UNIVERSITÉ PIERRE ET MARIE CURIE - PARIS 6

Rapport de stage

Master ATIAM 2012-2013 Mention : Sciences de l'ingénieur Spécialité : Psychophysique

VINCENT ISNARD

Mécanismes de perception auditive pour la localisation de sources sonores à travers les fréquences

31 juillet 2013

Encadrement : Romain Brette & Victor Benichoux







Université Pierre et Marie Curie - Paris 6 Bât. Esclangon, secrétariat du M2 Mécanique 2ème étage, porte 233 4, place Jussieu

75252 Paris cedex 05

Institut de Recherche et Coordination Acoustique/Musique 1, place Igor Stravinsky 75004 Paris

École Normale Supérieure

Département d'Études Cognitives, Équipe Audition 29, rue d'Ulm 75230 Paris cedex 05

Remerciements

Je remercie Romain Brette pour m'avoir accueilli dans son équipe et pour m'avoir permis de me rendre compte de l'ampleur de la tâche des neurosciences.

Je remercie également Victor Benichoux pour ses nombreux conseils, pour avoir su me redonner le cap quand ça débordait un peu du projet initial et pour m'avoir laissé déborder un peu quand même.

Et pour toute la vivacité qui régne dans ce labo, merci à : Romain, Victor, Marcel, Pierre, Betty & Giulietta, François, Aloïs, Martin, Bernhard, Dorothée, Trevor, Jane, Nelson, Yi, Claire 1, Claire 2, Claire 3, Tim, Jennifer, Marie, Yves, Daniel, et Christian.

Table des matières

L	exique	7		
Introduction				
Ι	Physiologie de l'audition	11		
1	Indices de localisation et perception de l'espace sonore	11		
2	Codage neurophysiologique	12		
II	Projet n°1 : Seuil fréquentiel de détection de l'IPD	15		
1	Postulats acoustiques versus neurophysiologiques	15		
2	Expérimentation 2.1 Procédure 2.2 Résultats 2.3 Discussion	16 16 17 19		
3	Test avec des plongeurs 3.1 Hypothèses 3.2 Résultats 3.3 Discussion	19 19 20 21		
4	Projets complémentaires 4.1 Apprentissage 4.1.1 Hypothèses 4.1.2 Dispositif expérimental proposé 4.2 Vision 4.2.1 Hypothèses 4.2.2 Résultats 4.2.3 Discussion	 22 22 22 22 22 22 23 24 		
5	Conclusion du projet n°1	24		
III Projet n°2 : Modèle unitaire de la localisation auditive : poids des ITDs à travers les fréquences 27				
1	Modèle bayésien 1.1 Pondération du poids des indices de localisation	27 27 27		

2	\mathbf{Exp}	érimentation	29	
	2.1	Prédictions	29	
	2.2	Protocole	29	
		2.2.1 Expérimentation pilote : matching	29	
		2.2.2 Paradigme 2AFC et méthode adaptative	30	
	2.3	Résultats	33	
	2.4	Discussion	36	
3	Con	clusion du projet n°2	36	
Conclusion				
Bibliographie				

Lexique

2AFC : Two-Alternative Forced Choice ANOVA : ANalysis Of VAriance BF : Basses Fréquences dB : décibel HF : Hautes Fréquences HRTF : Head-Related Transfer Function ILD : Interaural Level Difference IPD : Interaural Phase Difference ITD : Interaural Time Difference JND : Just-Noticeable Difference MAA : Minimum Audible Angle MF : Moyennes Fréquences PA : Potentiel d'Action

Introduction

G. T. Fechner posa en 1860, dans son ouvrage *Elemente der Psychophysik*, les bases de la psychophysique comme étant la « science exacte des relations fonctionnelles ou relations de dépendance entre le corps et l'esprit ». Depuis, un grand nombre d'expérimentations ont permis de cerner les problématiques majeures de la perception, et de poser précisément les cadres d'étude pour la compréhension du système perceptif humain. Mais l'apparition des neurosciences a bouleversé le schème psychophysique en proposant une investigation au sein même du système nerveux. Cette démarche peut toutefois être remise en cause lorsqu'on met en évidence la complexité inhérente au système nerveux, à laquelle doivent faire face les neurosciences. C'est donc pour garder un point de vue d'ensemble, en plus d'une recherche approfondie au cœur des neurones, que l'équipe Audition du Département d'études cognitives (DEC) de l'École normale supérieure (ENS) s'attache à traiter des problématiques allant de la psychophysique à la modélisation et aux neurosciences en lien avec nos capacités de perception des sons.

La localisation des sons, en particulier, est une problématique majeure dans la mesure où il s'agit d'exprimer et d'expliquer quels sont les indices dans un stimulus qui nous permettent de déduire la localisation d'une source placée dans l'espace. Nos observations doivent donc se focaliser sur les transformations acoustiques générées par le trajet de l'onde sonore jusqu'à la cochlée, et sur le traitement de l'information à l'intérieur du système auditif. Nous présentons ici deux projets que nous avons mené durant ce stage : un premier projet sur les différences interaurales de phase (IPDs : Interaural Phase Differences), et second projet sur les différences interaurales de temps (ITDs : Interaural Time Differences). Ces deux projets sont très liés mais visent à montrer deux aspects de la perception auditive en localisation. Dans le premier cas, il s'agit de chercher à montrer si nos capacités de localisation auditive dépendent de paramètres acoustiques, donc morphologiques, ou bien si nos limites de perception sont à attribuer davantage à notre constitution neurophysiologique. En ce qui concerne le second projet, on s'attache cette fois-ci aux différences interaurales de temps, afin de construire un modèle unitaire de localisation, suite à la modélisation introduite par Rayleigh en 1907, qui dissocie au contraire notre perception en deux indices, différences interaurales de temps et différences interaurales de niveau (ILDs : Interaural Level Differences).

Première partie Physiologie de l'audition

1 Indices de localisation et perception de l'espace sonore

D'un point de vue acoustique, le problème de la localisation revient à s'interroger sur les paramètres clefs qui fait qu'on perçoit une source placée dans l'espace simplement parce que nous sommes placés nous-mêmes dans cet espace et que nous disposons de récepteurs sensibles à cette relation entre la source sonore et nous. Ainsi, on peut considérer qu'une source sonore émet en vibrant un certain contenu spectral dans le temps. Nos deux oreilles sont disposées de telle manière qu'elles vont capter les ondes sonores émises par une source latéralisée avec une différence de temps (ITD) et d'intensité (ILD), constituant ainsi deux indices robustes de localisation en azimut (cf. Fig. 1). L'azimut est le paramètre qui va nous intéresser en particulier au cours de cette étude car c'est celui pour lequel on a le plus d'informations et qui peut nous permettre d'aborder la neurophysiologie avec des connaissances du principe de fonctionnement général déjà bien établies.



FIGURE 1 – Les différences interaurales.

Ces différences interaurales sont réparties sur l'axe des fréquences, et l'on considère, en suivant le modèle de la théorie duplex développée par Lord Rayleigh [Ray07], que les différences interaurales de temps sont perceptibles en basses fréquences (f < 1.5 kHz), et que les différences interaurales de niveaux sont perceptibles en hautes fréquences (f > 1.5 kHz). La théorie duplex décrit le retard et le filtrage que subit une source sonore lorsqu'elle atteint les oreilles d'un auditeur, pour générer respectivement des différences de temps et des différences d'intensité. Il est ainsi possible de reproduire artificiellement la spatialisation de sources sonores à partir de différences de temps et/ou d'intensité. Grâce à cela, on est en mesure de quantifier l'influence de chacun de ces paramètres en fonction des fréquences. Or, c'est justement ce qui va nous permettre de relativiser la limite fréquentielle théorique de 1.5 kHz. D'une part ces indices n'ont pas de bornes strictes, comme on peut le constater par exemple avec les ILDs sur la Figure 2, et d'autre part ils évoluent en interactions, des ITDs pouvant compenser des ILDs et inversement (cf. Fig. 3).

On se concentrera dans cette étude sur les différences interaurales de temps, et les différences interaurales de phase qui en sont directement issues en ce qui concerne des signaux périodiques, pour observer à travers plusieurs expérimentations comment ces indices évoluent le long des fréquences et ce qui en résulte pour notre perception, et comment on peut en tirer une information sur la latéralisation des sources sonores. On verra que ces indices varient en fonction de



FIGURE 2 – Evolution des ILDs en fonction des fréquences, d'après [Har60].



FIGURE 3 – ITD pour annuler une ILD donnée, représentée en fonction de l'ILD, d'après [Har60].

la fréquence, et l'on cherchera à déterminer comment le système auditif en tire l'informativité nécessaire à la localisation des sources sonores.

2 Codage neurophysiologique

La cochlée, au cœur de l'oreille interne (cf. Fig. 4), a pour fonction de transmettre les informations sonores au nerf auditif, par transduction des vibrations acoustiques en potentiels d'action (PA) électro-chimiques, qui sont les entités élémentaires du codage de l'information par le système nerveux. A l'intérieur de la cochlée, c'est la membrane basilaire qui réalise ce codage grâce à sa constitution mécanique. Elle permet en effet de réaliser une répartition tonotopique des fréquences : hautes au niveau de la base, basses au niveau de l'apex. Mais la répartition en fréquence n'est linéaire que jusqu'à 500 Hz, et devient logarithmique au-dessus de 500 Hz. Par ailleurs, la membrane basilaire se comporte comme un banc de filtres qui se recouvrent sous l'effet d'excitations acoustiques. L'échelle ERB (Equivalent Rectangular Bandwidth) permet d'approximer ces bandes critiques comme des filtres équivalents rectangulaires [MG83]. La largeur d'un filtre cochléaire sur l'échelle ERB est déterminée avec la formule suivante :



FIGURE 4 – Le système auditif humain

$$ERB(f_c) = 0.108 \cdot f_c + 24,7 \tag{1}$$

où $ERB(f_c)$ est exprimé en Hz, avec f_c la fréquence centrale du filtre.

Durant une phase d'excitation d'un filtre cochléaire, le liquide cochléaire entraîne le mouvement des cellules ciliées internes (CCI) de l'organe de Corti (cf. Fig. 5) qui se dépolarisent au passage des ions potassium dans la cellule. L'impulsion nerveuse, ou potentiel d'action, est transmise par le nerf auditif au cerveau. Les cellules qui émettent les potentiels d'action ont certaines contraintes temporelles liées à leur constitution électro-chimique. En particulier, leur période réfractaire, qui correspond à un temps de rechargement pour pouvoir émettre à nouveau un potentiel d'action, et qui limite le taux de décharge à 1000 PAs/s, selon l'estimation de [Mol99]. Ainsi, la technique de codage des tons purs dépend de la fréquence : en BF (f < 1kHz), le codage est temporel, c'est-à-dire que la fréquence transmise acoustiquement jusqu'à la membrane basilaire est codée par des PAs à la même fréquence que la fréquence acoustique; en MF (1 < f < 4 kHz), on parle de « volleying », comme l'on dispose en fait de 10 à 30 fibres nerveuses (95% de fibres afférentes) par CCI, les neurones se répartissent la tâche en émettant des PAs à une fréquence multiple de la fréquence acoustique, et la somme des contributions de chaque fibre permet de reconstituer une fréquence de PAs égale à la fréquence acoustique, la synchronisation est conservée entre la fréquence acoustique et la fréquence de décharge des neurones (cf. Fig. 6); en HF (f > 4 kHz), le système auditif profite de la répartition tonotopique des fréquences le long de la membrane basilaire pour le codage des vibrations acoustiques.



FIGURE 6 – Synchronisation de réponses d'une fibre cochléaire à des tons basses fréquences, illustrée par un enregistrement de la réponse d'une fibre d'un cochon d'Inde à un ton de fréquence 300 Hz et à un niveau faible (proche du seuil), d'après [APS71].



 ${\rm FIGURE}$ 5 – L'organe de Corti, posé sur la membrane basilaire, permet de transformer l'information acoustique des vibrations de la membrane en potentiels d'action.

Deuxième partie Projet nº1 : Seuil fréquentiel de détection de l'IPD

1 Postulats acoustiques versus neurophysiologiques

Positionnement du problème. Le problème de la localisation revient à s'interroger sur les paramètres clefs qui font que l'on perçoit une source placée dans l'espace simplement parce que nous sommes placés nous-mêmes dans cet espace et que nous disposons de récepteurs sensibles à cette relation entre la source sonore et nous. On considère qu'une source sonore émet en vibrant un certain contenu spectral dans le temps. Nos deux oreilles sont disposées de telle manière qu'elles vont capter les ondes émises par une source latéralisée avec une différence de temps (ITD : Interaural Time Difference) en basses fréquences (f < 1,5 kHz) et d'intensité (ILD : Interaural Level Difference) en hautes fréquences (f > 1,5 kHz), comme le décrit la théorie duplex, constituant ainsi deux indices robustes de localisation auditive répartis sur l'axe des fréquences. On s'intéresse ici plus précisément aux différences interaurales de phase, qui sont une manière d'aborder les différences interaurales de temps, et l'on s'interroge sur la cause de la limite fréquentielle supérieure de perception de cet indice : est-elle d'origine acoustique ou neurophysiologique?

Postulats acoustiques. D'un point de vue acoustique, la limite de perception de l'IPD serait due à notre morphologie, c'est-à-dire basiquement à la taille de notre tête. Si l'on considère une onde plane sinusoïdale de pulsation ω , plus la taille de la tête est importante, plus un son latéralisé met du temps à atteindre une oreille puis l'autre, et un biais de phase serait introduit avec des fréquences suffisamment élevées puisqu'on sortirait plus rapidement de l'intervalle $[-\pi; \pi]$ en termes de différences interaurales de phase. Les résultats pour un modèle de tête sphérique [Kuh77] permettent de prédire des ITDs constants jusqu'à 600 Hz, via la relation $ITD = IPD/\omega$, puis des variations en fonction de la fréquence *a priori* difficilement exploitables, et l'on se reporte généralement ensuite sur le retard d'enveloppe pour expliquer l'utilisation de l'indice d'ITD. Mais la limite théorique d'un biais de phase, à partir de laquelle on sortirait de l'intervalle $[-\pi; \pi]$, selon un modèle de tête sphérique, est autour de 1,5 kHz en fonction du rayon de la sphère. Ces résultats font abstraction d'une limite de l'encodage de la phase par le système nerveux, puisqu'elles se basent uniquement sur des considérations acoustiques.

Postulats neurophysiologiques. D'un point de vue neurophysiologique, le codage de l'information de localisation se fait dans la cochlée, par transduction des informations acoustiques en informations électrochimiques via les potentiels d'action. Ainsi, en supposant implicitement un codage temporel, la technique de codage par *phase-locking* dépend de la fréquence (cf. [Ric10]). Il s'agit, en basses fréquences, de faire une correspondance entre les fréquences des vibrations acoustiques et les fréquences de décharge des neurones.

Mise à l'épreuve des hypothèses. Des données sur la localisation auditives de plusieurs espèces mammifères [HH92a] montrent une corrélation avec la fréquence limite d'ambiguité de phase calculée à partir de la taille de la tête des animaux. La formule employée pour calculer la fréquence limite de détection de l'IPD est la suivante :

$$f_{lim}^{IPD} = \frac{1}{6 \cdot (a/c) \cdot \sin \theta} \tag{2}$$

où a est le rayon de la tête, c est la vitesse du son, et θ est l'angle de la source à partir du plan médian de l'animal.

Ces résultats montrent que la limite de perception de la phase est physique avant d'être neurophysiologique, et motivent notre démarche en ce qui concerne les humains puisque cela permettrait de montrer le potentiel du système nerveux et de redéfinir les problématiques neuroscientifiques vers une plasticité de celui-ci, plutôt que de poser des limites trop bas niveau. Autrement dit, chaque indice de localisation serait exploité selon sa rentabilité, et au-dessus de 1.5 kHz, notre mécanisme de localisation des sources sonores se reporterait sur les ILDs par souci de rentabilité. Cependant, la variabilité des tailles de tête est moins importante que celle mesurée entre différents animaux, et la significativité des résultats peut s'avérer plus difficile à atteindre. C'est pourquoi, il s'agit pour nous de maximiser les possibilités de trouver une corrélation entre des données morphologiques et des données psychophysiques en choisissant les paramètres appropriés dans le cadre d'une expérimentation sur la fréquence limite de détection de l'IPD.

2 Expérimentation

2.1 Procédure

Protocole expérimental. La procédure employée ici s'inspire de celle employée par Grose & Mamo dans [GM10], qui visait à évaluer le codage de la structure fine en fonction de l'âge. Puisqu'on s'intéresse aux différences de phase, c'est bien le codage de la structure fine de signaux qui nous intéresse également pour notre étude. Il s'agit de mesurer notre capacité à détecter une impression de spatialisation au casque, pour s'affranchir de l'espace acoustique naturel et mesurer directement nos capacités de manière isolée, pour un niveau d'écoute normal (65 dBSPL), dans une cabine d'écoute (acoustique atténuée). La procédure de test consiste à écouter successivement trois sons purs modulés en amplitude, d'une durée de 800 ms. L'un des trois, le stimulus cible, est différent des deux autres car on y aura introduit une différence de phase sensée générer cette impression de spatialisation. La modulation est de 100% en amplitude, pour une fréquence de 5 Hz, générant ainsi 4 périodes de modulation pour la durée d'un stimulus. La 2e et la 4e modulation du stimulus cible, sensé générer cette impression de spatialisation, sont en opposition de phase entre les deux canaux (cf. Fig. 7).



FIGURE 7 – Organisation temporelle des stimuli cible (à gauche) et référence (à droite), en fonction de l'IPD par période de modulation, d'après [GM10].

La tâche consiste à écouter les trois stimuli, qui sont diffusés dans un ordre aléatoire à chaque essai, et à cliquer sur l'un des trois boutons correspondants aux trois sons pour indiquer lequel semble différent des deux autres. Le participant dispose d'un feedback visuel lorsqu'il effectue la tâche.

La méthode employée est adaptative : la fréquence de départ est de 1000 Hz, si le sujet détecte correctement trois fois de manière consécutive le stimulus contenant une différence de phase, alors la fréquence porteuse est multipliée par un certain facteur, s'il se trompe, la fréquence porteuse est divisée par le même facteur. Ce facteur dépend du nombre d'inversions dans le sens de variation de la fonction des réponses : une inversion correspond à trois réponses correctes puis une réponse fausse, ou une réponse fausse puis trois réponses correctes (cf. Fig. 8). Au départ, le pas de modification de la fréquence porteuse est de 1/4 d'octave, puis au bout d'une inversion, il passe à 1/8 d'octave, puis au bout de 8 inversions, il passe à 1/12 d'octave. Le test se termine au bout de 15 inversions. On calcule la fréquence limite de détection de l'IPD en faisant la moyenne des fréquences porteuses au niveau des 6 dernières inversions.



FIGURE 8 – Allure de l'évolution de la fréquence porteuse des stimuli en fonction des réponses du sujet à chaque essai.

Stimuli. Les stimuli sont des sinus modulés en amplitude par un sinus à 5 Hz, à fréquence porteuse variable, d'une durée de 0.8 seconde, c'est-à-dire générant quatre modulations de la fréquence porteuse par stimulus. En ce qui concerne le stimulus cible, la modulation permet de dissimuler la transition vers l'opposition de phase sur les deuxième et quatrième modulations, sans générer d'indice externe quant à la présence de l'opposition de phase. Présenter deux fois sur quatre l'opposition de phase rallonge le temps des stimuli, mais simplifie la tâche en proposant une comparaison avec et sans opposition de phase.

Dispositif expérimental. L'interface de test a été programmée pour fonctionner sous Matlab. Les sons sont émis à travers un casque Sennheiser HD600 (circumaural, ouvert), étalonné à 65 dBSPL en moyenne entre 1000 Hz et 2000 Hz, grâce à une oreille artificielle Bruël & Kjær Type 4153 reliée au sonomètre Type 2250. Le casque est connecté à une carte son RME Fireface 400.

2.2 Résultats

Un panel de 17 sujets âgés de moins de 30 ans a participé à l'expérimentation. Tous les sujets ont passé un audiogramme au moment de l'expérimentation pour vérifier qu'ils étaient normoentendants entre 500 et 4000 Hz, c'est-à-dire que leur seuil de perception était inférieur à 20 dBHL. La mesure de la distance interaurale a été effectuée à l'aide d'un compas anthropométrique à bout olivaires GPM. Une mesure du tour de la tête a également été réalisée à l'aide d'un mètre couturier, au niveau du front. En plus de ces mesures, chaque sujet a dû indiquer plusieurs informations susceptible d'influencer ses résultats : son âge, s'il a une pratique de la musique, sa vue (cf. §II.4.2). Tous les sujets ont été rémunérés pour effectuer cette expérimentation.

Les résultats sont affichés sur la Figure 9 : on y présente la fréquence limite de détection de l'IPD en fonction de la fréquence théorique calculée à partir de la taille de la tête du sujet,

d'après l'Equation 2. Chaque fréquence limite de détection de l'IPD est calculée en faisant la moyenne des 3 meilleurs résultats sur 4 passages du test au total. On cherche ainsi à limiter la variabilité intra-individuelle due notamment aux défauts de concentration.



FIGURE 9 – Fréquence limite de détection de l'IPD (en Hz) pour des sons purs, en fonction de la fréquence limite théorique calculée à partir de la mesure de la distance interaurale d = 2a (en haut), et de la mesure du tour de la tête $T = 2\pi a$ (en m) (en bas), d'après l'Equation 2, pour $c = 342 \text{ m.s}^{-1}$ et $\theta = \pi/2$. Les barres d'erreur indiquent les écart-types. En vert : sujets non-musiciens; en rouge : sujets musiciens.

ANOVA. Réaliser une analyse de la variance (ANOVA) permet de déterminer la significativité de l'effet des variables expérimentales dans notre expérimentation. Concernant le facteur musical (musiciens vs. non-musiciens), elle vient ici confirmer ce qu'on pouvait observer sur la Figure 8, à savoir que ce groupe ne permet pas de générer de différences significatives dans notre expérimentation, en donnant une valeur p de 0.36 et ne nous permettant pas de rejetter l'hypothèse nulle.

2.3 Discussion

Ces premiers résultats ne montrent pas clairement une corrélation entre la fréquence limite de détection de l'IPD et la taille de la tête. Pour expliquer ces résultats vis-à-vis de nos hypothèses de départ, on peut invoquer le fait que les extremums des tailles de têtes ne sont pas très éloignés (de 12.2 à 14.2 cm), tandis que la mesure de la taille de la tête n'est pas particulièrement fiable (élasticité de la peau, points de mesure). D'autre part, on peut s'interroger sur la variabilité intra-individuelle de ces résultats : à quoi sont dues les différences de résultats pour un même sujet et comment pourrait-on les optimiser? L'attention semble être un premier facteur crucial pour parvenir à déceler la présence de différences de phase dans le stimulus cible. Certaines techniques permettent de focaliser l'attention, comme les saccades occulaires (cf. [Des04]), mais elles déplaceraient la problématique étudiée et c'est pourquoi nous n'avons pas donné de conseil particulier aux participants, sinon en donnant une brève description du test, où le but est d'atteindre le meilleur score.

Reformulation des hypothèses. L'hypothèse de départ visait à remettre en cause celle concernant la capacité limitée de nos neurones à déclencher des potentiels d'action à des fréquences élevées, à cause de leur période réfractaire. Il faudrait en déduire que c'est le hasard qui fait que cette fréquence limite de décharge est justement proche de la fréquence limite théorique de détection de l'IPD, calculée à partir de la taille de la tête. Face à cela, l'hypothèse qui a motivé le test garde encore de la consistance, en particulier en ce qui concerne la plasticité du système nerveux. De plus, nos statistiques peuvent être contournées si l'on considère que notre panel de sujets n'est pas suffisant. En effet, on a évoqué la faible discriminabilité des tailles de têtes, mais cette mesure serait aussi à reconsidérer si par exemple on pouvait mettre à jour l'existence d'une période critique dans la formation du système auditif humain, comme en ce qui concerne l'oreille absolue. Selon [LR05], il n'existe pas de cas d'adultes acquérant l'oreille absolue, et la période critique pour l'acquérir serait de 9 ans. L'étude [SG05] montre qu'il existe une période critique chez le gerbille en ce qui concerne la localisation auditive. Or, pour les humains, le développement de la tête atteint les 90% de la taille adulte à l'âge de 6 ans (cf. [Gou98]), la variabilité de nos résultats s'expliquerait donc par la variabilité dans le développement restant de la tête entre cette période critique (enfance) et le moment de passage de l'expérimentation (adulte). Pour s'affranchir le plus possible de cette variabilité, il s'agirait de faire passer le test sur un panel beaucoup plus important de sujets pour s'affranchir au maximum de cette variabilité, ou bien de faire passer le test à des jeunes participants. En tout cas, au vu des résultats qui ont conduit à poser les hypothèses pour cette expérimentation, et celles qui restent à vérifier, il est difficile de tirer une conclusion franche à ce stade de notre étude.

3 Test avec des plongeurs

3.1 Hypothèses

Selon Savel [Sav01], il existe des capacités chez l'humain permettant de localiser des sons sous l'eau jusqu'à 6 kHz. De plus, Savel met en évidence que les sons sous l'eau sont transmis essentiellement par conduction osseuse à cause de l'impédance de l'eau de mer, 3500 fois plus importante que celle de l'air, et conduisant à la perte des mécanismes d'amplification de l'oreille : perte des résonances du pavillon, du conduit auditif, et par suite de l'effet de levier des osselets. Enfin, la vitesse de propagation du son sous l'eau est 4 fois plus élevée que dans l'air, donc la longueur d'onde des signaux est 4 fois plus élevée, d'après la relation entre la longueur d'onde et la vitesse du son :

$$\lambda = \frac{c}{f} \tag{3}$$

avec λ la longueur d'onde (en m), c la vitesse du son (en m.s⁻¹), f la fréquence (en Hz). En posant que les atténuations dans le crâne sont faibles, les ILDs deviennent négligeables, et la localisation des sons jusqu'à 6 kHz doit être issue de la détection de différences temporelles uniquement. On peut donc relier la fréquence limite de localisation des sons à 6 kHz à la détection des différences interaurales de temps jusqu'à 6 kHz. Ces conclusions tendraient à montrer que la limite fréquentielle de détection des IPDs n'est pas due à des contraintes neuronales, mais qu'un apprentissage est possible pour des fréquences élevées.

3.2 Résultats

Le protocole utilisé pour cette expérimentation était exactement celui décrit précédemment pour les participants non-plongeurs. Un panel de 5 sujets plongeurs a participé à cette expérimentation. Tous les sujets ont passé un audiogramme, tous les sujets étaient normoentendants entre 500 et 8000 Hz, sauf un présentant des pertes à 4000 et 6000 Hz, et un autre présentant des pertes à 6000 Hz, pour des âges compris entre 27 et 47 ans. Tous pouvaient attester de niveaux moyens et élevés en plongée : 3 participants avaient respectivement les niveaux N2, N3 et N4, les niveaux allant de 1 à 4, de l'initiation à l'encadrement ; les deux autres participants étaient moniteurs de plongée. Tous les sujets ont été rémunérés pour participer à l'expérimentation.



FIGURE 10 – Fréquence limite de détection de l'IPD (en Hz) pour des sons purs, en fonction de la fréquence limite théorique calculée à partir de la mesure de la distance interaurale d = 2a, d'après l'Equation 2, pour $c = 342 \text{ m.s}^{-1}$ et $\theta = \pi/2$. Les barres d'erreur indiquent les écart-types. En vert : sujets non-plongeurs ; en rouge : sujets plongeurs.

3.3 Discussion

Les tests effectués avec des plongeurs ne révèlent pas de capacités développées spécifiquement par les plongeurs pour la détection d'IPDs à des fréquences élevées, et l'on ne trouve pas de corrélation entre la fréquence maximale de détection de l'IPD et la taille de la tête pour des participants plongeurs. Les résultats obtenus avec les plongeurs sont problématiques car ils ne remettent pas en cause directement nos hypothèses mais nous conduisent bien plutôt à questionner la compatibilité de ces résultats avec ceux des études effectuées sous l'eau. Tout d'abord, les études qui montrent les capacités des plongeurs à localiser des sons sous l'eau (cf. [Sav01], [Ber03]) : quels indices sont employés si ce ne sont pas les indices temporels? L'extension de la longueur d'onde sous l'eau et la conduction osseuse sont résolument consistants. En ce qui concerne la conduction osseuse toutefois, l'étude [RG08] avance la possibilité d'une conversion des ITDs en ILDs dans le crâne, à cause d'interférences acoustiques. Cette dernière hypothèse expliquerait de manière directe les résultats obtenus par notre expérimentation ainsi que ceux obtenus sous l'eau, puisque la localisation des sons en hautes fréquences serait rétablie avec des différences d'intensités sous l'eau, sans hypothèse supplémentaire d'un apprentissage particulier pour les sujets plongeurs. Une étude acoustique sur la conduction solidienne dans le crâne serait à fournir pour venir compléter cette étude. D'un point de vue perceptif, une expérimentation complémentaire serait à effectuer en remplaçant le casque audio par des vibrateurs. On pourrait ainsi comparer les deux modes d'écoute : par conduction aérienne versus par conduction osseuse. Il s'agirait de placer des vibrateurs au niveau des mandibules à la place du casque audio. On s'attend, pour des sujets quelconques, à observer une fréquence limite de détection d'IPDs supérieure avec les vibrateurs qu'avec le casque audio. Cela permettrait d'aller dans le sens d'une conversion des IPDs en ILDs dans les os du crâne, par des effets d'interférences. Mais ces résultats iraient à l'encontre des études posant des capacités spécifiques développées par les plongeurs.

4 Projets complémentaires

4.1 Apprentissage

4.1.1 Hypothèses

Un autre facteur important pouvant influencer la fréquence limite de détection de l'IPD concerne l'apprentissage. Les hypothèses que l'on défend visent à remettre en cause les limites que l'on pourrait formuler vis-à-vis du système nerveux (contre l'hypothèse d'une période critique, §II.2.3). Si les limites de perception de l'IPD sont effectivement acoustiques, on peut postuler la possibilité d'un apprentissage de la détection d'IPDs à des fréquences plus élevées. Il en résulterait une adaptation du système auditif au niveau neuronal. Malgré les résultats obtenus avec les plongeurs, il y a de bonnes raisons de penser que le système auditif est suffisamment plastique pour se réadapter à une nouvelle configuration des indices de localisation auditive. On peut citer notamment l'étude [TS11], qui montre qu'en ajoutant un délai entre les deux oreilles par l'intermédiaire de bouchons électroniques, les sujets réapprennent à localiser au bout de 50 heures de port. Il s'agirait pour nous à l'inverse de diminuer le temps d'arrivée des sons entre les deux oreilles, afin de forcer la décharge des neurones à des fréquences plus élevées.

4.1.2 Dispositif expérimental proposé

Pour cette expérimentation, le sujet porte un casque audio sur lequel sont fixés deux microphones omnidirectionnels, placés suivant la direction de l'axe interaural au-dessus de la tête du sujet, et espacés d'une distance de 8 cm, générant une différence de temps de 234 μs pour une source sonore placée dans l'axe des microphones. La fréquence correspondant à la longueur d'onde pour cette espacement est de 4275 Hz. Au-delà de cette fréquence, il existe un biais quant au nombre de périodes comprises dans une telle distance. En-deça de cette fréquence, il n'existe pas de biais, aussi, on postule que le système auditif doit être capable de s'adapter à ces nouvelles conditions d'écoute. Le casque audio et les microphones sont branchés sur un enregistreur numérique portable Zoom H2, de telle manière que le casque diffuse le son capté par les microphones à travers le circuit de monitoring. Les gains des microphones sont réglés pour avoir un rapport signal sur bruit correct, en fonction de l'ambiance générale. Le niveau d'écoute est réglé à un niveau confortable, correspondant à l'écoute sans le port du casque. Le circuit de monitoring est converti analogique-digital/digital-analogique et le signal numérique a une fréquence d'échantillonnage de 96 kHz.

4.2 Vision

4.2.1 Hypothèses

L'influence de la vision sur la localisation auditive n'est pas une problématique nouvelle. L'étude [HH92b], en particulier, montrait déjà une corrélation entre les capacités de localisation sonore et la largeur du champ visuel, déduite à partir de la taille de la fovéa, pour différentes espèces animales. On se demande quel rôle peut avoir la vision vis-à-vis de la localisation auditive. D'une certaine manière, on peut comprendre le traitement de nos perceptions par la modalité de l'instinct de causalité. On postule instinctivement qu'un effet doit avoir une cause, à tel point qu'on pourrait se laisser leurrer pour peu qu'on nous fournisse une cause plausible, comme on peut le constater dans la plupart des illusions qui tendraient à décorréler une même source entre ses aspects visuels et auditifs (par ex. le cas du ventriloque). Par ailleurs, on s'attend à retrouver le même effet pour une même cause. Ainsi, lorsqu'on perçoit une source sonore, on doit pouvoir la retrouver visuellement, ou par les autres sens, car sans cette cause unitaire il n'y aurait pas l'effet observé. Par suite, si ce qui correspond à la cause, c'est-à-dire la source sonore, n'apparaît pas également de manière visuelle, ou du moins pas au bon endroit, alors on doit réajuster nos mécanismes cognitifs en comptant sur la plasticité du cerveau. Celui-ci est en perpétuelle reconstruction du monde grâce aux sensations et à la proprioception. Pour s'en convaincre, il suffit de constater que l'image faite du monde sur la rétine est à l'envers, et que le cerveau traite les informations visuelles pour les replacer en corrélation avec le reste des autres sens. D'une manière générale, ce ne sont pas directement les décharges neuronales qui peuvent produire notre perception, sans une interprétation intermédiaire. À nouveau (cf. II.1), on part d'un résultat inter-espèces à partir duquel on voudrait mettre à jour un résultat inter-humains. Il s'agit pour cela de mettre en avant une variabilité inter-humains, ici concernant la vision, pour montrer une corrélation avec la localisation auditive. Després [Des04] donne un certain nombre de résultats significatifs entre vision et audition, comme le fait que les myopes localiseraient mieux les sources sonores que les sujets avant une vision normale. L'explication qu'il suggère est que leur vision périphérique est détériorée (port de lunettes), réduisant ainsi la largeur du champ visuel, et que l'audition, calibrée en tournant la tête, vient compenser la vision.

4.2.2 Résultats

Résultats du §II.2.2. On reprend ici les résultats exposés au §II.2.2, en distinguant cette fois-ci le groupe vision normale/myopie (cf. Fig. 11).



FIGURE 11 – Fréquence limite de détection de l'IPD (en Hz) pour des sons purs, en fonction de la fréquence limite théorique calculée à partir de la mesure de la distance interaurale d = 2a, d'après l'Equation 2, pour $c = 342 \text{ m.s}^{-1}$ et $\theta = \pi/2$. Les barres d'erreur indiquent les écart-types. En vert : sujets à la vision normale ; en rouge : sujets myopes.

Saccades occulaires. Une expérimentation complémentaire a été effectuée sur 2 sujets comprenant l'expérimentateur. Il s'agissait de faire le test du §II.2, mais en tournant les yeux autant que possible lors de l'écoute des stimuli. On affiche les résultats sur la Figure 12.



FIGURE 12 – Fréquence limite de détection de l'IPD (en Hz) pour des sons purs, pour les sujets VI (l'expérimentateur) et PY. Les barres d'erreur indiquent les écart-types. En vert : sans saccades occulaires; en rouge : avec saccades occulaires.

4.2.3 Discussion

L'impact de la vision ne semble pas avoir d'effet significatif sur la détection des IPDs. Cependant, en ce qui concerne les saccades occulaires, les résultats sont significatifs. Il ne faut pas non plus en tirer de conclusion trop hâtive, puisque l'expérimentation devrait être réalisée sur un panel plus important de sujets. De plus, il y a une indétermination sur l'influence de la vision, car il est certain que les saccades occulaires augmentent la concentration en forçant le sujet à rester actif à la tâche. Il serait donc judicieux de se reporter sur des expérimentations liées à la concentration et de tester d'autres situations pour juger de l'influence de la vision dans notre expérimentation.

5 Conclusion du projet n°1

Nous avons mis en place un protocole expérimental pour déterminer la fréquence limite de détection de l'IPD. L'objectif était de mettre en évidence une corrélation entre cette fréquence limite de détection de l'IPD et la taille de la tête, sur une base d'hypothèses consistantes. Toutefois elle s'est révélée infructueuse, aussi bien pour des sujets quelconques que pour des sujets plongeurs. On peut donc pas conclure positivement quant à une corrélation entre la fréquence limite de détection de l'IPD et la taille de la tête. Les raisons qui peuvent expliquer nos résultats malgré les hypothèses émises à l'origine sont principalement le défaut d'attention et/ou l'existence d'une période critique dans la formation du système auditif. L'expérimentation sur les

sujets plongeurs nous a également amené à une impasse quant à l'explication de la localisation auditive humaine sous l'eau. Pour conclure sur ce projet, il s'agirait d'étudier plus précisément la conduction osseuse au niveau du crâne. D'une manière générale, il n'est pas possible de conclure à ce stade sur les mécanismes de détection de l'IPD, au cœur du système nerveux, si ce n'est en se rattachant aux explications fournies dans la littérature, puisque nos résultats correspondent à la valeur limite généralement admise de 1.5 kHz. Nous pouvons, de plus, faire le constat des différentes échelles de développement du système nerveux à travers lesquelles se répartissent nos différentes problématiques : l'échelle de l'évolution inter-espèces, l'échelle humaine (période critique), l'échelle restreinte (apprentissage), l'échelle immédiate (l'attention). Ces échelles permettent de replacer les hypothèses dans leur contexte pour des expérimentations ultérieures.

Troisième partie

Projet n°2 : Modèle unitaire de la localisation auditive : poids des ITDs à travers les fréquences

1 Modèle bayésien

1.1 Pondération du poids des indices de localisation

Il semble difficile a priori d'expliquer l'unité des indices de localisation pour la seule information de localisation, au sein du système auditif, d'après les modèles dérivés de la théorie duplex. Pour constituer un modèle unitaire de la perception de la localisation, il s'agit de poser une hypothèse bayésienne d'informativité des indices : en basses fréquences (f < 1.5 kHz), les ILDs étant plus faibles que les ITDs, on postule que les premiers sont moins informatifs que les seconds, et inversement en hautes fréquences (f > 1.5 kHz). En effet, la valeur des indices ITDs et ILDs varie en fonction de la fréquence, et donc ne fournissent pas la même informativité, un indice pouvant prédominer sur l'autre en fonction de la fréquence. Ainsi, l'ILD est perceptible dès 200 Hz mais est négligeable jusqu'à environ 2 kHz face à l'ITD (cf. [Mar96]). Cela est dû à la configuration acoustique de la tête qui induit un filtrage des ondes sonores faible en basses fréquences, tandis qu'en hautes fréquences les différences d'intensité peuvent atteindre 20 dB (cf. [Gel98]). Il s'agit ensuite de vérifier que les sujets agissent comme des observateurs bayésiens. Pour cela, on souhaiterait déterminer le poids de chaque indice en fonction de la fréquence afin de prédire la localisation d'une source latéralisée. Pour un observateur bayésien, il s'agit de se fier à la probabilité de rencontrer une source à un endroit donné, à partir d'indices de localisation. Dans l'Equation 4, A représenterait l'angle réel d'azimuth de la source, et B les indices de localisation :

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)} \tag{4}$$

1.2 ITDs à travers les fréquences

Analyse de bases de données HRTFs. En commençant en basses fréquences par les ITDs, Benichoux *et al.* [BRB12] ont mis en évidence, à partir d'une base de données de HRTFs (Head-Related Transfer Functions), que les ITDs sont 50 % plus importants en basses fréquences (BF) jusqu'à 600 Hz, que pour des fréquences plus élevées, jusqu'à 1,5 kHz, que nous appelerons dans la suite de notre étude hautes fréquences (HF), relativement à l'indice ITDs (cf. Fig. 13).



FIGURE 13 – Moyennes (lignes bleues) et écart-types (zones bleues) des ITDs normalisés (cf. Eq. 5) moyennés sur 130 sujets des bases de données HRTFs de l'IRCAM, du CIPIC, et de l'ARI, pour le plan horizontal, d'après [RBB13]. Le symbole \bigstar donne une approximation de l'ITD normalisé en BF, le symbole \blacksquare de la fréquence de transition normalisée (cf. Eq. 8), et le symbole \bullet de l'ITD normalisé en HF.

Il doit donc exister un biais d'ITD lorsque les contenus fréquentiels varient pour localiser les sources au même endroit. Si l'on veut prédire le comportement, il s'agit donc de prédire ce biais qui existe entre les ITDs en basses fréquences et en hautes fréquences pour localiser les deux sources respectives au même endroit. La modélisation des ITDs à travers les fréquences est rendue complexe à cause de la diffraction du son par la tête, et du fait que les morphologies sont toutes différentes et induisent des ITDs qui varient en fonction des fréquences pour toutes les directions d'une source, d'une manière propre à chaque sujet. On ne réécrit pas ici les démonstrations mathématiques de cette dépendance en fréquence mais on pourra consulter par exemple la thèse de Daniel [Dan01] pour les explications acoustiques de l'évolution des ITDs en basses fréquences (retard de phase) et en hautes fréquences (retard de groupe). On peut simplement retenir ici que la diffraction des ondes sonores sur la tête accroît les ITDs en basses fréquences. Or, cette différence d'ITDs doit être perceptible.

Maximisation du biais. L'analyse de la base de données HRTFs nous permet de déterminer la différence d'ITDs entre les zones BF (f < 600 Hz) et HF (600 < f < 1,5 kHz) pour lesquels le biais est maximisé, en fonction de l'azimut de la source et pour une élévation nulle. Les deux zones fréquentielles BF et HF correspondent aux paliers d'ITDs qu'on peut observer sur la Figure 13, pour une même position de source sonore. On observe également une bande de transition autour de 600 Hz. En prévision de notre expérimentation, on va donc chercher 2 ITDs, correspondants à 2 angles, qui vont maximiser la valeur absolue de la différence d'ITDs, correspondants à des angles compris entre -90 et +90°, entre les zones BF et HF. L'analyse de la base de données HRTFs nous permet de choisir les angles 60° et 50°.

2 Expérimentation

2.1 Prédictions

Le but de cette expérimentation est de rendre compte du biais d'ITD qui doit exister dans la localisation d'une source BF et d'une source HF perçues au même endroit. Ce biais doit être prédictible de manières qualitative et quantitative.

- Qualitatif :
 - Côtés de la tête : le biais des côtés gauche et droite de la tête, pour la perception de sources BF et HF au même endroit, doivent être les mêmes (au signe près), en supposant que la tête est symétrique.
 - Azimuts nul et non nul : si une source est frontale (ITD nul), le biais doit être nul; si une source est latéralisée (ITD non nul), le biais doit être non nul (cf. §III.1.2 Maximisation du biais).
 - Spectre des sons : si le spectre est identique entre les 2 sons, on prédit un biais nul, si le spectre est différent, on prédit un biais non nul.
- Quantitatif : d'une manière générale, les biais sont prédictibles de manière quantitative si l'on dispose des HRTFs du sujet. Sinon, on peut se reporter sur les moyennes effectuées sur les bases de données HRTFs. Les HRTFs nous fournissent directement les ITDs propres à chaque sujet en fonction de la fréquence. Chercher le biais d'ITD entre des sources aux contenus fréquentiels différents revient à calculer la différence qui existe entre les ITDs à ces fréquences.

2.2 Protocole

2.2.1 Expérimentation pilote : matching

Une première expérimentation pilote avait été codée par Benichoux en Python et consistait en une tâche de matching entre 2 bruits de bande latéralisés avec des ITDs. Le sujet écoute des paires de bruits et ajuste l'ITD du second bruit pour le placer à la même position que celle du premier. Il peut rejouer les sons et ajuster l'ITD jusqu'à être satisfait du matching, avant de passer à la paire suivante. Le problème de cette méthode est l'influence des sujets sur leur propre jugement du matching. Néanmoins, les résultats sont significatifs (cf. Fig. 14) et sont concordants avec les prédictions.



FIGURE 14 – Résultats pilotes avec la méthode de matching, sur 9 sujets, d'après [Ben12].

2.2.2 Paradigme 2AFC et méthode adaptative

Just-noticeable difference d'ITD. On s'inspire ici du paradigme du Minimum Audible Angle (MAA), développé par Mills [Mil58]. Il permet de mesurer la précision de localisation de sources sonores, mais peut nous servir ici à mesurer le biais de latéralisation entre nos deux sons. On cherche plus précisément ici la différence juste audible (Just-Noticeable Difference ou JND) d'ITD entre les deux sons. De cette manière, en choisissant deux sons respectivement BF et HF, la mesure du JND nous fournit le biais de latéralisation entre des sons BF et HF en valeur d'ITD.

Protocole expérimental.

Jeux de conditions en fonction des stimuli. On présente successivement deux bruits de bande, ayant chacun une durée de 800 ms, comprenant une rampe de départ et une rampe de fin de 50 ms respectivement, et séparés par un invervalle de silence de 500 ms. Leur contenu fréquentiel est soit basses fréquences (f < 600 Hz), soit hautes fréquences (f > 600 Hz), pour une largeur de bande de 3 ERBs. Leur fréquence centrale est déterminée en fonction de la distance interaurale du sujet d (en m), mesurée à l'aide d'un compas anthropométrique, suivant le système d'Equations 10 énoncé plus bas. Les stimuli sont latéralisés dans l'espace interaural avec un certain ITD. Sur les 2 stimuli qui composent la paire, l'un des deux est la référence et l'autre la cible. La référence est fixe et la cible se déplace entre chaque essai suivant les réponses du sujet, la tâche du sujet étant d'indiquer si le 2e stimuli était à gauche ou à droite du 1er en cliquant sur l'un des deux boutons correspondants. 24 jeux de conditions sont employés en parallèle, par méthodes adaptatives, et dans un ordre aléatoire afin de déterminer le plus petit ITD discriminable par le sujet, suivant 5 critères :

 2 positions correspondant aux 2 angles de 60° et 50° choisis pour maximiser le biais attendu.
 Les ITDs correspondants sont calculés en fonction de la taille de la tête du sujet, suivant la formule suivante :

$$ITD^{BF/HF} = \left| \frac{d \cdot ITD_{norm}^{BF/HF}}{2 \cdot c} \right| \tag{5}$$

où $ITD^{BF/HF}$ est exprimé en μs , c est la vitesse du son (342 m.s⁻¹), d la distance interaurale du sujet (en m), et ITD_{norm} les ITDs normalisés d'après une moyenne de mesures HRTFs pour chaque angle. Pour chacun des 2 bruits, on obtient donc 2 ITDs pour chaque angle. Pour avoir un ordre de grandeur, pour une distance interaurale de 13.5 cm, on obtient des ITDs de 361 et 302 μs en HF, et de 463 et 403 μs en BF, pour les angles respectivement de 60° et 50° à partir desquels sont calculés les ITDs normalisés.

- Les 2 côtés gauche et droit du sujet, en prenant les ITDs opposés de ceux obtenus à partir des angles 60° et 50° pour le côté opposé.
- Le spectre est identique (par ex. BF/BF) ou non (par ex. BF/HF).
- La référence est BF ou HF.
- La référence est jouée en 1er ou en 2nd (dans les cas BF/BF et HF/HF, ce critère n'est pas retenu, d'où les 24 jeux de conditions pour 5 critères au lieu de 32).

Méthode adaptative. Lors du déroulement de la procédure, la cible a un ITD ajouté en positif ou en négatif à la référence, dont l'ITD est fixe. L'ITD ajouté est choisi au départ aléatoirement dans $[-200; -150] \cup [150; 200]$ (en μs). Lorsque le sujet répond "à gauche" ou "à droite", l'ITD de la cible est modifié dans le sens qu'indique le sujet en ajoutant ou en soustrayant (suivant le côté du sujet) un pas d'ITD pour générer un nouvel ITD de la cible, utilisé lors de la prochaine écoute de la paire de stimuli (cf. Fig. 15).



FIGURE 15 – Allure de l'évolution des différences d'ITD entre les stimuli de différentes paires, pour des stimuli latéralisé d'un côté de la tête, au contenu fréquentiel identique (HF/HF et BF/BF) et différent (BF/HF et HF/BF). Les lignes en pointillés indiquent la valeur finale de la plus petite différence d'ITD discriminable entre les deux stimuli.

Les méthodes adaptatives concernent le pas d'ITD pour déplacer la cible, et le nombre d'inversions fixé pour arrêter la procédure. Le pas d'ITD varie à chaque inversion de la courbe de réponse du sujet (par exemple le sujet répond "à gauche" puis "à droite" ou inversement) suivant une évolution hyperbolique. Il est fixé au départ à 110us et atteint 50us, l'équivalent d'un JND (cf. [LBD13]), à la fin de la procédure pour cette paire de stimuli, ce qui donne en fonction des essais :

$$Pas(n_{inversions}) = 110 + \frac{n_{max}}{n_{max} - 1} \cdot \frac{60}{n_{inversions} + 1} \tag{6}$$

où $Pas(n_{inversions})$ s'exprime en μs , avec $n_{max} = 6$ le nombre d'inversions fixés pour arrêter la procédure. La variation du pas en fonction des inversions est donc déterminée par :

$$Pas(n_{inversions}) = 110 + \frac{360}{5 \cdot (n_{inversions} + 1)}$$
(7)

De plus, une sécurité bloque la différence d'ITDs entre la référence et la cible à 400 μs en valeur absolue, afin que l'ITD de la cible ne dépasse pas le seuil de latéralisation par ITD, auquel cas on entendrait un écho plutôt qu'un unique stimulus latéralisé. Avant d'effectuer le test formel, le sujet effectue un pré-test de 10 essais, afin de comprendre la tâche, avec des stimuli où le choix dans les 5 critères est aléatoire. Le plus petit ITD discriminable entre la référence et la cible est calculé en faisant la moyenne sur les deux dernières inversions.

Stimuli. Les stimuli sont des bruits de bande dont la fréquence centrale dépend de la distance interaurale du sujet. Les fréquences normalisées, déterminées pour chaque bruit, nous permettent de calculer les fréquences centrales des bruits de manière inversement proportionnelle à la distance interaurale. On calcule les fréquences normalisées grâce à la formule suivante (d'après [RBB13]) :

$$f_{norm} = \frac{f}{\hat{f}} \tag{8}$$

où $\hat{f} = \frac{c}{2\pi a}$, avec $c = 342 \text{ m.s}^{-1}$ la vitesse du son, a le rayon de la tête (en m). On a cherché à optimiser les résultats pour d = 2a = 0.135 (en m), en choisissant un bruit BF de fréquence centrale 300 Hz et un bruit HF de fréquence centrale 1300 Hz. On obtient les fréquences normalisées suivantes :

$$\begin{cases} f_{c_{norm}}^{BF} &= 0.3720 \\ f_{c_{norm}}^{HF} &= 1.6121 \end{cases}$$
(9)

On peut en déduire les fréquences centrales (en Hz) des bruits pour d'autres tailles de têtes :

$$\begin{cases} f_c^{BF} &= \frac{0.3720 \cdot c}{\pi \cdot d} \\ f_c^{HF} &= \frac{1.6121 \cdot c}{\pi \cdot d} \end{cases}$$
(10)

où $c = 342 \text{ m.s}^{-1}$ est la vitesse du son, et *d* la distance interaurale (en m). Les largeurs de bruit ont été fixées à 3 ERBs. Pour avoir un ordre de grandeur, pour une distance interaurale de 13.5 cm, cela donne des bandes passantes de 171.3 Hz en BF et 495.3 Hz en HF. Les ITDs utilisées pour latéraliser les bruits sont calculées d'après l'Equation 5.

Dispositif expérimental. Les stimuli sont générés sous Matlab, à la fréquence d'échantillonnage de 192 kHz (la durée d'un échantillon vaut donc 5.2 μs), et envoyés dans un casque Sennheiser HD600 (circumaural, ouvert) à travers une carte son RME Fireface 400. On étalonne les gains pour chaque écouteur et pour chaque bruit BF et HF, afin de les diffuser en moyenne à 65 dBA. Comme ils varient en fonction de la taille de la tête, on calcule le gain moyen pour chaque écouteur et pour chaque bruit, entre les valeurs extrêmes correspondant aux distances interaurales de 12 cm ($f_c^{BF} = 338$ Hz et $f_c^{HF} = 1463$ Hz) à 14 cm ($f_c^{BF} = 289$ Hz et $f_c^{HF} = 1254$ Hz), par pas de 0.5 cm. Le sujet est placé dans une cabine d'écoute à acoustique atténuée pour effectuer le test.

2.3 Résultats

La Figure 15 présentait l'allure de l'évolution des différences d'ITD pour des stimuli latéralisés en fonction des réponses du sujet. On présente ici nos résultats obtenus pour 3 sessions du test, respectivement pour une tâche de contrôle où les ITDs des références ont été fixés à 0 μs (cf. Fig. 16), et pour la tâche proprement dite, où les ITDs des références correspondent à des angles de 60° et 50° (cf. Fig. 17). Les moyennes et écart-types sont calculés sur les résultats des 3 sessions et affichés pour les 24 jeux de conditions, ainsi que les valeurs de prédiction, calculées d'après les données HRTFs. Les valeurs de prédiction sont opposées entre les 2 côtés de la tête, et suivant que la référence est BF ou HF. Elles sont égales, par hasard, pour les ITDs correspondants aux angles de 60° et 50°. Elles sont nulles pour les ITDs correspondants à l'angle de 0°.



FIGURE 16 – Tâche de contrôle : ITDs des références fixées à 0 μs . Moyennes sur 3 tests des différences d'ITDs entre la référence et la cible pour les 24 combinaisons des 5 critères, pour le sujet-expérimentateur.



FIGURE 17 – Maximisation des biais. Moyennes sur 3 tests des différences d'ITDs entre la référence et la cible pour les 24 combinaisons des 5 critères, pour le sujet-expérimentateur.

2.4 Discussion

Bruits de même contenu fréquentiel. Conformément à ce qui était attendu, les résultats pour les paires BF/BF et HF/HF ne montrent pas de différences significatives, avec une valeur p supérieure à 0.5 pour les trois conditions restantes : position 60° ou 50°, côtés de la tête, référence BF ou HF.

Bruits au contenu fréquentiel différent. Pour les paires de bruits au contenu fréquentiel différent, respectivement avec une référence BF et HF, les résultats sont significatifs pour la condition correspondant au côté de la tête (valeur p < 0.01), ce qui était un résultat attendu, avec des valeurs attendues opposées. Les résultats ne sont pas significatifs pour les autres conditions : position 60° ou 50°, la référence est présentée en 1ère ou en 2e, ce qui était aussi des résultats attendus.

3 Conclusion du projet n°2

Ce projet sur la localisation de sources sonores latéralisées en ITD en fonction du contenu fréquentiel a été mené à partir d'un constat acoustique sur une base de données HRTFs : les ITDs varient en fonction des fréquences et sont 50% plus importantes en basses fréquences qu'en hautes fréquences (jusqu'à 1.5 kHz). Les implications sur la perception ont été vérifiées pour 1 sujet et elles sont significatives dans la mesure où elles sont importantes pour la perception de sources aux contenus fréquentiels différents. Tandis que pour des sources de même contenu fréquentiel, leur perception à une position donnée n'implique pas d'ITDs significativement différents. L'expérimentation pour déterminer le JND d'ITD en fonction du contenu fréquentiel de la source sonore pourra être effectuée sur un panel plus important de sujets. Les résultats de ce projet pourront être intégrés, par la suite, à un modèle général de la localisation auditive pour un observateur bayésien, en y joignant les ILDs.

Conclusion

Ces deux projets nous ont permis de discerner des paramètres fondamentaux de la localisation auditive, de concevoir des protocoles expérimentaux adaptés à chaque situation, de les optimiser en fonction des résultats attendus, pour enfin mettre en œuvre les expérimentations et confronter les résultats aux hypothèses. Dans un premier temps, le projet sur la fréquence limite de détection des IPDs a montré l'importance de la variabilité des données morphologiques et physiologiques. Bien que n'ayant pas obtenu les résultats espérés, ceux-ci nous permettent de redistribuer la problématique de départ vers les problématiques suivantes : formation du système auditif et période critique pour expliquer la variabilité de la perception de l'IPD chez des sujets quelconques, autour de 1.4 kHz, mécanismes d'interférences en conduction osseuse pour expliquer la perception de la localisation auditive sous l'eau chez des sujets plongeurs, mécanismes d'apprentissage au niveau du système auditif, rôle de l'attention en psychométrie. Dans un second temps, le projet sur la perception des ITDs à travers les fréquences a permis de faire un lien direct entre les données acoustiques des HRTFs et la perception des ITDs. La perception de la localisation à partir d'indices temporels dépend non seulement du placement de la source dans l'espace, mais aussi de la diffraction du champ acoustique sur la tête. Le projet sur la perception des ITDs nous a permis de montrer comment la perception de la localisation évolue en fréquences. On peut retenir notamment que les ITDs en basses fréquences peuvent atteindre 1.5 fois les ITDs en hautes fréquences (jusqu'à 1.5 kHz) pour des sources latéralisées, et que ces différences sont perceptibles. Le traitement de ces informations fréquentielles par le système nerveux serait à approfondir sur la base des pistes proposées en conclusion du projet sur la fréquence limite de détection de l'IPD.

Bibliographie

[APS71] Arthur R. M., Pfeiffer R. R., Suga N., Properties of "two tone inhibition" in primary auditory neurons, Journal of Physiology (London), 1971, 212, 593-609.

[Ben12] Benichoux V., Frequency dependence of ITDs, Binaural Bash, 2012.

[BRB12] Benichoux V., Rébillat M., Brette R., On the variations of inter-aural time differences (ITDs) with frequency, Acoustics 2012 Nantes, 2012.

[Ber03] Bernaschina F., Localisation spatiale acoustique en milieu subaquatique, Thèse de doctorat, Genève, 2003.

[Dan01] Daniel J., Représentation de champs acoustiques, application à la transmission et à la reproduction de scènes sonores complexes dans un contexte multimédia, Thèse de doctorat, Paris 6, 2000.

[**Des04**] Després O., Mécanismes de localisation spatiale chez l'homme : interaction entre le système visuel et le système auditif, Thèse de doctorat, Strasbourg I, 2004.

[Gel98] Gelfand S. A., *Hearing : An introduction to psychological and physiological acoustics*, New York : Marcel Dekker, 1998.

[Gou98] Goussard J. P., Croissance et maturation de l'enfant à l'adulte, Développement de l'enfant, 1998.

[GM10] Grose J. H., Mamo S. K., Processing of temporal fine structure as a function of age, Ear Hear, 2010 December; 31(6): 755-760.

[Har60] Harris G. G., *Binaural interactions of impulsive stimuli and pure tones*, J. Acoust. Soc. Am., vol. 32, 1960.

[HH92a] Heffner R. S., Heffner H. E., *Evolution of sound localization in mammals*, The Evolutionary Biology of Hearing, pp.691-715, NY : Springer-Verlag, 1992.

[HH92b] Heffner R. S., Heffner H. E., Visual Factors in Sound Localization in Mammals, J. of Comparative Neurology, 317 : 219-232, 1992.

[Kuh77] Kuhn G. F., Model for the interaural time differences in the azimuthal plane, J. Acoust. Soc. Am., 62(1): 157-167, 1977.

[Mil58] Mills A. W., On the minimum audible angle, J. Acoust. Soc. Am., 30 : 237-246, 1958.

[LBD13] Le Goff N., Buchholz J. M., Dau T., Spectral integration of interaural time differences in auditory localization, ICA, Montréal, 2013.

[LR05] Levitin D. J., Rogers S. E., Absolute pitch : perception, coding, and controversies, Trends Cogn. Sci., 9(1), 26-33, 2005.

[Mar96] Martin M., Etude de la localisation en restitution du son pour la téléconférence de haute qualité, Thèse de doctorat, Université du Maine, 1996.

[Mol99] Moller A. R., *Review of the roles of temporal and place coding of frequency in speeck discrimination*, Acta Otolaryngol 119, 424-30, 1999.

[MG83] Moore B. C. J., Glasberg B. R., Suggested formulae for calculating auditory-filter bandwidths and excitation pattern, JASA, vol. 74, Sept. 1983.

[Ray07] Rayleigh, On our perception of sound direction, Philos. Mag. 13, 1907.

[**RBB13**] Rébillat M., Benichoux V., Brette R., On the coupled frequency and direction dependence of Inter-aural Time Differences, 2013.

[Ric10] Richard C., Étude de l'encodage des sons de parole par le tronc cérébral dans le bruit, Thèse de doctorat, Lyon 2, 2010.

[RG08] Rowan D., Gray M., Lateralization of high-frequency pure tones with interaural phase difference and bone conduction, International Journal of Audiology, 47:7, 404-411, 2008.

[Sav01] Savel S., Localisation auditive subaquatique : de l'identification des mécanismes impliqués à la démonstration de la plasticité du système auditif et à l'application sur les plongeurs, Thèse de doctorat, Paris V, 2001. [SG05] Seidl A. H., Grothe B., Development of sound localization mechanisms in the mongolian gerbil is shaped by early acoustic experience, J. Neurophysiol., 94 :1028-1036, 2005.

[TS11] Trapeau R., Schönwiesner M., *Relearning sound localization with digital earplugs*, Canadian Acoustics, Vol.39 No.3, 2011.

Résumé

Le traitement de l'information sonore par le système auditif pour la perception de la localisation de sources sonores est loin d'être élucidé. L'un des objectifs de la psychoacoustique est d'aider à cerner les problématiques majeures de ce traitement, comme dans cette étude en ce qui concerne la localisation auditive. Nous y présentons deux projets liés aux différences interaurales de temps. Un premier projet sur la fréquence limite de détection des différences interaurales de phase (Interaural Phase Differences ou IPDs), et un second projet sur les différences interaurales de temps (Interaural Time Differences ou ITDs) au travers des fréquences. Ces deux projets visent à réévaluer l'implication du paramètre fréquentiel dans la perception des différences interaurales temporelles. Ce paramètre est traité dans le premier projet au niveau du codage de l'information auditive par le système nerveux, et dans le deuxième projet, au niveau des diffractions acoustiques induites par la tête sur les ondes sonores. Ainsi, on montre que le paramètre fréquentiel a son rôle à jouer à tous les niveaux du traitement de l'information sonore avant d'atteindre le dernier stade de la perception de la localisation des sources sonores.