the measured tuning curves, we hypothesized that auditory cortex encodes

the perception of vertical sound location by a decrease in activity with increasing elevation. Under this hypothesis, we predict less informative (flatter and/or noisier) tuning curves when elevation perception is reduced by the earmolds, and a recovery of the tuning curve shape when participants adapt to the earmolds and regain elevation perception. Steepness of the tuning curves was quantified by their EG, and noisiness was quantified by the R^2 of the EG regression line, in voxels that showed a significant covariation with elevation. Results showed a significant flattening of the participants' mean tuning curve from the first to the second session (mean tuning curve EG from session 1 to session 2: 0.73 ± 0.03 to 0.49 ± 0.06 ; one-tailed Wilcoxon signed-rank test, $p = 0.0098$, $n = 10$). Noisiness also increased from session 1 to 2, as shown by the reduction of R^2 (0.56 ± 0.04 for session 1, 0.33 ± 0.03 for session 2; $p = 0.0068$, $n = 10$).

To observe the effects of behavioral adaptation on the shape of elevation tuning curves, we then examined differences between sessions 2 and 3. Two of our participants showed no behavioral adaptation to the earmolds and were therefore excluded from the following comparison (the behavioral EG across sessions of the remaining participants is shown Figure 5.9.B). From session 2 to session 3, tuning curves changed in the predicted direction, i.e. the opposite direction of the changes observed between sessions 1 and 2. Tuning curve slope increased (mean tuning curve EG from session 2 to session 3: 0.5 ± 0.07 to 0.67 ± 0.03 ; $p = 0.0273$, $n = 8$; Figure 5.9.C) and the curves were less noisy in session 3 (R^2 from session 2 to session 3: 0.33 ± 0.04 to 0.46 ± 0.04 ; $p = 0.0195$, $n = 8$; Figure 5.9.D). These differences between the sessions are also visible in the mean tuning curves across participants. The mean tuning curve with molds (session 2) was flatter and less monotonic than the one with native ears (session 1), and recovered after adaptation to the molds (session 3, Figure 5.9.E).

5.4.4 Secondary narrow tuning to elevation

Other features of the response, apart from the overall slope of the tuning functions, may change with sound elevation. To identify such changes, we subtracted the mean tuning function of each participant from all voxel tuning functions and examined the remainder separately in voxels that responded maximally to each of the seven elevations. Without subtracting

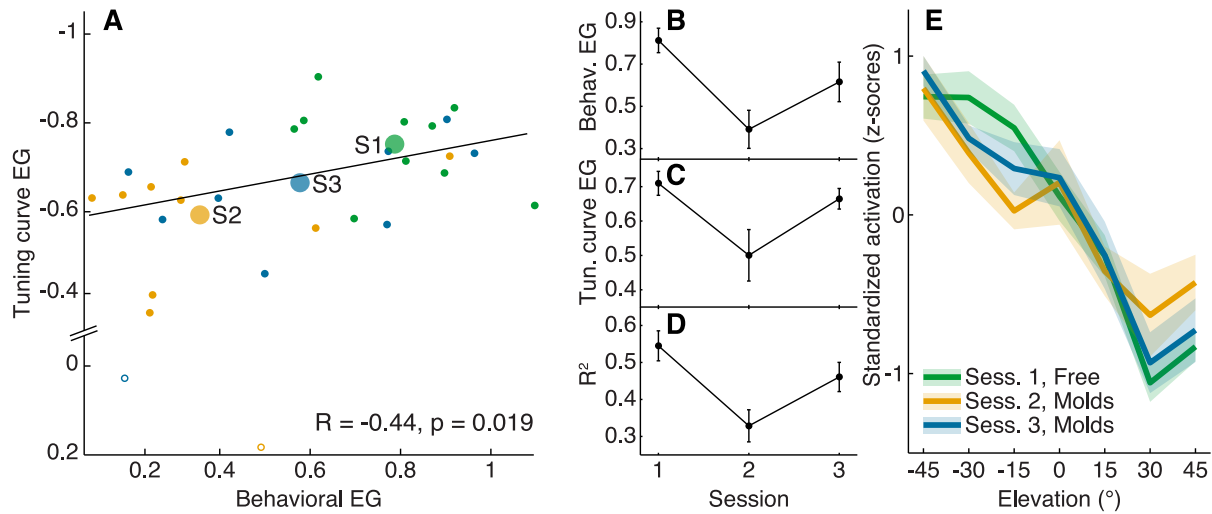


Figure 5.9 – Behavioral and tuning curve results across sessions. A) Correlation between behavioral and tuning curve EG across all participants and sessions. The Spearman correlation coefficient (lower right inset) was calculated after excluding two outlier data points (unfilled dots; coefficient including outliers: $R = -0.48$, $p = 0.008$). Black dots indicate results of session 1, light grey dots indicate results of session 2, and dark grey dots indicate results of session 3. Large dots represent the mean of each session (outliers excluded). Data in panels A–D shows clear covariation of reported elevation perception with mean voxel tuning curve slope. The position of session 3 about halfway in between sessions 1 and 2 in panel D suggests a proportional recovery of elevation perception and tuning curve shape during adaptation. B) Behavioral EG across participants who showed adaptation to the earmolds ($n = 8$). Error bars indicate the standard error across participants. C) Mean tuning curve EG across the same participants. The absolute EG dropped in session 2 (shallower tuning curves) and partially recovered in session 3. D) Mean R^2 of the EG regression line across participants. R^2 dropped in session 2 (tuning curves noisier) and partially recovered in session 3. E) Mean elevation tuning curves. Tuning curves showed a steep decrease towards higher elevations with free ears (session 1, green). Tuning curves flattened when molds were inserted (session 2, orange), but recovered after adaptation to the molds (session 3, blue). Shaded areas indicate the standard error across participants.

the mean tuning, local maxima in the cross-validated tuning function are obscured by the overall decreasing tuning curve (Figures 5.5.G–M). Cross-validated mean tuning curves after subtracting the mean had narrow peaks at each selected maximum. Each of these peaks was significantly higher than the mean activation at other elevations (as tested by a one-tailed Wilcoxon signed-rank test, Figure 5.10.A). The mean height of the peaks compared to other elevations corresponded to $\sim 20\%$ of the BOLD signal change of the global decreasing tuning curve. No topographical pattern or gradient was evident in the distribution of these peaks across the surface of the auditory cortex. If the peaks relate to elevation perception, the loss in elevation perception caused by the earmolds should be reflected by a decrease in peak height from session 1 to 2, and adaptation to the molds should be reflected by an increase from session 2

to 3. This was observed. After insertion of the molds, narrow tuning was no longer evident (Figure 5.10.B) and the height of the peaks decreased (mean standardized peak height from session 1 to session 2: 0.2 ± 0.03 to 0.1 ± 0.03 ; one-tailed Wilcoxon signed-rank test, $p = 0.015$, $n = 15$) and most of them were no longer significantly higher than the mean activity at other elevations. After adaptation, all peaks, except one, were significantly above the mean activity (Figure 5.10.C), and peak height increased significantly from session 2 to 3 (mean standardized peak height: from 0.1 ± 0.03 to 0.19 ± 0.03 ; one-tailed Wilcoxon signed-rank test, $p = 0.011$, $n = 15$).

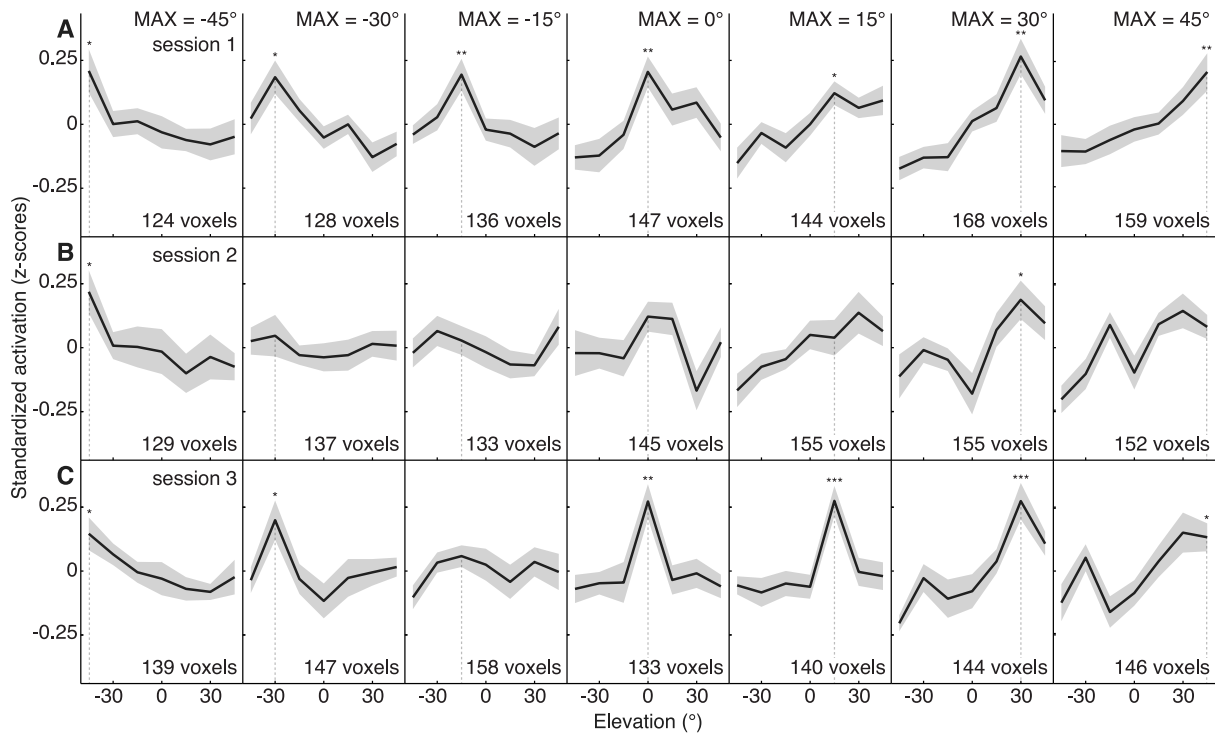


Figure 5.10 – Mean tuning curves after subtracting the global mean. Mean tuning curves of voxels binned according to their tuning curve maximum, after subtracting the global mean tuning of each participant. Shaded areas indicate the standard error across participants ($n = 15$). In each bin, corresponding tuning curves were found in all participants. To avoid selection bias, maxima were computed in one half of the data and mean tuning curves were calculated in the other half of the data (and inversely). In each plot, mean tuning curves are shown in black and the bin maximum is indicated by a vertical dotted line. The mean number of voxels across participants selected by cross-validation is listed in each plot. Whether the peaks at each selected elevation were significantly higher than the mean activation at other elevations was tested by a one-tailed Wilcoxon signed-rank test, and the number of stars above each peak indicate p -values below 0.05 (*), 0.01 (**), and 0.001 (***). The peaks were significant at each selected elevation in session 1 (A). Only two out of seven peaks were significant in session 2, after insertion of the molds (B). All but one peak were significant in session 3, after adaptation to the molds (C).

5.5 DISCUSSION

5.5.1 Behavioral adaptation to modified spectral cues

Consistent with previous studies, the insertion of earmolds produced a marked decrease in elevation performance (Carlile, 2014; Hofman et al., 1998; Trapeau et al., 2016). Most participants recovered a portion of their elevation perception after wearing the molds during seven days and taking part in daily training sessions. As observed in previous earmold adaptation studies (Carlile & Blackman, 2013; Hofman et al., 1998; Trapeau et al., 2016; M. M. V. Van Wanrooij & Van Opstal, 2005), individual differences in adaptation were large and ranged from full recovery to no adaptation. Various factors might explain these differences, such as the individual capacity for plasticity or differences in lifestyle that may have lead to different amounts of daily sound exposure or multisensory interactions (Carlile, 2014; Javer & Schwarz, 1995). In addition, the acoustical effects of the earmolds vary from one ear to another and acoustical differences between directional transfer functions with and without molds explain a portion of individual differences in the amount of adaptation (Trapeau et al., 2016).

Only closed-field stimuli can be presented in typical human neuroimaging setups, and the difficulty to achieve proper elevation perception for such stimuli hampered earlier attempts to measure elevation tuning (Fujiki et al., 2002; Lewald et al., 2008). To circumvent this problem we trained participants in an audio-visual sensory-motor task using the same earphones as in the scanner, which allowed participants to achieve accurate perception of the fMRI stimuli.

5.5.2 Elevation tuning in auditory cortex

We measured fMRI responses to sounds perceived at different elevations and computed voxel-wise elevation tuning curves. Elevation tuning curves were wide and approximated a linear decrease in activation with increasing elevation. Their slope was correlated with elevation perception as measured by behavioral responses: modifying spectral cues with earmolds, thereby reducing elevation perception, resulted in noisier and flatter tuning curves; changes reflecting the perceptual adaptation to the modified cues were observed from session 2 to 3, which were

identical in all aspects. After adaptation, the shape of the tuning curves re-approached the one measured with free ears. The difference in the tuning observed for physically identical stimuli in sessions 2 and 3 indicates that voxel tuning functions depend on listening experience with spectral cues and thus do not solely represent physical features of the spectral cues. The change in tuning covaried with changes in elevation perception between the sessions, which may indicate that the tuning functions in low-level auditory cortex underlie the perceived elevation of a sound source. At the population level, cortical neurons may thus encode perceived sound source elevation using the slope of the tuning curves, as previously proposed for the coding of horizontal sound direction in the human auditory cortex (Magezi & Krumbholz, 2010; McLaughlin, Higgins, & Stecker, 2016; Salminen, Tiitinen, Miettinen, et al., 2010; Salminen et al., 2009; Trapeau & Schönwiesner, 2015). A potential neural mechanism for the observed voxel-level responses is a rate code.

Why would cortical neurons respond most strongly to lower elevations? In cats, an early stage of spectral cues processing is carried out by notch-sensitive neurons of the dorsal cochlear nucleus (Imig et al., 2000; E. D. Young et al., 1992) and by their projections to type-O units of the central nucleus of the inferior colliculus (Davis et al., 2003). Deeper spectral notches at low elevations might result in increased sensitivity of these cells to low elevations and might have influenced the general cortical tuning preference for low rather than high elevations during evolution. Previous plots of directional transfer functions suggest deeper spectral notches at low elevations than at high elevations (Cheng & Wakefield, 1999; Middlebrooks, 1999, 1997; Raykar et al., 2005; Trapeau et al., 2016). The acoustic elevation gain computed from our stimuli shows that low elevations contained deeper notches in frequency bands previously identified as crucial for elevation perception (Langendijk & Bronkhorst, 2002; Trapeau et al., 2016). These acoustic features could conceivably drive stronger responses at low elevations. However, such spectral features do not solely determine the shape of cortical elevation tuning curves: swapping stimuli between participants disrupted elevation perception and led to tuning curves that no longer exhibited a linear decrease with increasing elevation. Additionally, the adaptation to modified spectral cues showed that differences in perception of physically identical stimuli change the

shape of tuning curves in low-level auditory cortex. These results suggest that the tuning curves observed in auditory cortex do not encode acoustical elevation cues, but the source elevation computed from these cues. This provides evidence for the notion that low-level auditory cortex represents not only physical stimulus features, but also the subjective perception of these features (Kilian-Hütten, Valente, Vroomen, & Formisano, 2011; Staeren, Renvall, Martino, Goebel, & Formisano, 2009). A model of elevation processing posits that spectral profiles of an incoming sound are compared to templates for different elevations (Hofman & Van Opstal, 1998). In terms of this model, auditory cortex activity may reflect the result of this comparison, and a correct mapping would enable accurate perception. In turn, experience may reshape this mapping or templates.

Obscured by the main broad tuning, a secondary narrow tuning was observed after subtracting each participant's mean tuning curve from the data. Voxels exhibiting significant secondary tuning peaks were found at each elevation and the height of these peaks was approximately a fifth of the size of the overall decrease with increasing elevation. With modified spectral cues, secondary peaks were initially almost absent, but reappeared after adaptation to the modified cues. Therefore, both broad tuning aspect and secondary narrow tuning aspect of the elevation tuning curves covary with elevation perception. This finding may indicate that sound elevation is encoded by two complementary tuning mechanisms. In gerbils, the representation of ITDs seems to be transformed from broadly-tuned channels in the midbrain to a uniform sampling by sharply-tuned responses in the primary auditory cortex (Belliveau, Lyamzin, & Lesica, 2014), and a representation of ITDs involving multiple types of codes in human auditory cortex has been recently speculated about (McLaughlin et al., 2016). Two populations of neurons may be necessary for elevation encoding because pinna and head/torso contribute independent spectral cues in different frequency ranges. Pinna cues lie above 4 kHz, whereas torso cues exist only below 3 kHz (Algazi et al., 2001; Asano et al., 1990). Different optimal coding strategies could then arise from the different spectral cue frequency ranges (Harper & McAlpine, 2004).

Comprehensive data from animal models would aid in interpreting the voxel-level results in humans. No clear pattern for an overall elevation tuning preference was observed among the

few elevation sensitive neurons found in macaque primary auditory cortex and caudomedial field (Recanzone et al., 2000; Recanzone, 2000). One study in cats reported distributed elevation tuning preference across neurons in nontonotopic auditory cortex, with some evidence for an overall preference for upper frontal elevations (Xu et al., 1998). However, elevation encoding may differ between cats and humans due to differences in body height, ear shape, and ear mobility.

We observed experience-dependent changes in auditory cortex, but this plasticity may emerge already in subcortical structures, where spectral cues are extracted. The corticofugal system has been suggested to play a role in the modification of subcortical sensory maps in response to sensory experience (Yan & Suga, 1998) and plays a critical role in experience-dependent auditory plasticity (Suga et al., 2002). Bajo et al. (2010) demonstrated that auditory corticocollicular lesions abolish experience-dependent recalibration of horizontal sound localization. The integrity of this descending pathway may be needed in the adaptation to modified spectral cues.

5.5.3 Limitations

The supine position in the scanner may have influenced our results. Body tilt affects horizontal sound localization judgments (Comalli & Altshuler, 1971; Lackner, 1974) and sound location is perceived in a world coordinate reference frame (Goossens & Van Opstal, 1999; Vliegen, Grootel, & Van Opstal, 2004). At the level of the auditory cortex it is still unclear whether the representation is craniocentric (Altmann, Wilczek, & Kaiser, 2009), or both craniocentric and allocentric (Schechtman, Shrem, & Deouell, 2012; Town, Brimijoin, & Bizley, 2017). In the latter case, subsets of neurons representing sound location in a world coordinate frame might have influenced the shape of the elevation tuning curves.

We acquired a total of 52 one-hour sessions of fMRI data from 16 participants, considerably exceeding the amount of data in a typical fMRI study. A higher number of participants would have been useful to identify sources of the interindividual variability in adaptation to our ear manipulation and the potential correlation between behavioral and brain data within sessions, but was technically and economically not feasible.

5.5.4 Outlook

Our results suggest that sound elevation is represented in the auditory cortex by broad tuning functions preferring lower elevations as well as secondary narrow tuning functions preferring individual elevation directions. The shape of these tuning functions covaried with the perception of the stimuli, highlighting the link between low-level cortical activity and perception. Our results emphasize the complexity of cortical processing of sound location. A distinct neural population or tuning might be needed to represent sounds in the rear field. Regions involved in vertical and horizontal auditory space processing seem to largely overlap (compare with [Trapeau & Schönwiesner, 2015](#)), and there is growing evidence for an integrated cortical processing of localization cues ([Edmonds & Krumbholz, 2014](#); [Pavani et al., 2002](#)). In animal models, single-unit recordings demonstrated that cortical neurons are sensitive to multiple localization cues ([Edmonds & Krumbholz, 2014](#); [Pavani et al., 2002](#)). Horizontal and vertical auditory space might be represented by integrated codes, and future experiments should combine these dimensions.

5.5.5 Acknowledgments

The authors would like to thank Vincent Migner-Foisy for collecting most of the data. Funding was provided by an operating grant from the Canadian Institutes of Health Research to M.S. (MOP-133716).

CHAPITRE 6

DISCUSSION GÉNÉRALE

Cette recherche avait pour objectif d'étudier la plasticité du système auditif humain dans sa capacité à localiser les sources sonores. Dans chacune des quatre études regroupées dans cette thèse, cette plasticité a été observée. En introduisant artificiellement un retard d'une durée proche de l'ITD maximale que peut procurer l'écartement de nos oreilles, les deux premières études ont démontré qu'il était possible de rapidement s'adapter à des ITDs largement décalées. La capacité d'adaptation à des indices spectraux modifiés a été observée dans les deux études suivantes. L'étude de la plasticité en localisation auditive par des expériences comportementales et de neuroimagerie, a permis d'émettre des hypothèses sur l'encodage des indices de localisation auditive, ainsi que sur les mécanismes adaptatifs par lesquels cette plasticité s'exprime. Ce chapitre est l'occasion de discuter les résultats obtenus dans les différentes études, d'examiner leurs implications et potentielles limites, en ayant une vision d'ensemble de la recherche menée au cours de ce doctorat.

6.1 ENCODAGE DES INDICES BINAURAUX ET PLASTICITÉ

Les deux premières études ont démontré sans équivoque que le système auditif pouvait s'adapter rapidement, et parfois même en seulement quelques heures, à un important décalage des ITDs. En 1995, Javer et Schwarz suggéraient déjà qu'une telle rapidité n'était théoriquement pas compatible avec le modèle de Jeffress d'encodage des ITDs et que certains mécanismes additionnels devaient exister. Vingt ans plus tard, nous savons que le modèle de Jeffress semble ne pas s'appliquer aux mammifères et quelques études récentes suggèrent que ce constat serait également valable pour l'être humain (cf. 1.1.3. Encodage des indices acoustiques). Le système auditif humain représenterait alors le plan horizontal de l'espace auditif par deux populations de neurones à sélectivités directionnelles opposées. Une population préférerait alors l'hémichamp gauche et l'autre l'hémichamp droit. L'étendue de la sélectivité de chacune de ces deux

populations serait large et serait maximale pour l'extrémité d'un des deux hémichamps. La comparaison entre l'activité des deux populations de neurones permettrait alors de déduire la provenance d'un son sur le plan horizontal. En raison de la très forte contralatéralité observée par les études physiologiques animales, il a d'abord été suggéré que les deux populations occupaient chacune un hémisphère et que c'est la comparaison entre l'activité des deux hémisphères qui permettait la localisation sur le plan horizontal (McAlpine, 2005). Mais ce postulat est en désaccord avec les études de lésions unilatérales (ex. Haeske-Dewick, Canavan, & Hömberg, 1996 ; Jenkins & Masterson, 1982), qui montrent que même si la précision en est réduite¹, il est encore possible de localiser des sons sur le plan horizontal avec un seul hémisphère fonctionnel. Un hémisphère seul contiendrait alors le matériel neuronal nécessaire pour pouvoir localiser des sons sur le plan horizontal (Schnupp, Nelken, & King, 2011). Plus récemment, il a été proposé que chaque hémisphère devait contenir une portion de chacune des deux populations, même si leur répartition resterait majoritairement contralatérale (*hemifield code* : Salminen et al., 2009, 2012).

Les résultats fMRI présentés dans le deuxième article concordent parfaitement avec ce modèle. La plupart des voxels du cortex auditif présentaient des courbes d'accord larges et dont le maximum se situait à des positions extrêmes d'un hémichamp ou l'autre. Ces observations ont été confirmées grâce à l'élaboration d'un modèle de représentation de l'espace auditif par des courbes d'accord neuronales. Un algorithme génétique nous a permis de trouver les valeurs optimales de chacun des paramètres de ce modèle, afin qu'il reflète au mieux les données réellement observées. L'un des paramètres du modèle était la proportion de neurones répondant comme suggéré par l'*hemifield code* contre ceux répondant suivant le modèle de Jeffress. À chacune des 50 convergences de l'algorithme génétique, un schéma dans lequel la grande majorité des neurones répondait suivant l'*hemifield code* était proposé. Nous pensons alors que nos résultats constituent une forte preuve de l'existence de ce type d'encodage pour la représentation de l'espace auditif horizontal dans le cortex auditif. Nous aurons été les premiers

¹En particulier dans l'hémichamp contralatéral à la lésion.

à fournir cette preuve en employant la fMRI et en montrant les courbes d'accord directionnelles des voxels du cortex auditif.

Le modèle nous a également aidé à mieux comprendre certains résultats qui étaient surprenants de prime abord. Pour simuler la situation avec bouchons d'oreille numériques, les stimuli de chaque session fMRI comprenaient le retard de $625 \mu\text{s}$ dans le canal gauche. En moyenne, ces stimuli étaient perçus avec un décalage de 45° vers la droite avant l'adaptation. Cependant, les courbes d'accord directionnelles moyennes de chaque hémisphère se croisaient à quelques 8° à gauche de la ligne médiane (Figure 3.8) et les distributions des centres de gravité (COG, *center of gravity*) de chaque hémisphère étaient positionnées de manière symétrique, à 15° de chaque côté de la ligne médiane (Figure 3.9). Étant donné le décalage perceptif de 45° , nous nous attendions plutôt à ce que le croisement des courbes moyennes et la position des distributions des COGs soient décalés par rapport à la ligne médiane. Les courbes et distributions construites à partir du modèle et de ses valeurs de paramètres optimales trouvées par l'algorithme génétique, ressemblaient néanmoins de près aux courbes obtenues expérimentalement. De plus, l'algorithme génétique avait trouvé pour le paramètre *shift*², une valeur optimale de 59° , dans la direction prévue. Ceci montre que les positions des courbes et distributions observées sont bien compatibles avec la présence du décalage perceptif. Mais pourquoi ce décalage n'est-il pas visible sur ces figures ?

Au moment d'écrire ce chapitre, un nouvel examen des résultats de cet article nous a permis de répondre à cette question. Si les courbes d'accord sont bien décalées d'une soixantaine de degrés, alors la courbe moyenne de l'hémisphère gauche devrait avoir son maximum autour de 30° à droite. C'est le cas pour la courbe modélisée et cela semble l'être également pour les données fMRI. Du côté de l'hémisphère droit, le maximum devrait se trouver autour de -150° . Ceci est impossible à observer sur les données fMRI car la fenêtre d'analyse ne couvre que $\pm 60^\circ$. Cependant, si c'est le cas, cela veut dire que l'on n'observe que la partie minimale de la courbe de l'hémisphère droit, et seulement la partie maximale de la courbe de l'hémisphère gauche.

²Qui décale l'azimut à l'entrée du modèle.

Comme les données sont normalisées³, la partie haute de la courbe de l'hémisphère gauche et la partie basse de celle de l'hémisphère droit, seraient alors placées à hauteur équivalente sur l'axe des ordonnées. On remarque sur la figure 3.12 que cela semble être le cas. L'introduction de ce biais de placement des courbes par la normalisation déplace leur point de croisement et explique pourquoi il est si proche de la ligne médiane et non à des valeurs avoisinant les 60°.

Nous ne pouvons pas agrandir la fenêtre d'analyse des données fMRI, cependant ceci est possible avec le modèle. Un agrandissement de la fenêtre d'analyse à $\pm 180^\circ$ permet de couvrir toute l'étendue des courbes d'accord du modèle et d'éviter le biais introduit par la normalisation. Le résultat est visible figure 6.1. Les courbes se croisent bien à un azimuth proche du *shift* de 59° proposé par l'algorithme génétique.

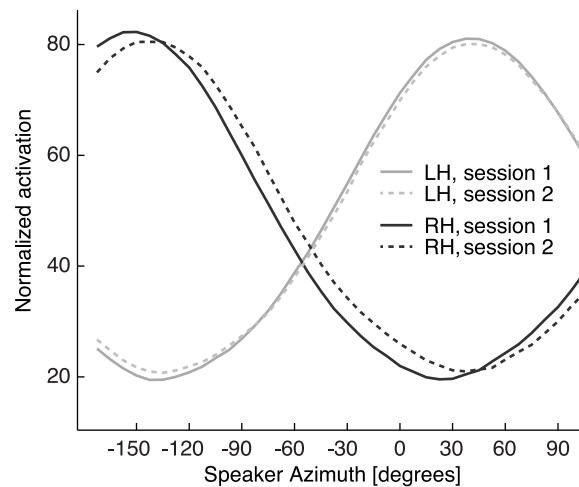


Figure 6.1 – Courbes d'accord directionnelles moyennes créées à partir du modèle. Courbes d'accord moyennes du modèle calculées sur une plage d'azimuts de $\pm 180^\circ$, plutôt que celle de $\pm 60^\circ$ avec laquelle les courbes sont calculées en figure 3.12. Ainsi calculées, la position en ordonnée des courbes d'accord n'est pas biaisée par la normalisation et les courbes se croisent à environ 58° pour la session 1 et 51° pour la session 2.

Concernant les distributions des COG, leur placement relativement à la ligne médiane s'explique par le fait que pour des courbes traversant tout l'espace d'analyse, le COG dépend plus de la pente de la courbe que de sa position. Une courbe de pente positive aura un COG positif. Les pentes absolues des courbes étant assez proches, les distributions des COGs sont symétriques par rapport à l'axe médian. Une manipulation des paramètres du modèle montre

³Car on veut observer et analyser la forme des courbes d'accord avant toute autre propriété.

que plus le paramètre *shift* est proche de zéro, plus ces distributions sont écartées, et inversement, plus le *shift* augmente, plus ces distributions se rapprochent de la ligne médiane. Ces dernières observations avaient déjà été faites dans l'article.

L'aspect le plus important des distributions des COG est la hauteur de chaque distribution, car ces hauteurs reflètent la taille des populations de neurones qui préfèrent l'hémichamp gauche ou l'hémichamp droit. La figure 3.9) montre que le modèle a très fidèlement reproduit les distributions observées en session 1, alors que le résultat est plus mitigé pour la session 2. En session 2, la distribution des COGs observée pour l'hémisphère droit augmente nettement par rapport à la session 1. Cette augmentation est beaucoup moins nette dans les résultats du modèle et la distribution modélisée ne reflète alors pas complètement les données. Aucune des 50 solutions générées par l'algorithme génétique n'est arrivée à reproduire cette augmentation de la taille de la distribution des COGs dans l'hémisphère droit. Ceci est forcément le reflet d'une faiblesse du modèle, ce qui suit explique laquelle.

Les données fMRI présentées dans l'article montraient une augmentation, entre la première et la deuxième session, du nombre de voxels répondant de manière significative dans l'hémisphère droit. Nous avons alors suggéré que ce changement de latéralisation hémisphérique devait être causé, soit par un changement de la taille relative des deux populations neuronales, soit par un changement de sélectivité de ces populations. Les deux seuls paramètres du modèle qui ne présentaient pas des valeurs similaires entre les solutions des sessions 1 et 2 étaient la direction préférée des populations (qui passe de 94° à 89°) et le *shift* à l'entrée du modèle (qui diminue de 59° à 52°). Nous avons alors conclu qu'en plus du changement de latéralisation hémisphérique, un petit *shift*, c'est-à-dire un déplacement de la sélectivité neuronale, semblait être un des corrélats de l'adaptation aux ITDs décalées. Cependant, lors de l'évaluation des paramètres du modèle par l'algorithme génétique, la taille en nombre de voxels de chaque hémisphère correspondait à la latéralisation hémisphérique réellement observée. Ce paramètre était donc fixe, et le *shift* était alors l'un des seuls paramètres avec lequel l'algorithme génétique pouvait jouer pour répliquer les données observées en session 2. De plus, le paramètre qui déterminait le niveau de contralatéralité dans la distribution hémisphérique des deux populations neuronales

était le même pour les deux hémisphères. Cela implique que la contralatéralité du modèle était forcément symétrique entre les deux hémisphères. Ceci n'a pas freiné l'algorithme génétique pour trouver une très bonne solution pour la session 1, mais il est probable que ce soit ce qui l'a empêché de trouver une aussi bonne solution pour la session 2. Cette propriété de symétrie de la contralatéralité avait été implémentée au modèle pour limiter la convergence de l'algorithme génétique vers des maxima locaux, mais il est très probable qu'elle ait dissimulé une partie des résultats. L'algorithme génétique n'avait en effet pas de contrôle sur la taille relative des deux populations neuronales dans chacun des deux hémisphères. L'idée que cette propriété ait freiné l'algorithme génétique est renforcée par l'observation dans les données fMRI d'un changement du niveau de contralatéralité des préférences aux stimuli avec ITDs décalées dans l'hémisphère droit seulement.

C'est pour ces raisons qu'à travers le présent chapitre, nous nuancions l'hypothèse émise dans le deuxième article de déplacement de sélectivité neuronale comme mécanisme adaptatif à des ITDs décalées. Les données ne permettent pas d'émettre de conclusion ferme, mais l'hypothèse d'un changement de la taille relative des deux populations neuronales⁴ ne doit pas être mise de côté. Si l'encodage des indices binauraux tient en partie grâce à la comparaison entre l'activation des deux populations à sélectivité opposée comme proposé par McAlpine (McAlpine, 2005), il est probable qu'un des mécanismes adaptatifs soit de maintenir la balance entre ces deux populations. Avant adaptation aux ITDs décalées, sur l'ensemble des voxels des deux hémisphères dont nous avons extrait les courbes d'accord, 32% avaient un COG à gauche de la ligne médiane des stimuli présentés, contre 68% à droite⁵. Après adaptation ces valeurs sont passées à 37% pour la gauche et 63% pour la droite. Pour compenser le décalage perceptif vers le droite, un changement des tailles relatives des deux populations en faveur de la population préférant l'hémichamp gauche aurait donc eu lieu. Comme les stimuli comportaient le décalage en ITDs il est impossible de savoir qu'elle était la répartition initiale des deux populations dans les deux

⁴Changement de taille pouvant causer des changements du niveau de contralatéralité différents dans chaque hémisphère.

⁵Données non présentées dans l'article.

hémisphères. Les techniques d'analyse utilisées dans cette étude permettraient de révéler cette répartition à travers une expérience où des stimuli non modifiés seraient présentés.

Les éclaircissements apportés dans cette partie de la discussion générale nous ont également permis de montrer que le décalage des ITDs dans les stimuli fMRI est bien observable au niveau du croisement des courbes d'accord du modèle (Figure 6.1), même s'il n'est pas directement visible dans les données fMRI.

6.2 ENCODAGE DES INDICES SPECTRAUX ET PLASTICITÉ

La capacité d'adaptation à de nouveaux indices spectraux était déjà connue et beaucoup plus documentée que la capacité à s'adapter à des ITDs décalées. Nos deux dernières études l'ont à nouveau démontrée. La troisième étude était destinée à explorer les mécanismes d'apprentissages de nouveaux indices spectraux et a montré qu'une mise en correspondance des indices spectraux avec les positions de l'espace dans laquelle plusieurs DTFs pourraient indiquer une même position (*many-to-one mapping*), semblait exister⁶. La quatrième étude n'avait pas pour but d'étudier l'adaptation à de nouveaux indices spectraux, mais tirait profit de cette plasticité pour explorer les mécanismes encore inconnus de l'encodage de l'élévation sonore dans le cortex auditif.

Une modification des indices binauraux entraîne un décalage de la perception sur le plan horizontal, mais n'affecte pas la discrimination de sources sonores placées à des azimuts différents. Modifier les indices spectraux n'entraîne pas de décalage mais plutôt une abolition de la perception sur le plan vertical. Cette particularité, ajoutée à notre capacité d'adaptation à des indices spectraux modifiés, offre l'opportunité unique de pouvoir présenter à un auditeur des stimuli qu'il sera incapable de discriminer lors d'une première séance de neuroimagerie, mais pour lesquels il aura une bonne perception de l'élévation, lors d'une seconde séance, une fois que l'adaptation a pris place. La comparaison entre les données fonctionnelles de ces deux séances permet alors d'énoncer des hypothèses relatives à la perception de l'élévation et d'exclure des

⁶Ce résultat est discuté plus en détail à la section 6.3. **Rapidité d'adaptation, absence d'effet consécutif et persistance.**

interprétations basées seulement sur les propriétés physiques de stimuli. C'est par ce paradigme que nous avons pu formuler une proposition de l'encodage de l'élévation au niveau du cortex auditif dans le quatrième article. Il semblerait donc que l'élévation soit encodée de manière distribuée dans le cortex auditif de chaque hémisphère, par une population de neurones à sélectivité large et préférant les élévations basses.

Cet encodage partage plusieurs propriétés avec l'encodage des indices binauraux, c'est un codage fréquentiel, effectué par une population de neurones dont la sélectivité est étendue. L'existence d'un même type d'encodage pour les différents indices de localisation n'est pas très surprenante. L'organisation tonotopique du cortex auditif oblige à avoir une représentation de l'espace qui soit distribuée, et effectuée selon un schéma de codage par population (Recanzone & Sutter, 2008). Un tel encodage a également la capacité de transporter différentes propriétés d'un stimulus simultanément (Abbott & Sejnowski, 1999) et peut donc permettre l'intégration des indices de localisation en un seul percept de la position spatiale. Les régions du cortex auditif qui montrent une sensibilité aux indices binauraux et aux indices spectraux semblent en effet largement se chevaucher (si l'on compare les figures 3.10 et ??), et des études physiologiques chez l'animal ont montré que des mêmes neurones corticaux montraient une sélectivité à plusieurs indices de localisation (Brugge et al., 1994; Mrsic-Flogel et al., 2005; Xu et al., 1998).

Il semble que l'encodage des indices spectraux ne soit porté que par une seule population de neurones, alors que deux populations encodent l'espace horizontal. Comme nous l'avons discuté dans le quatrième article, il est possible que l'encodage des indices spectraux ne soit pas limité à cette population. Nous n'avons en effet pas sondé l'espace auditif situé à l'arrière de l'auditeur. Il est possible que cette partie de l'espace, pour laquelle d'autres indices spectraux sont en correspondance avec les positions du plan vertical, soit encodée par une autre population de neurones montrant quant à elle une préférence pour des élévations hautes.

6.3 RAPIDITÉ D'ADAPTATION, ABSENCE D'EFFET CONSÉCUTIF ET PERSISTANCE

La rapidité d'adaptation n'est pas la même suivant l'indice qui a été modifié. Dans les deux premières études, les participants qui se sont adaptés avec succès aux ITDs décalées, l'ont fait en moins de 48 h, et ne bénéficiaient d'aucun entraînement spécifique. Dans les études 3 et 4, même avec un entraînement quotidien, rares sont ceux qui ont retrouvé une performance identique à l'originale après une semaine d'adaptation à de nouveaux indices spectraux. Sans entraînement, il a été montré que cette adaptation prenait entre trois et huit semaines (Carlile & Blackman, 2013 ; Hofman et al., 1998). Une telle différence de rapidité d'adaptation doit pouvoir s'expliquer par plusieurs facteurs. Comme nous l'avons mentionné précédemment, une modification des indices binauraux ou des indices spectraux n'a pas du tout la même incidence perceptive. L'adaptation à des indices binauraux modifiés consiste essentiellement à ajuster un espace auditif pivoté, alors que l'adaptation à de nouveaux indices spectraux nécessite l'apprentissage de nouvelles associations entre indices et positions spatiales. Les résultats du troisième article suggèrent que ce sont bien de nouvelles associations qui sont faites car les anciennes sont toujours disponibles après adaptation. L'adaptation nécessite alors sûrement la création de nouvelles connexions neuronales, ce qui a peu de chances de pouvoir se produire en l'espace de quelques heures. L'existence des indices binauraux et spectraux repose sur des principes physiques fondamentalement différents et leur extraction au niveau sous-cortical doit mettre en jeu des mécanismes très différents eux aussi. Le temps nécessaire pour mettre en place ou modifier ces mécanismes doit donc varier d'un indice à l'autre. Enfin, il est possible qu'un des facteurs expliquant les différences de rapidité d'adaptation soit l'exposition préalable aux modifications induites. Les chances pour que les participants aient déjà été exposés à la modification causée par les moulages en silicone sont inexistantes, alors qu'une exposition antérieure à des ITDs décalées est très probable. Si elle est unilatérale, l'otite séro-muqueuse peut en effet induire un décalage des ITDs aussi important que celui causé par le retard que nous introduisons dans l'un des bouchons d'oreille numériques. Ce type d'otite étant très fréquent chez le bébé et l'enfant, il est possible que beaucoup d'entre nous aient eu à s'adapter à des

ITDs décalées par le passé. De plus, comme nous l'avons vu dans les études sur la plasticité développementale (cf. 1.2.2. **Existence d'une période critique ?**), l'adaptation à une contrainte à l'âge adulte est facilitée si cette adaptation a déjà eu lieu pendant la période de développement, or c'est principalement durant cette période que ce type d'otite survient le plus fréquemment. Il est donc possible qu'un des facteurs expliquant la rapidité d'adaptation à des ITDs décalées, soit l'exposition antérieure à cette contrainte.

Immédiatement après l'adaptation à des indices de localisation modifiés, aucun effet consécutif n'est observé, que ce soit pour les indices binauraux ou spectraux. À travers le troisième article, nous avons inspecté l'absence d'effet consécutif à l'adaptation à de nouveaux indices spectraux dans de multiples conditions. Dès la toute première présentation d'un stimulus contenant les indices spectraux d'origine, la performance est du même niveau qu'avant l'adaptation. Ceci est vrai même dans la condition où les stimuli sont des enregistrements binauraux présentés aux écouteurs. Dans cette condition, les participants portent encore les moulages et ignorent que les stimuli correspondent à la situation d'origine avant adaptation. Parce que les stimuli ne sont pas présentés en champ libre, les indices de localisation dynamiques sont également absents dans cette condition. Ces observations nous ont amené à formuler l'hypothèse que plusieurs DTFs pouvaient simultanément être associées à une même position spatiale, hypothèse de *many-to-one mapping*. Nous avons à nouveau vérifié cette observation au cours de la quatrième étude. Dans cette étude, les participants étaient entraînés tous les jours à localiser les enregistrements avec indices spectraux modifiés et leur performance de localisation de ces enregistrements était également mesurée tous les jours. Au cours d'une de ces tâches quotidiennes nous avons présenté, à l'insu des participants, les enregistrements contenant les indices spectraux d'origine⁷. La performance des participants lors de cette tâche était identique à celle mesurée avant adaptation.

Dans les études où les ITDs étaient décalées, nous n'avons pas observé d'effet consécutif non plus. Ultérieurement à la publication de ces études, nous avons analysé le test où les bouchons d'oreille numériques sont retirés, en utilisant la même méthode que celle présentée dans le troisième article. Cette analyse montre que dès le premier essai, les participants localisaient

⁷Données non présentées dans l'article.

correctement les stimuli. Dans le cas des indices spectraux, l'absence d'effet consécutif est facilement concevable. De nouveaux indices ont été appris, cet apprentissage n'a pas effacé le savoir-faire antérieur. Mais dans le cas des indices binauraux, pour lesquels il n'y a *a priori* pas eu d'apprentissage de nouveaux indices, mais un ajustement des indices existants, cette particularité est remarquable. Elle l'est d'autant plus qu'une persistance de l'adaptation aux ITDs modifiées est observée. Dans la deuxième étude, les participants qui ont à nouveau été testés avec les bouchons d'oreille numériques, trois jours après les avoir retirés, localisaient aussi bien les stimuli que lors du dernier test de la période d'adaptation. Une participante, celle qui s'était adaptée en quelques heures, a été testée un mois après avoir retiré les bouchons. À nouveau, sa performance égalait celle qu'elle avait obtenue en fin de période d'adaptation.

Un bémol doit cependant être mis à ces observations. Dans les études où nous décalions les ITDs, le retour aux indices d'origine nécessitait de retirer les bouchons. Or un effet secondaire des bouchons était de modifier les indices spectraux. Donc en retirant les bouchons, non seulement les ITDs redevenaient ceux d'origine, mais c'était également le cas pour les indices spectraux. Il est possible que le retour aux indices spectraux d'origine indique au système auditif que la situation a changé. Il est également possible que l'ajustement des ITDs ait été effectué en relation avec les indices spectraux modifiés et avec ces indices seulement. Les bouchons d'oreille numériques ne permettaient pas de modification ou d'annulation instantanée du retard ajouté dans une oreille. Il fallait retirer les bouchons des oreilles des participants et quelques minutes étaient nécessaires pour effectuer le changement. Sans cela, nous aurions pu tester si l'annulation subite du retard entraînait un effet consécutif ou non. Ceci reste donc à vérifier dans une étude future, qui inclurait des tests de localisation dans un environnement auditif virtuel, par le biais duquel il est facile de passer instantanément des ITDs décalées aux ITDs normaux, sans que les autres indices ne soient affectés.

6.4 VARIABILITÉ INTERINDIVIDUELLE

Dans chacune des études présentées dans cette thèse, les différences interindividuelles en matière d'adaptation étaient importantes. Après une altération ou une modification de leurs indices de localisation, certains participants retrouvent rapidement des aptitudes de localisation très proches de celles précédant la manipulation, d'autres ne montrent aucun signe d'adaptation. Les différentes performances des participants ne permettent pas de les classer en deux groupes distincts qui comprendraient seulement ceux qui s'adaptent et ceux qui ne s'adaptent pas. Les différents degrés d'adaptation observés sont plutôt uniformément répartis entre les deux extrêmes.

Une large variabilité interindividuelle dans le degré d'adaptation à des indices de localisation biaisés a également été reportée par la majorité des études antérieures du domaine (Carlile et al., 2014 ; Hofman et al., 1998 ; Javer & Schwarz, 1995 ; Shinn-Cunningham et al., 1998 ; M. M. V. Van Wanrooij & Van Opstal, 2005). Il est possible que cette variabilité reflète des prédispositions générales à la plasticité. Cependant une telle hypothèse est difficile à vérifier étant donné les innombrables facteurs qui peuvent entrer en jeu durant la période d'adaptation. Il faudrait pouvoir observer l'adaptation dans un environnement entièrement contrôlé, mais étant donné l'importance des interactions multisensorielles en milieu naturel dans le processus d'adaptation, il est possible que peu ou pas d'adaptation ne prenne place en laboratoire. Il a d'ailleurs été proposé que la quantité et la qualité de ces interactions jouent un rôle dans l'adaptation (Carlile et al., 2014 ; Javer & Schwarz, 1995). Différentes activités ou différents styles de vie (par exemple, travailler toute la journée sur un ordinateur dans un environnement silencieux *versus* pratiquer des activités sportives à l'extérieur avec des amis) ont certainement différents impacts sur la rapidité d'adaptation, mais encore une fois, ces facteurs sont très difficiles à contrôler.

D'autres facteurs sont plus facilement observables et quantifiables. Les performances de localisation auditive dans des conditions normales varient d'un individu à l'autre (Makous & Middlebrooks, 1990 ; Populin, 2008 ; Savel, 2009 ; Wightman & Kistler, 1989a). Il est légitime

de penser que ces différences puissent expliquer une partie de la variabilité observée dans l'adaptation aux indices de localisation modifiés. Les données de nos différentes études n'ont cependant révélé aucune corrélation entre les performances initiales et la quantité ou la rapidité d'adaptation. Néanmoins, il faut garder à l'esprit que nos expériences n'étaient pas spécialement conçues pour étudier ces possibles corrélations, et nos participants étaient de jeunes adultes qui montraient tous de bonnes performances initiales. Un facteur qui en revanche a expliqué une partie de la variabilité de l'adaptation à une modification des indices spectraux et binauraux, est le degré de perturbation des indices spectraux. Dans la troisième étude, nous avons proposé une mesure de la perturbation acoustique des indices spectraux⁸ et montré que cette mesure corrélait avec la perte de performance sur le plan vertical causée par les moulages. Cette corrélation existait encore à la fin de la période d'adaptation. Les participants qui avaient le plus été perturbés étaient aussi ceux qui s'adaptaient le moins. Cette corrélation a également été observée pour l'adaptation aux ITDs décalées. La perte de performance que causaient les bouchons d'oreille sur le plan vertical, avant qu'un retard ne soit ajouté dans l'un des bouchons, corrélait avec la vitesse d'adaptation sur le plan horizontal. Le degré de perturbation des indices spectraux semble donc être un important prédicteur des chances d'adaptation à des indices de localisation modifiés, et suggère qu'il est impossible de s'adapter à de trop fortes modifications.

Ce dernier point a des conséquences pour la conception d'appareils auditifs. Une part importante des appareils auditifs sont retournés, ou ne sont pas utilisés par les patients atteints de pertes auditives (Kochkin, 1999 ; Kochkin et al., 2010). Ces défauts de satisfaction sont imputables aux insuffisants gains perceptifs qu'offrent les appareils, en particulier en termes de compréhension de la parole dans le bruit (Wong, Hickson, & McPherson, 2003). L'importance de la localisation auditive dans l'intelligibilité et la compréhension de la parole dans un environnement bruyant ou dans une situation dite de *cocktail party*, n'est plus à démontrer (Bronkhorst, 2000 ; Dirks & Wilson, 1969 ; Hirsh, 1950 ; Kidd et al., 2005 ; Kock, 1950 ; MacKeith & Coles, 1971 ; Roman et al., 2003). Or les appareils auditifs perturbent la localisation auditive (Byrne &

⁸VSI dissimilarity, voir 4.2.10. Directional transfer functions.

Noble, 1998 ; Köbler & Rosenhall, 2002). Ils peuvent décaler les ITDs⁹, et modifient fortement les indices spectraux (Kuk & Korhonen, 2014). Pour que le gain en amplitude acoustique que fournissent les appareils auditifs soit bénéfique dans des situations où la compréhension de la parole est mise à l'épreuve, il faut donc que les patients soient capables de réapprendre à localiser les sons avec leurs appareils. Nos résultats montrent que cette adaptation sera d'autant plus difficile que la perturbation des indices de localisation causée par les appareils est grande. Ceci est particulièrement vrai concernant la modification des indices spectraux. Nous avons en effet vu que l'adaptation, tant sur le plan horizontal que vertical, dépendait du degré de perturbation initial de ces indices. Pour qu'une adaptation prenne place et pour que les gains perceptifs offerts par les appareils auditifs soient maximaux, il est donc important que ces dispositifs soient miniaturisés et que le micro se trouve au plus proche de l'entrée du canal auditif, afin de perturber le moins possible les indices spectraux.

⁹D'autant plus si portés unilatéralement.

6.5 CONCLUSION

Cette recherche avait pour objectif d'étudier la plasticité du système auditif humain dans des tâches de localisation sonore, ainsi que les mécanismes d'encodage des indices de localisation auditive. La plasticité en localisation auditive a été observée dans chacune des quatre études regroupées dans cette thèse. Son analyse par des techniques comportementales et de neuroimagerie, a permis d'émettre des hypothèses sur l'encodage des indices de localisation auditive, ainsi que sur les mécanismes adaptatifs par lesquels cette plasticité s'exprime.

La première étude avait pour principal objectif de savoir si l'adaptation à des ITDs décalées de 625 μ s était possible. Les résultats ont montré que de jeunes adultes peuvent rapidement s'adapter à cet important décalage perceptif. Au moyen de la fMRI haute résolution, nous avons pu confirmer l'hypothèse de codage par hémichamp comme représentation de l'espace auditif horizontal, et avons observé des changements de l'activité corticale auditive accompagnant l'adaptation. Ces changements s'expriment en termes de latéralisation hémisphérique, que nous imputons à une probable modification des tailles relatives des deux populations qui encodent l'espace auditif horizontal. Les résultats d'une modélisation de la représentation de l'espace auditif horizontal par des neurones corticaux virtuels, associée à une optimisation des paramètres de ce modèle par un algorithme génétique, suggèrent qu'un fin réajustement de la sélectivité des populations de neurones qui encodent l'espace auditif horizontal aurait eu lieu. Les résultats de la modélisation renforcent également l'hypothèse de codage par hémichamp.

Dans une troisième étude, nous avons modifié les indices spectraux à l'aide de moulages en silicone. Nous avons montré que l'adaptation à cette modification n'était suivie d'aucun effet consécutif au retrait des moulages, même lors de la toute première présentation d'un stimulus sonore. Ce résultat concorde avec l'hypothèse d'un mécanisme dit de *many-to-one mapping*, à travers lequel plusieurs profils spectraux peuvent être associés avec une même position spatiale. La quatrième étude n'avait pas pour but d'étudier l'adaptation à de nouveaux indices spectraux, mais tirait profit de cette plasticité pour explorer les mécanismes encore inconnus de l'encodage de l'élévation sonore dans le cortex auditif. Les résultats fMRI nous ont permis de révéler

que l'élévation semble être encodée de manière distribuée dans le cortex auditif de chaque hémisphère, par une population de neurones à sélectivité large et préférant les élévations basses.

Cette recherche constitue une première tentative d'observation de la plasticité corticale dans des tâches de localisation auditive par la neuroimagerie fonctionnelle. Elle a permis de divulguer certaines stratégies employées par le système auditif pour s'adapter à une perturbation des indices de localisation auditive et a montré que la plasticité pouvait être un puissant outil pour l'exploration des mécanismes neuronaux de la perception.

BIBLIOGRAPHIE

- Abbott, L., & Sejnowski, T. J. (1999). *Neural Codes and Distributed Representations : Foundations of Neural Computation*. MIT Press.
- Ahveninen, J., Kopčo, N., & Jääskeläinen, I. P. (2014). Psychophysics and neuronal bases of sound localization in humans. *Hearing Research*, 307, 86–97.
- Akeroyd, M. A. (2006). The psychoacoustics of binaural hearing : La psicoacústica de la audición binaural. *International journal of audiology*, 45(S1), 25–33.
- 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.
- Altmann, C. F., Wilczek, E., & Kaiser, J. (2009). Processing of Auditory Location Changes after Horizontal Head Rotation. *The Journal of Neuroscience*, 29(41), 13074–13078.
- Andéol, G., Macpherson, E. A., & Sabin, A. T. (2013). Sound localization in noise and sensitivity to spectral shape. *Hearing Research*, 304, 20–27.
- Andéol, G., Savel, S., & Guillaume, A. (2014). Perceptual factors contribute more than acoustical factors to sound localization abilities with virtual sources. *Auditory Cognitive Neuroscience*, 8, 451.
- Asano, F., Suzuki, Y., & Sone, T. (1990). Role of spectral cues in median plane localization. *The Journal of the Acoustical Society of America*, 88(1), 159–168.
- Ashida, G., & Carr, C. E. (2011). Sound localization : Jeffress and beyond. *Current Opinion in Neurobiology*, 21(5), 745–751.
- Bajo, V. M., Nodal, F. R., Moore, D. R., & King, A. J. (2010). The descending corticocollicular pathway mediates learning-induced auditory plasticity. *Nature neuroscience*, 13(2), 253–260.
- Batteau, D. W. (1967). The Role of the Pinna in Human Localization. *Proceedings of the Royal Society of London B : Biological Sciences*, 168(1011), 158–180.

- Bauer, R. W., Matuzsa, J. L., Blackmer, R. F., & Glucksberg, S. (1966). Noise localization after unilateral attenuation. *The Journal of the Acoustical Society of America*, 40, 441.
- Bazwinsky, I., Hilbig, H., Bidmon, H.-J., & Rübsamen, R. (2003). Characterization of the human superior olivary complex by calcium binding proteins and neurofilament H (SMI-32). *The Journal of Comparative Neurology*, 456(3), 292–303.
- Beitel, R. E., & Kaas, J. H. (1993). Effects of bilateral and unilateral ablation of auditory cortex in cats on the unconditioned head orienting response to acoustic stimuli. *Journal of Neurophysiology*, 70(1), 351–369.
- Belliveau, L. A. C., Lyamzin, D. R., & Lesica, N. A. (2014). The Neural Representation of Interaural Time Differences in Gerbils Is Transformed from Midbrain to Cortex. *Journal of Neuroscience*, 34(50), 16796–16808.
- Bergan, J. F., Ro, P., Ro, D., & Knudsen, E. I. (2005). Hunting increases adaptive auditory map plasticity in adult barn owls. *The journal of neuroscience*, 25(42), 9816–9820.
- Bernstein, L. R., van de Par, S., & Trahiotis, C. (1999). The normalized interaural correlation : accounting for NoS pi thresholds obtained with Gaussian and "low-noise" masking noise. *The Journal of the Acoustical Society of America*, 106(2), 870–876.
- Bizley, J. K., Nodal, F. R., Parsons, C. H., & King, A. J. (2007). Role of Auditory Cortex in Sound Localization in the Midsagittal Plane. *Journal of Neurophysiology*, 98(3), 1763–1774.
- Blauert, J. (1997). *Spatial hearing : the psychophysics of human sound localization* (Vol. 494). The MIT press, Cambridge, MA.
- Brainard, M. S., & Knudsen, E. I. (1993). Experience-dependent plasticity in the inferior colliculus : a site for visual calibration of the neural representation of auditory space in the barn owl. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, 13(11), 4589–4608.
- Brand, A., Behrend, O., Marquardt, T., McAlpine, D., & Grothe, B. (2002). Precise inhibition is essential for microsecond interaural time difference coding. *Nature*, 417(6888), 543–547.
- Bregman, A. S. (1994). *Auditory scene analysis : The perceptual organization of sound*. MIT press.

- Broadbent, D. E. (1954). The role of auditory localization in attention and memory span. *Journal of experimental psychology*, 47(3), 191.
- Bronkhorst, A. W. (2000). The Cocktail Party Phenomenon : A Review of Research on Speech Intelligibility in Multiple-Talker Conditions. *Acta Acustica united with Acustica*, 86(1), 117–128.
- Brugge, J. F., Reale, R. A., Hind, J. E., Chan, J. C. K., Musicant, A. D., & Poon, P. W. F. (1994). Simulation of free-field sound sources and its application to studies of cortical mechanisms of sound localization in the cat. *Hearing Research*, 73(1), 67–84.
- Brunetti, M., Belardinelli, P., Caulo, M., Del Gratta, C., Della Penna, S., Ferretti, A., ... Tartaro, A. (2005). Human brain activation during passive listening to sounds from different locations : An fMRI and MEG study. *Human brain mapping*, 26(4), 251–261.
- Brungart, D. S., & Rabinowitz, W. M. (1999). Auditory localization of nearby sources. Head-related transfer functions. *The Journal of the Acoustical Society of America*, 106(3), 1465–1479.
- Burke, K. A., Letsos, A., & Butler, R. A. (1994). Asymmetric performances in binaural localization of sound in space. *Neuropsychologia*, 32(11), 1409–1417.
- Butler, R. A. (1987). An analysis of the monaural displacement of sound in space. *Perception & Psychophysics*, 41(1), 1–7.
- Butler, R. A., Humanski, R. A., & Musicant, A. D. (1990). Binaural and monaural localization of sound in two-dimensional space. *Perception*, 19(2), 241–256.
- Byrne, D., & Noble, W. (1998). Optimizing Sound Localization with Hearing Aids. *Trends in Amplification*, 3(2), 51–73.
- Carlile, S. (2014). The plastic ear and perceptual relearning in auditory spatial perception. *Frontiers in Neuroscience*, 8.
- Carlile, S., Balachandar, K., & Kelly, H. (2014). Accommodating to new ears : The effects of sensory and sensory-motor feedback. *The Journal of the Acoustical Society of America*, 135(4), 2002–2011.

- Carlile, S., & Blackman, T. (2013). Relearning Auditory Spectral Cues for Locations Inside and Outside the Visual Field. *Journal of the Association for Research in Otolaryngology*, 15(2), 249–263.
- 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., Martin, R., & McAnally, K. (2005). Spectral Information in Sound Localization. In *International Review of Neurobiology* (Vol. Volume 70, pp. 399–434). Academic Press.
- Carr, C. E., & Konishi, M. (1988). Axonal delay lines for time measurement in the owl's brainstem. *Proceedings of the National Academy of Sciences of the United States of America*, 85(21), 8311–8315.
- Carr, C. E., & Konishi, M. (1990). A circuit for detection of interaural time differences in the brain stem of the barn owl. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, 10(10), 3227–3246.
- Casseday, J. H., & Covey, E. (1987). Central Auditory Pathways in Directional Hearing. In W. A. Yost & G. Gourevitch (Eds.), *Directional Hearing* (pp. 109–145). Springer US. doi: 10.1007/978-1-4612-4738-8_5
- Casseday, J. H., Fremouw, T., & Covey, E. (2002). The Inferior Colliculus : A Hub for the Central Auditory System. In D. Oertel, R. R. Fay, & A. N. Popper (Eds.), *Integrative Functions in the Mammalian Auditory Pathway* (pp. 238–318). Springer New York. doi: 10.1007/978-1-4757-3654-0_7
- Chang, E. F., & Merzenich, M. M. (2003). Environmental noise retards auditory cortical development. *Science (New York, N.Y.)*, 300(5618), 498–502.
- Chase, S. M., & Young, E. D. (2006). Spike-timing codes enhance the representation of multiple simultaneous sound-localization cues in the inferior colliculus. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, 26(15), 3889–3898.
- Cheng, C. I., & Wakefield, G. H. (1999). Introduction to Head-Related Transfer Functions (HRTFs) : Representations of HRTFs in Time, Frequency, and Space. Audio Engineering Society.

- Cherry, E. C. (1953). Some Experiments on the Recognition of Speech, with One and with Two Ears. *The Journal of the Acoustical Society of America*, 25(5), 975–979.
- Clifton, R. K., Gwiazda, J., Bauer, J. A., Clarkson, M. G., & Held, R. M. (1988). Growth in head size during infancy : Implications for sound localization. *Developmental Psychology*, 24(4), 477–483.
- Colburn, H. S., Isabelle, S. K., & Tollin, D. J. (1997). Modeling binaural detection performance for individual masker waveforms. *Binaural and spatial hearing in real and virtual environments*, 533–556.
- Comalli, P. E., & Altshuler, M. W. (1971). Effect of body tilt on auditory localization. *Perceptual and Motor Skills*, 32(3), 723–726.
- Cross, D. V. (1973). Sequential dependencies and regression in psychophysical judgments. *Perception & Psychophysics*, 14(3), 547–552.
- Davis, K. A., Ramachandran, R., & May, B. J. (2003). Auditory processing of spectral cues for sound localization in the inferior colliculus. *Journal of the Association for Research in Otolaryngology : JARO*, 4(2), 148–163.
- Deb, K. (2001). *Multi-objective optimization using evolutionary algorithms* (Vol. 2012). John Wiley & Sons Chichester.
- Dechent, P., Schönwiesner, M., Voit, D., Gowland, P., & Krumbholz, K. (2007). Basic coding mechanisms in the human auditory cortex – a view through high-resolution fMRI. In (p. 164). Chicago, USA.
- Delgutte, B., Joris, P. X., Litovsky, R. Y., & Yin, T. C. T. (1995). Relative importance of different acoustic cues to the directional sensitivity of inferior-colliculus neurons. *Advances in hearing research*, 288–299.
- Deouell, L. Y., Heller, A. S., Malach, R., D'Esposito, M., & Knight, R. T. (2007). Cerebral responses to change in spatial location of unattended sounds. *Neuron*, 55(6), 985–996.
- de Villers-Sidani, E., Chang, E. F., Bao, S., & Merzenich, M. M. (2007). Critical period window for spectral tuning defined in the primary auditory cortex (A1) in the rat. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, 27(1), 180–189.

- Dirks, D. D., & Wilson, R. H. (1969). The effect of spatially separated sound sources on speech intelligibility. *Journal of Speech, Language and Hearing Research*, 12(1), 5.
- Doucet, M.-E., Guillemot, J.-P., Lassonde, M., Gagné, J.-P., Leclerc, C., & Lepore, F. (2005). Blind subjects process auditory spectral cues more efficiently than sighted individuals. *Experimental Brain Research*, 160(2), 194–202.
- Edmonds, B. A., & Krumbholz, K. (2014). Are interaural time and level differences represented by independent or integrated codes in the human auditory cortex? *Journal of the Association for Research in Otolaryngology*, 15(1), 103–114.
- Ekstrom, A. (2010). How and when the fMRI BOLD signal relates to underlying neural activity : The danger in dissociation. *Brain research reviews*, 62(2), 233–244.
- Flannery, R., & Butler, R. A. (1981). Spectral cues provided by the pinna for monaural localization in the horizontal plane. *Perception & Psychophysics*, 29(5), 438–444.
- Florentine, M. (1976). Relation between lateralization and loudness in asymmetrical hearing losses. *Ear and Hearing*, 1(6), 243–251.
- Fujiki, N., Riederer, K. A. J., Jousmäki, V., Mäkelä, J. P., & Hari, R. (2002). Human cortical representation of virtual auditory space : differences between sound azimuth and elevation. *The European Journal of Neuroscience*, 16(11), 2207–2213.
- Gaese, B. H., & Johnen, A. (2000). Coding for auditory space in the superior colliculus of the rat. *The European Journal of Neuroscience*, 12(5), 1739–1752.
- Gardner, M. B. (1973). Some monaural and binaural facets of median plane localization. *The Journal of the Acoustical Society of America*, 54(6), 1489–1495.
- Gardner, M. B., & Gardner, R. S. (1973). Problem of localization in the median plane : effect of pinnae cavity occlusion. *The Journal of the Acoustical Society of America*, 53(2), 400–408.
- Glendenning, K. K., Baker, B. N., Hutson, K. A., & Masterton, R. B. (1992). Acoustic chiasm V : inhibition and excitation in the ipsilateral and contralateral projections of LSO. *The Journal of Comparative Neurology*, 319(1), 100–122.
- Goense, J. B. M., & Logothetis, N. K. (2008, May). Neurophysiology of the BOLD fMRI Signal in Awake Monkeys. *Current Biology*, 18(9), 631–640.

- Goodyear, B. G., & Menon, R. S. (1998). Effect of luminance contrast on BOLD fMRI response in human primary visual areas. *Journal of Neurophysiology*, 79(4), 2204–2207.
- Goossens, H. H. L. M., & Van Opstal, A. J. (1999). Influence of Head Position on the Spatial Representation of Acoustic Targets. *Journal of Neurophysiology*, 81(6), 2720–2736.
- Griffiths, T. D., & Hall, D. A. (2012, September). Mapping Pitch Representation in Neural Ensembles with fMRI. *Journal of Neuroscience*, 32(39), 13343–13347.
- Haeske-Dewick, H., Canavan, A. G., & Hömberg, V. (1996). Sound localization in egocentric space following hemispheric lesions. *Neuropsychologia*, 34(9), 937–942.
- Hall, D. A., Haggard, M. P., Akeroyd, M. A., Palmer, A. R., Summerfield, A. Q., Elliott, M. R., ... Bowtell, R. W. (1999). Sparse temporal sampling in auditory fMRI. *Human brain mapping*, 7(3), 213–223.
- Hancock, K. E., & Delgutte, B. (2004). A Physiologically Based Model of Interaural Time Difference Discrimination. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 24(32), 7110–7117.
- Harper, N. S., & McAlpine, D. (2004). Optimal neural population coding of an auditory spatial cue. *Nature*, 430(7000), 682–686.
- Harrington, I. A., Stecker, G. C., Macpherson, E. A., & Middlebrooks, J. C. (2008). Spatial sensitivity of neurons in the anterior, posterior, and primary fields of cat auditory cortex. *Hearing research*, 240(1-2), 22–41.
- Hartley, D., & King, A. J. (2010). Development of the auditory pathway. *The Oxford Handbook of Auditory Science : The Auditory Brain*, 2, 361–386.
- Hartley, D., & Moore, D. R. (2003). Effects of conductive hearing loss on temporal aspects of sound transmission through the ear. *Hearing Research*, 177(1-2), 53–60.
- Hartmann, W. M. (1983). Localization of sound in rooms. *The Journal of the Acoustical Society of America*, 74, 1380–1391.
- Hay, J. C., & Pick Jr., H. L. (1966). Visual and proprioceptive adaptation to optical displacement of the visual stimulus. *Journal of Experimental Psychology*, 71(1), 150–158.

- Heffner, H. E., & Heffner, R. S. (1990). Effect of bilateral auditory cortex lesions on sound localization in Japanese macaques. *Journal of Neurophysiology*, 64(3), 915–931.
- Held, R. (1955). Shifts in binaural localization after prolonged exposures to atypical combinations of stimuli. *The American Journal of Psychology*, 68(4), 526–548.
- Henning, G. B. (1974). Detectability of interaural delay in high-frequency complex waveforms. *The Journal of the Acoustical Society of America*, 55(1), 84–90.
- Hensch, T. K. (2004). Critical period regulation. *Annual Review of Neuroscience*, 27, 549–579.
- Hilbig, H., Beil, B., Hilbig, H., Call, J., & Bidmon, H.-J. (2009). Superior olivary complex organization and cytoarchitecture may be correlated with function and catarrhine primate phylogeny. *Brain Structure & Function*, 213(4-5), 489–497.
- Hirsh, I. J. (1950). The Relation between Localization and Intelligibility. *The Journal of the Acoustical Society of America*, 22(2), 196–200.
- Hofman, P. M., & Van Opstal, A. J. (1998). Spectro-temporal factors in two-dimensional human sound localization. *The Journal of the Acoustical Society of America*, 103(5), 2634–2648.
- Hofman, P. M., & Van Opstal, A. J. (2003). Binaural weighting of pinna cues in human sound localization. *Experimental Brain Research*, 148(4), 458–470.
- Hofman, P. M., Van Riswick, J. G., & Van Opstal, A. J. (1998). Relearning sound localization with new ears. *Nature neuroscience*, 1(5), 417–421.
- Hofman, P. M., Vlaming, M. S. M. G., Termeer, P. J. J., & Van Opstal, A. J. (2002). A method to induce swapped binaural hearing. *Journal of Neuroscience Methods*, 113(2), 167–179.
- Holland, J. H. (1975). *Adaptation in natural and artificial systems : An introductory analysis with applications to biology, control, and artificial intelligence*. U Michigan Press.
- Humphries, C., Liebenthal, E., & Binder, J. R. (2010, April). Tonotopic organization of human auditory cortex. *NeuroImage*, 50(3), 1202–1211.
- Imig, T. J., Bibikov, N. G., Poirier, P., & Samson, F. K. (2000). Directionality Derived From Pinna-Cue Spectral Notches in Cat Dorsal Cochlear Nucleus. *Journal of Neurophysiology*, 83(2), 907–925.

- Irvine, D. R., Park, V. N., & McCormick, L. (2001). Mechanisms underlying the sensitivity of neurons in the lateral superior olive to interaural intensity differences. *Journal of Neurophysiology*, 86(6), 2647–2666.
- Javer, A. R., & Schwarz, D. W. (1995). Plasticity in human directional hearing. *The Journal of otolaryngology*, 24(2), 111–117.
- Jeffress, L. A. (1948). A place theory of sound localization. *J. comp. physiol. Psychol*, 41(1), 35–39.
- Jeffress, L. A., & Taylor, R. W. (1961). Lateralization vs localization. *The Journal of the Acoustical Society of America*, 33(4), 482–483.
- Jenkins, W. M., & Masterson, R. B. (1982). Sound localization : effects of unilateral lesions in central auditory system. *Journal of Neurophysiology*.
- Jenkins, W. M., & Merzenich, M. M. (1984). Role of cat primary auditory cortex for sound-localization behavior. *Journal of Neurophysiology*, 52(5), 819–847.
- Jäncke, L., Buchanan, T., Lutz, K., Specht, K., Mirzazade, S., & Shah, N. J. (1999). The time course of the BOLD response in the human auditory cortex to acoustic stimuli of different duration. *Brain Research. Cognitive Brain Research*, 8(2), 117–124.
- Jäncke, L., Shah, N. J., Posse, S., Grosse-Ryken, M., & Müller-Gärtner, H. W. (1998). Intensity coding of auditory stimuli : an fMRI study. *Neuropsychologia*, 36(9), 875–883.
- Jäncke, L., Wüstenberg, T., Schulze, K., & Heinze, H. (2002). Asymmetric hemodynamic responses of the human auditory cortex to monaural and binaural stimulation. *Hearing Research*, 170(1-2), 166–178.
- Kacelnik, O., Nodal, F. R., Parsons, C. H., & King, A. J. (2006). Training-induced plasticity of auditory localization in adult mammals. *PLoS biology*, 4(4), e71.
- Kaiser, J., Lutzenberger, W., Preissl, H., Ackermann, H., & Birbaumer, N. (2000). Right-hemisphere dominance for the processing of sound-source lateralization. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, 20(17), 6631–6639.
- Karni, A., Meyer, G., Jezzard, P., Adams, M. M., Turner, R., & Ungerleider, L. G. (1995). Functional MRI evidence for adult motor cortex plasticity during motor skill learning. *Nature*, 377(6545), 155–158.

- Kavanagh, G. L., & Kelly, J. B. (1987). Contribution of auditory cortex to sound localization by the ferret (*Mustela putorius*). *Journal of neurophysiology*, 57(6), 1746–1766.
- Kidd, G. J., 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.
- Kilian-Hütten, N., Valente, G., Vroomen, J., & Formisano, E. (2011, February). Auditory Cortex Encodes the Perceptual Interpretation of Ambiguous Sound. *Journal of Neuroscience*, 31(5), 1715–1720.
- King, A. J., Dahmen, J. C., Keating, P., Leach, N. D., Nodal, F. R., & Bajo, V. M. (2011). Neural circuits underlying adaptation and learning in the perception of auditory space. *Neuroscience & Biobehavioral Reviews*, 35(10), 2129–2139.
- King, A. J., & Hutchings, M. E. (1987). Spatial response properties of acoustically responsive neurons in the superior colliculus of the ferret : a map of auditory space. *Journal of Neurophysiology*, 57(2), 596–624.
- King, A. J., Hutchings, M. E., Moore, D. R., & Blakemore, C. (1988). Developmental plasticity in the visual and auditory representations in the mammalian superior colliculus. *Nature*, 332(6159), 73–76.
- King, A. J., Parsons, C. H., & Moore, D. R. (2000). Plasticity in the neural coding of auditory space in the mammalian brain. *Proceedings of the National Academy of Sciences*, 97(22), 11821–11828.
- Knudsen, E. I. (1998). Capacity for plasticity in the adult owl auditory system expanded by juvenile experience. *Science (New York, N.Y.)*, 279(5356), 1531–1533.
- Knudsen, E. I. (2002). Instructed learning in the auditory localization pathway of the barn owl. *Nature*, 417(6886), 322–328.
- Knudsen, E. I., Esterly, S. D., & Knudsen, P. F. (1984). Monaural occlusion alters sound localization during a sensitive period in the barn owl. *The Journal of Neuroscience*, 4(4), 1001–1011.

- Knudsen, E. I., & Knudsen, P. F. (1989). Vision calibrates sound localization in developing barn owls. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, 9(9), 3306–3313.
- Knudsen, E. I., Knudsen, P. F., & Esterly, S. D. (1984). A critical period for the recovery of sound localization accuracy following monaural occlusion in the barn owl. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, 4(4), 1012–1020.
- Knudsen, E. I., & Mogdans, J. (1992). Vision-independent adjustment of unit tuning to sound localization cues in response to monaural occlusion in developing owl optic tectum. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, 12(9), 3485–3493.
- Kochkin, S. (1999). Reducing hearing instrument returns with consumer education. *Hearing Review*, 6(10), 18–20.
- Kochkin, S., Beck, D. L., Christensen, L. A., Compton-Conley, C., Fligor, B. J., Kricos, P. B., & Turner, R. G. (2010). MarkeTrak VIII : The impact of the hearing healthcare professional on hearing aid user success. *Hearing Review*, 17(4), 12–34.
- Kock, W. E. (1950). Binaural Localization and Masking. *The Journal of the Acoustical Society of America*, 22(6), 801–804.
- Krumbholz, K., Schönwiesner, M., Cramon, D. Y. v., Rübsamen, R., Shah, N. J., Zilles, K., & Fink, G. R. (2005). Representation of Interaural Temporal Information from Left and Right Auditory Space in the Human Planum Temporale and Inferior Parietal Lobe. *Cerebral Cortex*, 15(3), 317–324.
- Krumbholz, K., Schönwiesner, M., Rübsamen, R., Zilles, K., Fink, G. R., & Von Cramon, D. Y. (2005). Hierarchical processing of sound location and motion in the human brainstem and planum temporale. *European Journal of Neuroscience*, 21(1), 230–238.
- Kuhn, G. F. (1987). Physical acoustics and measurements pertaining to directional hearing. In *Directional hearing* (pp. 3–25). Springer.
- Kuk, F., & Korhonen, P. (2014). *Localization 101 : Hearing Aid Factors in Localization*.
- Kulesza, R. J., & Grothe, B. (2015). Yes, there is a medial nucleus of the trapezoid body in humans. *Frontiers in Neuroanatomy*, 9.

- Kumpik, D. P., Kacelnik, O., & King, A. J. (2010). Adaptive reweighting of auditory localization cues in response to chronic unilateral earplugging in humans. *The Journal of Neuroscience*, 30(14), 4883–4894.
- Köbler, S., & Rosenhall, U. (2002). Horizontal localization and speech intelligibility with bilateral and unilateral hearing aid amplification. *International Journal of Audiology*, 41(7), 395–400.
- Lackner, J. R. (1974). Changes in Auditory Localization During Body Tilt. *Acta Oto-Laryngologica*, 77(1-6), 19–28.
- Langendijk, E. H. A., & Bronkhorst, A. W. (2002). Contribution of spectral cues to human sound localization. *The Journal of the Acoustical Society of America*, 112(4), 1583–1596.
- Lesica, N. A., Lingner, A., & Grothe, B. (2010). Population coding of interaural time differences in gerbils and barn owls. *The Journal of Neuroscience*, 30(35), 11696–11702.
- Lessard, N., Paré, M., Lepore, F., & Lassonde, M. (1998). Early-blind human subjects localize sound sources better than sighted subjects. *Nature*, 395(6699), 278–280.
- Lewald, J., & Guski, R. (2003). Cross-modal perceptual integration of spatially and temporally disparate auditory and visual stimuli. *Cognitive brain research*, 16(3), 468–478.
- Lewald, J., Riederer, K. A. J., Lentz, T., & Meister, I. G. (2008). Processing of sound location in human cortex. *The European Journal of Neuroscience*, 27(5), 1261–1270.
- Linkenhoker, B. A., & Knudsen, E. I. (2002). Incremental training increases the plasticity of the auditory space map in adult barn owls. *Nature*, 419(6904), 293–296.
- Logothetis, N. K., Pauls, J., Augath, M., Trinath, T., & Oeltermann, A. (2001, July). Neurophysiological investigation of the basis of the fMRI signal. *Nature*, 412(6843), 150–157.
- Lupo, J. E., Koka, K., Thornton, J. L., & Tollin, D. J. (2011). The effects of experimentally induced conductive hearing loss on spectral and temporal aspects of sound transmission through the ear. *Hearing research*, 272(1-2), 30–41.
- MacKeith, N. W., & Coles, R. R. (1971). Binaural advantages in hearing of speech. *The Journal of Laryngology and Otology*, 85(3), 213–232.
- Macpherson, E. A. (2013). Cue weighting and vestibular mediation of temporal dynamics in sound localization via head rotation. *Proceedings of Meetings on Acoustics*, 19(1), 050131.

- Macpherson, E. A., & Middlebrooks, J. C. (2002). Listener weighting of cues for lateral angle : the duplex theory of sound localization revisited. *The Journal of the Acoustical Society of America*, 111, 2219.
- Magezi, D. A., & Krumbholz, K. (2010). Evidence for Opponent-Channel Coding of Interaural Time Differences in Human Auditory Cortex. *Journal of Neurophysiology*, 104(4), 1997–2007.
- Majdak, P., Baumgartner, R., & Laback, B. (2014). Acoustic and non-acoustic factors in modeling listener-specific performance of sagittal-plane sound localization. *Auditory Cognitive Neuroscience*, 5, 319.
- 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.
- Malhotra, S., Hall, A. J., & Lomber, S. G. (2004). Cortical control of sound localization in the cat : unilateral cooling deactivation of 19 cerebral areas. *Journal of neurophysiology*, 92(3), 1625–1643.
- Malhotra, S., Stecker, G. C., Middlebrooks, J. C., & Lomber, S. G. (2008). Sound Localization Deficits During Reversible Deactivation of Primary Auditory Cortex and/or the Dorsal Zone. *Journal of Neurophysiology*, 99(4), 1628–1642.
- Marcas, V. L., Straessle, A., Girard, F., Loenneker, T., & Martin, E. (2004). When more means less : a paradox BOLD response in human visual cortex. *Magnetic Resonance Imaging*, 22(4), 441–450.
- Masterton, B., & Diamond, I. T. (1973). Hearing- Central neural mechanisms. *Handbook of perception.*, 3, 407–448.
- Masterton, B., Diamond, I. T., Harrison, J. M., & Beecher, M. D. (1967). Medial superior olive and sound localization. *Science (New York, N.Y.)*, 155(3770), 1696–1697.

- McAlpine, D. (2005). Creating a sense of auditory space. *The Journal of physiology*, 566(1), 21–28.
- McAlpine, D., Jiang, D., & Palmer, A. R. (2001). A neural code for low-frequency sound localization in mammals. *Nature neuroscience*, 4(4), 396–401.
- McFadden, D., & Pasanen, E. G. (1976). Lateralization at high frequencies based on interaural time differences. *The Journal of the Acoustical Society of America*, 59(3), 634–639.
- McLaughlin, S. A., Higgins, N. C., & Stecker, G. C. (2016). Tuning to Binaural Cues in Human Auditory Cortex. *Journal of the Association for Research in Otolaryngology*, 17(1), 37–53.
- McPartland, J. L., Culling, J. F., & Moore, D. R. (1997). Changes in lateralization and loudness judgements during one week of unilateral ear plugging. *Hearing research*, 113(1), 165–172.
- Mendonca, C. (2014). A review on auditory space adaptations to altered head-related cues. *Frontiers in Neuroscience*, 8.
- Middlebrooks, J. C. (1997). Spectral shape cues for sound localization. In *Binaural and spatial hearing in real and virtual environments* (pp. 77–97). Lawrence Erlbaum Associates, Mahwah, NJ.
- 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), 1480–1492.
- Middlebrooks, J. C., & Bremen, P. (2013). Spatial Stream Segregation by Auditory Cortical Neurons. *Journal of Neuroscience*, 33(27), 10986–11001.
- Middlebrooks, J. C., & Green, D. M. (1991). Sound Localization by Human Listeners. *Annual Review of Psychology*, 42(1), 135–159.
- Middlebrooks, J. C., & Knudsen, E. I. (1984). A neural code for auditory space in the cat's superior colliculus. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, 4(10), 2621–2634.
- Middlebrooks, J. C., Xu, L., Eddins, A. C., & Green, D. M. (1998). Codes for sound-source location in nontopographic auditory cortex. *Journal of neurophysiology*, 80(2), 863–881.

- Miller, L. M., & Recanzone, G. H. (2009). Populations of auditory cortical neurons can accurately encode acoustic space across stimulus intensity. *Proceedings of the National Academy of Sciences of the United States of America*, 106(14), 5931–5935.
- Mills, A. W. (1958). On the Minimum Audible Angle. *The Journal of the Acoustical Society of America*, 30(4), 237–246.
- Mogdans, J., & Knudsen, E. I. (1992). Adaptive adjustment of unit tuning to sound localization cues in response to monaural occlusion in developing owl optic tectum. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, 12(9), 3473–3484.
- Mogdans, J., & Knudsen, E. I. (1993). Early monaural occlusion alters the neural map of interaural level differences in the inferior colliculus of the barn owl. *Brain Research*, 619(1-2), 29–38.
- Mohamed, F. B., Pinus, A. B., Faro, S. H., Patel, D., & Tracy, J. I. (2002). BOLD fMRI of the visual cortex : quantitative responses measured with a graded stimulus at 1.5 Tesla. *Journal of magnetic resonance imaging : JMRI*, 16(2), 128–136.
- Moore, J. K. (2000). Organization of the human superior olivary complex. *Microscopy Research and Technique*, 51(4), 403–412.
- Morimoto, M., & Ando, Y. (1980). On the simulation of sound localization. *Journal of the Acoustical Society of Japan (E)*, 1(3), 167–174.
- Mrsic-Flogel, T. D., King, A. J., & Schnupp, J. W. H. (2005). Encoding of Virtual Acoustic Space Stimuli by Neurons in Ferret Primary Auditory Cortex. *Journal of Neurophysiology*, 93(6), 3489–3503.
- Mukamel, R., Gelbard, H., Arieli, A., Hasson, U., Fried, I., & Malach, R. (2005). Coupling between neuronal firing, field potentials, and FMRI in human auditory cortex. *Science (New York, N.Y.)*, 309(5736), 951–954.
- 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.
- Nakahara, H., Zhang, L. I., & Merzenich, M. M. (2004). Specialization of primary auditory cortex processing by sound exposure in the "critical period". *Proceedings of the National Academy of Sciences of the United States of America*, 101(18), 7170–7174.

- Nir, Y., Fisch, L., Mukamel, R., Gelbard-Sagiv, H., Arieli, A., Fried, I., & Malach, R. (2007). Coupling between neuronal firing rate, gamma LFP, and BOLD fMRI is related to interneuronal correlations. *Current biology : CB*, 17(15), 1275–1285.
- Nodal, F. R., Kacelnik, O., Bajo, V. M., Bizley, J. K., Moore, D. R., & King, A. J. (2010). Lesions of the auditory cortex impair azimuthal sound localization and its recalibration in ferrets. *Journal of neurophysiology*, 103(3), 1209–1225.
- Oldfield, R. C. (1971). The assessment and analysis of handedness : the Edinburgh inventory. *Neuropsychologia*, 9(1), 97–113.
- Oldfield, S. R., & Parker, S. P. (1984a). Acuity of sound localisation : a topography of auditory space. II. Pinna cues absent. *Perception*, 13(5), 601–617.
- Oldfield, S. R., & Parker, S. P. (1984b). Acuity of sound localisation : a topography of auditory space. I. Normal hearing conditions. *Perception*, 13(5), 581–600.
- Oldfield, S. R., & Parker, S. P. (1986). Acuity of sound localisation : a topography of auditory space. III. Monaural hearing conditions. *Perception*, 15(1), 67–81.
- 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*, 14(2), 261–273.
- Palmer, A. R., & King, A. J. (1982). The representation of auditory space in the mammalian superior colliculus. *Nature*, 299(5880), 248–249.
- Palomäki, K. J., Tiitinen, H., Mäkinen, V., May, P. J., & Alku, P. (2005). Spatial processing in human auditory cortex : The effects of 3d, ITD, and ILD stimulation techniques. *Cognitive Brain Research*, 24(3), 364–379.
- Par, S. v. d., Trahiotis, C., & Bernstein, L. R. (2001). A consideration of the normalization that is typically included in correlation-based models of binaural detection. *The Journal of the Acoustical Society of America*, 109(2), 830–833.
- Parise, C. V., Knorre, K., & Ernst, M. O. (2014, April). Natural auditory scene statistics shapes human spatial hearing. *Proceedings of the National Academy of Sciences*, 111(16), 6104–6108.

- Park, T. J. (1998). IID sensitivity differs between two principal centers in the interaural intensity difference pathway : the LSO and the IC. *Journal of Neurophysiology*, 79(5), 2416–2431.
- Parseihian, G., & Katz, B. F. G. (2012). Rapid head-related transfer function adaptation using a virtual auditory environment. *The Journal of the Acoustical Society of America*, 131(4), 2948–2957.
- Pavani, F., Macaluso, E., Warren, J. D., Driver, J., & Griffiths, T. D. (2002). A common cortical substrate activated by horizontal and vertical sound movement in the human brain. *Current Biology*, 12(18), 1584–1590.
- 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.
- Polley, D. B., Steinberg, E. E., & Merzenich, M. M. (2006). Perceptual learning directs auditory cortical map reorganization through top-down influences. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, 26(18), 4970–4982.
- 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.
- Rademacher, J., Morosan, P., Schormann, T., Schleicher, A., Werner, C., Freund, H.-J., & Zilles, K. (2001, April). Probabilistic Mapping and Volume Measurement of Human Primary Auditory Cortex. *NeuroImage*, 13(4), 669–683.
- Rauschecker, J. P. (1999). Auditory cortical plasticity : a comparison with other sensory systems. *Trends in Neurosciences*, 22(2), 74–80.
- Raykar, V. C., Duraiswami, R., & Yegnanarayana, B. (2005, July). Extracting the frequencies of the pinna spectral notches in measured head related impulse responses. *The Journal of the Acoustical Society of America*, 118(1), 364–374.
- Razavi, B., O'Neill, W., & Paige, G. (2005, March). Both Interaural and Spectral Cues Impact Sound Localization in Azimuth. In *2nd International IEEE EMBS Conference on Neural Engineering, 2005. Conference Proceedings* (pp. 587–590).

- Recanzone, G. H. (1998). Rapidly induced auditory plasticity : The ventriloquism aftereffect. *Proceedings of the National Academy of Sciences*, 95(3), 869–875.
- Recanzone, G. H. (2000). Spatial processing in the auditory cortex of the macaque monkey. *Proceedings of the National Academy of Sciences*, 97(22), 11829–11835.
- Recanzone, G. H., Guard, D. C., Phan, M. L., & Su, T.-I. K. (2000). Correlation Between the Activity of Single Auditory Cortical Neurons and Sound-Localization Behavior in the Macaque Monkey. *Journal of Neurophysiology*, 83(5), 2723–2739.
- Recanzone, G. H., Schreiner, C. E., & Merzenich, M. M. (1993). Plasticity in the frequency representation of primary auditory cortex following discrimination training in adult owl monkeys. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, 13(1), 87–103.
- Recanzone, G. H., & Sutter, M. L. (2008). The Biological Basis of Audition. *Annual review of psychology*, 59.
- Roman, N., Wang, D., & Brown, G. J. (2003, October). Speech segregation based on sound localization. *The Journal of the Acoustical Society of America*, 114(4), 2236–2252.
- Röder, B., Teder-Sälejärvi, W., Sterr, A., Rösler, F., Hillyard, S. A., & Neville, H. J. (1999). Improved auditory spatial tuning in blind humans. *Nature*, 400(6740), 162–166.
- Salminen, N. H., Aho, J., & Sams, M. (2013). Visual task enhances spatial selectivity in the human auditory cortex. *Frontiers in Neuroscience*, 7.
- Salminen, N. H., May, P. J., Alku, P., & Tiitinen, H. (2009). A population rate code of auditory space in the human cortex. *PLoS One*, 4(10), e7600.
- Salminen, N. H., Tiitinen, H., & May, P. J. (2012). Auditory Spatial Processing in the Human Cortex. *The Neuroscientist*, 18(6), 602–612.
- Salminen, N. H., Tiitinen, H., Miettinen, I., Alku, P., & May, P. J. (2010). Asymmetrical representation of auditory space in human cortex. *Brain research*, 1306, 93–99.
- Salminen, N. H., Tiitinen, H., Yrttiaho, S., & May, P. J. C. (2010). The neural code for interaural time difference in human auditory cortex. *The Journal of the Acoustical Society of America*, 127(2), EL60–EL65.

- Savel, S. (2009). Individual differences and left/right asymmetries in auditory space perception. I. Localization of low-frequency sounds in free field. *Hearing Research*, 255(1–2), 142–154.
- Scharf, B. (1998). Auditory attention : The psychoacoustical approach. In H. Pashler (Ed.), *Attention* (pp. 75–117). Hove, UK : Psychology Press.
- Schechtman, E., Shrem, T., & Deouell, L. Y. (2012). Spatial Localization of Auditory Stimuli in Human Auditory Cortex is Based on Both Head-Independent and Head-Centered Coordinate Systems. *The Journal of Neuroscience*, 32(39), 13501–13509.
- Schnupp, J., Nelken, I., & King, A. (2011). *Auditory Neuroscience : Making Sense of Sound*. MIT Press.
- Schönwiesner, M., Dechent, P., Voit, D., Petkov, C. I., & Krumbholz, K. (2015). Parcellation of Human and Monkey Core Auditory Cortex with fMRI Pattern Classification and Objective Detection of Tonotopic Gradient Reversals. *Cerebral Cortex*, 25(10), 3278–3289.
- Schönwiesner, M., Krumbholz, K., Rübsamen, R., Fink, G. R., & von Cramon, D. Y. (2007). Hemispheric asymmetry for auditory processing in the human auditory brain stem, thalamus, and cortex. *Cerebral Cortex (New York, N.Y. : 1991)*, 17(2), 492–499.
- Schönwiesner, M., Voix, J., & Pango, P. (2009). Digital earplug for brain plasticity research. *Canadian Acoustics*, 37(3), 94–95.
- Seidl, A. H., Rubel, E. W., & Harris, D. M. (2010). Mechanisms for Adjusting Interaural Time Differences to Achieve Binaural Coincidence Detection. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 30(1), 70.
- Shinn-Cunningham, B. G., Durlach, N. I., & Held, R. M. (1998). Adapting to supernormal auditory localization cues. *The Journal of the Acoustical Society of America*, 103, 3656–3676.
- Singer, G., & Day, R. (1966). Spatial adaptation and aftereffect with optically transformed vision : Effects of active and passive responding and the relationship between test and exposure responses. *Journal of Experimental Psychology*, 71(5), 725–731.
- Siveke, I., Pecka, M., Seidl, A. H., Baudoux, S., & Grothe, B. (2006). Binaural response properties of low-frequency neurons in the gerbil dorsal nucleus of the lateral lemniscus. *Journal of neurophysiology*, 96(3), 1425–1440.

- Slattery, W. H., & Middlebrooks, J. C. (1994). Monaural sound localization : Acute versus chronic unilateral impairment. *Hearing Research*, 75(1–2), 38–46.
- Smith, A. L., Parsons, C. H., Lanyon, R. G., Bizley, J. K., Akerman, C. J., Baker, G. E., ... King, A. J. (2004). An investigation of the role of auditory cortex in sound localization using muscimol-releasing Elvax. *European Journal of Neuroscience*, 19(11), 3059–3072.
- Staeren, N., Renvall, H., Martino, F. D., Goebel, R., & Formisano, E. (2009, March). Sound Categories Are Represented as Distributed Patterns in the Human Auditory Cortex. *Current Biology*, 19(6), 498–502.
- Stecker, G. C., Harrington, I. A., & Middlebrooks, J. C. (2005). Location coding by opponent neural populations in the auditory cortex. *PLoS biology*, 3(3), e78.
- Steinhauser, A. (1879). The theory of binaural audition. A contribution to the theory of sound. *Philosophical Magazine Series 5*, 7(42), 181–197.
- Stern, R. M., & Shear, G. D. (1996). Lateralization and detection of low-frequency binaural stimuli : Effects of distribution of internal delay. *The Journal of the Acoustical Society of America*, 100(4), 2278–2288.
- Stevens, S. S., & Newman, E. B. (1936). The Localization of Actual Sources of Sound. *The American Journal of Psychology*, 48(2), 297–306.
- Strutt (Lord Rayleigh), J. W. (1907). On our perception of sound direction. *Philosophical Magazine*, 13, 214–232.
- Suga, N., Xiao, Z., Ma, X., & Ji, W. (2002). Plasticity and Corticofugal Modulation for Hearing in Adult Animals. *Neuron*, 36(1), 9–18.
- Syswerda, G. (1989). Uniform crossover in genetic algorithms. In (pp. 2–9). Morgan Kaufmann Publishers Inc.
- Thompson, S. P. (1877). On binaural audition. *Philosophical Magazine Series 5*, 4(25), 274–276.
- Thornton, J. L., Chevallier, K. M., Koka, K., Lupo, J. E., & Tollin, D. J. (2012). The Conductive Hearing Loss Due to an Experimentally Induced Middle Ear Effusion Alters the Interaural Level and Time Difference Cues to Sound Location. *JARO : Journal of the Association for Research in Otolaryngology*, 13(5), 641–654.

- Town, S. M., Brimijoin, W. O., & Bizley, J. K. (2017, June). Egocentric and allocentric representations in auditory cortex. *PLOS Biology*, 15(6), e2001878.
- Trapeau, R., Aubrais, V., & Schönwiesner, M. (2016). Fast and persistent adaptation to new spectral cues for sound localization suggests a many-to-one mapping mechanism. *The Journal of the Acoustical Society of America*, 140(2), 879–890.
- Trapeau, R., & Schönwiesner, M. (2011). Relearning sound localization with digital earplugs. *Journal of the Canadian Acoustical Association*, 39(3), 116–117.
- Trapeau, R., & Schönwiesner, M. (2015). Adaptation to shifted interaural time differences changes encoding of sound location in human auditory cortex. *NeuroImage*, 118, 26–38.
- Van der Maaten, L., & Hinton, G. (2008). Visualizing data using t-SNE. *Journal of Machine Learning Research*, 9(2579-2605), 85.
- Van Wanrooij, M. M., & Van Opstal, A. J. (2004). Contribution of Head Shadow and Pinna Cues to Chronic Monaural Sound Localization. *The Journal of Neuroscience*, 24(17), 4163–4171.
- Van Wanrooij, M. M., & Van Opstal, A. J. (2007). Sound Localization Under Perturbed Binaural Hearing. *Journal of Neurophysiology*, 97(1), 715–726.
- Van Wanrooij, M. M. V., & Van Opstal, A. J. (2005). Relearning sound localization with a new ear. *The Journal of neuroscience*, 25(22), 5413–5424.
- Vliegen, J., Grootel, T. J. V., & Van Opstal, A. J. (2004). Dynamic Sound Localization during Rapid Eye-Head Gaze Shifts. *The Journal of Neuroscience*, 24(42), 9291–9302.
- Vliegen, J., & Van Opstal, A. J. (2004). The influence of duration and level on human sound localization. *The Journal of the Acoustical Society of America*, 115(4), 1705–1713.
- Voix, J., & Laville, F. (2009). The objective measurement of individual earplug field performance. *The Journal of the Acoustical Society of America*, 125(6), 3722–3732.
- Wallach, H. (1939). On sound localization. *The Journal of the Acoustical Society of America*, 10(4), 270–274.
- Werner-Reiss, U., & Groh, J. M. (2008). A rate code for sound azimuth in monkey auditory cortex : implications for human neuroimaging studies. *The Journal of Neuroscience*, 28(14), 3747–3758.

- Wightman, F. L., & Kistler, D. J. (1989a). 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. (1989b). Headphone simulation of free-field listening. I : Stimulus synthesis. *The Journal of the Acoustical Society of America*, 85(2), 858–867.
- Wightman, F. L., & Kistler, D. J. (1997). Monaural sound localization revisited. *The Journal of the Acoustical Society of America*, 101, 1050.
- 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.
- Woldorff, M. G., Tempelmann, C., Fell, J., Tegeler, C., Gaschler-Markefski, B., Hinrichs, H., ... Scheich, H. (1999). Lateralized auditory spatial perception and the contralaterality of cortical processing as studied with functional magnetic resonance imaging and magnetoencephalography. *Human Brain Mapping*, 7(1), 49–66.
- Wong, L. L. N., Hickson, L., & McPherson, B. (2003). Hearing aid satisfaction : what does research from the past 20 years say ? *Trends in Amplification*, 7(4), 117–161.
- Woods, D. L., Stecker, G. C., Rinne, T., Herron, T. J., Cate, A. D., Yund, E. W., ... Kang, X. (2009). Functional Maps of Human Auditory Cortex : Effects of Acoustic Features and Attention. *PLoS ONE*, 4(4), e5183.
- Woods, T. M., Lopez, S. E., Long, J. H., Rahman, J. E., & Recanzone, G. H. (2006). Effects of Stimulus Azimuth and Intensity on the Single-Neuron Activity in the Auditory Cortex of the Alert Macaque Monkey. *Journal of Neurophysiology*, 96(6), 3323–3337.
- Worsley, K. J., Liao, C. H., Aston, J., Petre, V., Duncan, G. H., Morales, F., & Evans, A. C. (2002). A general statistical analysis for fMRI data. *Neuroimage*, 15(1), 1–15.
- Worsley, K. J., Marrett, S., Neelin, P., Vandal, A. C., Friston, K. J., & Evans, A. C. (1996). A unified statistical approach for determining significant signals in images of cerebral activation. *Human brain mapping*, 4(1), 58–73.

- Wright, B. A., & Fitzgerald, M. B. (2001). Different patterns of human discrimination learning for two interaural cues to sound-source location. *Proceedings of the National Academy of Sciences*, 98(21), 12307–12312.
- Wright, B. A., & Zhang, Y. (2006). A review of learning with normal and altered sound-localization cues in human adults : Revisión del aprendizaje en adultos con claves de localización sonora normales o alteradas. *International journal of audiology*, 45(S1), 92–98.
- Xu, L., Furukawa, S., & Middlebrooks, J. C. (1998). Sensitivity to Sound-Source Elevation in Nontopographic Auditory Cortex. *Journal of Neurophysiology*, 80(2), 882–894.
- Yan, W., & Suga, N. (1998). Corticofugal modulation of the midbrain frequency map in the bat auditory system. *Nature Neuroscience*, 1(1), 54–58.
- Yost, W. A. (1974). Discriminations of interaural phase differences. *The Journal of the Acoustical Society of America*, 55(6), 1299–1303.
- Yost, W. A., & Fay, R. R. (2007). *Auditory Perception of Sound Sources*. Springer Science & Business Media.
- Young, E. D., Spirou, G. A., Rice, J. J., Voigt, H. F., & Rees, A. (1992). Neural Organization and Responses to Complex Stimuli in the Dorsal Cochlear Nucleus [and Discussion]. *Philosophical Transactions of the Royal Society of London B : Biological Sciences*, 336(1278), 407–413.
- Young, P. T. (1928). Auditory localization with acoustical transposition of the ears. *Journal of Experimental Psychology*, 11(6), 399–429.
- Zatorre, R. J., Bouffard, M., Ahad, P., & Belin, P. (2002). Where is 'where' in the human auditory cortex? *Nature Neuroscience*, 5(9), 905–909.
- Zatorre, R. J., Bouffard, M., & Belin, P. (2004). Sensitivity to auditory object features in human temporal neocortex. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, 24(14), 3637–3642.
- Zatorre, R. J., & Penhune, V. B. (2001). Spatial localization after excision of human auditory cortex. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, 21(16), 6321–6328.

- Zhang, X., Zhang, Q., Hu, X., & Zhang, B. (2015). Neural representation of three-dimensional acoustic space in the human temporal lobe. *Frontiers in Human Neuroscience*, 9.
- Zimmer, U., Lewald, J., Erb, M., & Karnath, H.-O. (2006). Processing of auditory spatial cues in human cortex : an fMRI study. *Neuropsychologia*, 44(3), 454–461.
- Zwiers, M. P., Van Opstal, A. J., & Paige, G. D. (2003). Plasticity in human sound localization induced by compressed spatial vision. *Nature Neuroscience*, 6(2), 175–181.