

CFA '18 LE HAVRE ■ 23-27 avril 2018
14^{ème} Congrès Français d'Acoustique



HRTF prêt-à-porter pour le son binaural dans les futurs contenus d'Orange

R. Nicol^a, M. Emerit^b et L. Gros^a

^aOrange Lab, 2 Avenue Pierre Marzin, 22307 Lannion, France

^bOrange Lab, 4 Rue du Clos Courtel, 35510 Cesson-Sévigné, France

rozenn.nicol@orange.com

La synthèse binaurale permet de créer une scène sonore virtuelle 3D. Les sources sont spatialisées en convoluant leur son par des fonctions de transfert, dites HRTF (Head Related Transfer Function), qui reproduisent l'interaction entre les ondes acoustiques et la morphologie de l'auditeur. Ces fonctions dépendent fortement de l'individu. Pour garantir la meilleure qualité d'expérience, il est nécessaire d'utiliser des HRTF individuelles. Cependant, pour des applications destinées au grand public, il est difficile de fournir des mesures individuelles à tous les auditeurs. On préfère opter pour des méthodes de personnalisation. Une première solution simple est de proposer une sélection de type « prêt-à-porter » de jeux de HRTF, dans laquelle chaque auditeur vient chercher le jeu qui lui donne le meilleur rendu. Le problème consiste à identifier une sélection appropriée de jeux de HRTF, proposant un nombre minimal de jeux afin de ne pas alourdir le processus de choix par l'utilisateur, tout en permettant de satisfaire le maximum d'auditeurs. L'étude décrit une expérience visant à tester une telle sélection. Une réflexion a été menée pour construire un protocole de choix offrant une évaluation des HRTF à la fois intuitive et rapide. Il se base sur un test de localisation implémenté sous la forme d'un jeu de réalité virtuelle dans lequel le participant est entouré d'une sphère de haut-parleurs. Sa tâche consiste à pointer un laser sur le haut-parleur le plus proche de la direction perçue du stimulus binaural. L'objectif est de classer les jeux sur la base des performances de localisation de l'ensemble d'un panel d'auditeurs experts et non experts. Les résultats de l'expérience sont présentés et discutés.

1 Introduction

La technologie binaurale permet d'immerger l'auditeur dans une scène sonore 3D [1]. Elle se base sur des fonctions de transfert, dites "Head Related Transfer Function" (HRTF), qui caractérisent l'ensemble des interactions entre l'onde acoustique et la morphologie de l'auditeur (diffraction par la tête et les pavillons d'oreille, réflexion sur le haut du torse, etc...). Ces fonctions de transfert sont implémentées sous forme d'une paire de filtres binauraux (un filtre par oreille) pour spatialiser des sons autour de l'auditeur. Un point critique réside dans le fait que les HRTF dépendent fortement de la morphologie individuelle de l'auditeur. Idéalement, pour garantir la meilleure qualité de spatialisation et d'immersion, il faudrait mesurer les HRTF individuelles de chaque auditeur, ce qui n'est pas envisageable dans le cadre de services ou de produits destinés au grand public. Par ailleurs, comme nous le verrons plus loin, une synthèse binaurale individuelle ne garantit pas forcément la qualité de spatialisation. Pour contourner cette difficulté, des méthodes de personnalisation ont été développées. Une des plus simples consiste à proposer à l'auditeur de choisir les filtres qui lui procurent la meilleure qualité de rendu binaural dans une sélection de plusieurs jeux de HRTF (correspondant à différentes morphologies), une sorte de sélection "prêt-à-porter" par opposition à la solution "sur-mesure" que constitue la mesure de HRTF individuelles. Il s'agit alors d'identifier la sélection comportant le nombre minimal de jeux et qui permettra de satisfaire le plus grand nombre d'auditeurs. C'est l'objectif de l'étude présentée ici. La première partie de l'article décrit l'élaboration de la sélection prêt-à-porter de jeux de HRTF. Le protocole expérimental du test de localisation pour l'évaluer est ensuite précisé. Cette seconde partie se termine par les résultats et ouvre sur une discussion avant de conclure.

2 Proposition d'une sélection représentative de HRTF

Identifier une sélection de jeux de HRTF "prêt-à-porter" pour offrir la meilleure qualité de reproduction binaurale au plus grand nombre d'auditeurs est une question délicate. On ne compte à ce jour que deux propositions. La première a été constituée à partir d'une expérience dans laquelle 46

sujets de la base Listen [2] ont été invités à donner un note (échelle à 3 niveaux : "Bad", "Fair" et "Good") à chacun des 46 jeux contenus dans la base, en évaluant leur qualité de reproduction de trajectoires spatiales horizontales et verticales [3]. A l'issue de cette étude, une sélection de sept jeux de HRTF correspondant aux jeux classés au moins "Good" par l'un des sujets a été proposée.

La seconde étude a porté sur la même base de HRTF [4], mais, au lieu d'un jugement de qualité, une mesure de dissimilarité entre les jeux de HRTF a été évaluée. Les matrices de dissimilarité obtenues pour chaque sujet ont été exploitées pour mener une Analyse Multi Dimensionnelle (*Multi Dimensional Analysis* ou MDS) permettant d'identifier le nombre de dimensions gouvernant les jugements de dissimilarité. Il en est ressorti quatre dimensions perceptives :

- Modification spectrale : perception des déséquilibres spectraux (graves, aigus),
- Profondeur du champ sonore : perception du relief de la scène sonore,
- Latéralisation : perception des sons suivant l'axe gauche droite,
- Distance : perception proche ou lointaine des sons.

La mesure des dissimilarités entre toutes les paires possibles de HRTF se traduit sous la forme d'un dendrogramme, dans lequel les HRTF sont regroupées en fonction de leur similarité. Si on définit un seuil de dissimilarité, on obtient un nombre donné de classes de HRTF. Pour chaque classe, on choisit un jeu de HRTF représentatif des caractéristiques perçues des membres de cette classe. L'ensemble des jeux représentatifs constitue une proposition de sélection "prêt-à-porter" visant à couvrir l'étendue de la variété des HRTF de la base.

La première étude propose une sélection de type consensuelle, c'est-à-dire convenant au plus grand nombre. Avec la seconde, on cherche plutôt à offrir la plus grande variété de jeux pour couvrir la diversité des perceptions. Nous avons commencé à travailler avec la première solution qui propose un nombre *a priori* raisonnable de jeux de HRTF. On peut en effet craindre qu'il soit illusoire de chercher à représenter une grande variété avec moins de sept jeux.

Dans l'objectif de proposer aux utilisateurs un parcours client permettant de choisir un jeu parmi plusieurs, différents protocoles ont été testés, en s'inspirant d'une première étude décrite dans [5] et déployée sur le site HyperRadio de Radio France [6]. Tous les différents processus que nous avons testés ont échoué pour plusieurs raisons. Les résultats étaient peu reproductibles. Les sujets trouvaient la tâche très difficile quel que soit le scénario de choix proposé. Les retours des sujets sur leurs difficultés portaient sur deux points : le nombre trop important de jeux de HRTF à évaluer, et la trop forte similarité entre certains jeux. Nous avons donc décidé de réduire ce nombre. En reprenant les deux études citées précédemment et en réalisant des écoutes expertes, une nouvelle sélection comportant cette fois quatre jeux de HRTF a été retenue. Les résultats de la première étude [3] concernant cette sélection indiquent que tous les sujets ont noté chaque jeu au moins à "Fair". D'autre part, nous avons pris la précaution de choisir des jeux qui selon la seconde étude [4] sont le plus éloignés possible les uns des autres suivant les dimensions 3 et 4. Dans la suite, les quatre jeux sélectionnés seront baptisés 51, 52, 53 et 54.

3 Évaluation de la pertinence de la sélection de HRTF

L'objectif de notre expérience est de tester la sélection de quatre jeux précédemment obtenue afin de déterminer si elle permet de satisfaire un panel d'auditeurs en termes de rendu binaural. Nous avons choisi d'évaluer la qualité d'un jeu de HRTF sur la base des performances de localisation que l'auditeur obtient sur des sons virtuels synthétisés avec ce jeu. Nous faisons l'hypothèse que plus le taux d'erreur de localisation est faible, plus le jeu est adapté à l'auditeur. On se concentre donc sur la fidélité spatiale de la reproduction binaurale.

3.1 Paramètres expérimentaux

Chaque jeu de HRTF est évalué pour 10 directions. L'impact de l'individualisation des HRTF étant prépondérant sur la perception de l'élévation des sons, le choix des directions testées se focalise sur les élévations en sélectionnant 10 valeurs régulièrement distribuées entre -24 et 30 degrés. Pour définir une direction, un angle d'azimut tiré au hasard dans une sélection de 10 valeurs régulièrement réparties entre 0 et 324 degrés est associé à chaque angle d'élévation. Chaque sujet évalue avec ces 10 directions les quatre jeux de HRTF constituant la sélection "prêt-à-porter" proposée. De plus, pour tous les sujets, on dispose de leur jeu de HRTF individuelles qui a été mesuré préalablement [9]. Ce cinquième jeu est ajouté aux quatre précédents. Le jeu de HRTF individuelles constitue une sorte d'"ancrage" avec laquelle le sujet est censé présenter des performances optimales de localisation, car ce jeu contient l'encodage spatial que l'individu a appris dans son expérience d'écoute naturelle depuis sa naissance. Par ailleurs, les performances de localisation peuvent sensiblement varier d'un individu à l'autre. Dans la population, il existe des personnes présentant "naturellement" de piètres performances de localisation auditive. Ces personnes vont potentiellement obtenir des erreurs importantes de localisation, quel que soit le jeu

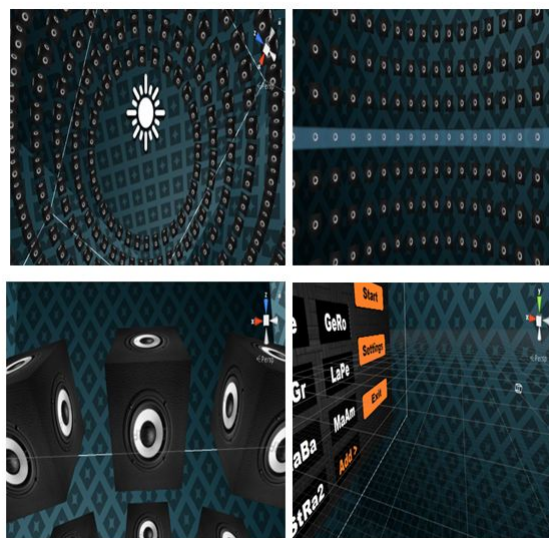


FIGURE 1 – Différentes vues intérieures (vue utilisateur) de la scène virtuelle 3D du test de localisation

d'HRTF, ce qui ne signifiera pas pour autant que ces jeux ne sont leur pas adaptés.

3.2 Protocole expérimental

Un test de localisation est réalisé au moyen d'une application de réalité virtuelle développée dans l'environnement Unity. Un problème remonté par certains sujets lors des tests décrits dans la section précédente (Section 2) et qui n'exploitaient que la modalité sonore, concernait le fait qu'il leur était difficile d'imaginer une source sonore externalisée en l'absence de support visuel. Nous avons donc décidé de développer une application associant une représentation visuelle des sources sonores aux stimuli sonores, ceci afin de faciliter au maximum la tâche du sujet. Le sujet est ainsi immergé, à l'aide du casque HTC-Vive associé à un casque d'écoute de type Beyerdynamic DT 990 Pro, dans un monde virtuel constitué d'une sphère de haut-parleurs placés tous les six degrés d'élévation entre -36 à 70 degrés et tous les six degrés en azimut (Fig. 1). Le rayon de la sphère est de 2 mètres. La tête de l'utilisateur est placée au centre de la sphère au lancement de l'application. Il peut ensuite tourner la tête et se déplacer librement dans une sphère de rayon 1.5 mètres (un cercle de couleur matérialise la limite à ne pas dépasser). Le stimulus sonore, constitué d'un train d'impulsions choisi pour sa facilité de localisation [7, 8], est émis en permanence. Il est apparu en effet lors de tests préalables qu'avec un stimulus court joué une seule fois, la tâche était jugée plus difficile par les sujets. De plus, une grande variabilité sur les réponses était observée.

Le sujet tient dans sa main une manette Vive qu'il voit dans le monde virtuel. Elle est prolongée par un faisceau laser avec lequel il lui est demandé de pointer le haut-parleur correspondant à la position perçue de la source sonore virtuelle. Un pointeur de la forme d'un petit point violet indique également au sujet la direction de son regard. Il valide son choix en pressant un bouton situé sous la manette. Il n'a pas de retour sur la justesse de sa réponse afin d'éviter un effet d'apprentissage. Dans les consignes lues en début d'expérience, le sujet est incité à bouger la tête, à se

déplacer, et à rapprocher de la position de la source perçue. Il est aussi demandé au sujet de fermer et d'ouvrir les yeux régulièrement, afin qu'il se construise une représentation mentale 3D de la scène. Si les HRTF utilisées pour la synthèse binaurale sont adaptées au sujet, la représentation sera correcte et l'erreur de localisation sera faible. Dans le cas contraire, le sujet aura des difficultés à se faire une représentation mentale de la scène, ce qui conduira à une augmentation des erreurs de localisation.

Le test comprend une première session d'apprentissage, dont le but est de familiariser le sujet avec la tâche. Pour cet apprentissage, trois directions sont tirées au sort entre les élévations -6 et 6 degrés. Le sujet localise 15 sources (3 directions x 5 jeux de HRTF). Cet apprentissage est suivi de la session principale qui comprend 50 conditions (10 directions x 5 jeux de HRTF). Les 50 conditions sont présentées dans un ordre aléatoire différent pour chaque sujet. La durée du test est comprise entre 20 et 30 minutes selon le sujet. Les résultats sont sauvegardés dans un fichier xml pour chaque sujet.

3.3 Sujets

Au total 24 sujets ont participé à l'expérience. Dix d'entre eux peuvent être considérés comme experts au sens où ils ont acquis une expérience d'écoute spatialisée (technologie binaurale et autres technologies de spatialisation sonore) de part leur activité (chercheurs ou ingénieurs du son) dans le domaine de l'audio 3D depuis de nombreuses années.

3.4 Implémentation de la synthèse binaurale dynamique

La synthèse binaurale dynamique temps réel est implémentée sous la forme d'un plugin Unity développé spécialement pour l'expérience. Ce plugin permet de lire les données HRTF au format SOFA. Les filtres binauraux sont calculés à chaque trame par interpolation en fonction de la position de la source par rapport à l'auditeur en prenant en compte les données du suivi de tête ("head-tracking"). La commutation des filtres entre deux positions se fait par la technique de "cross-fading" appliquée sur la sortie des unités de filtrage. La fréquence d'échantillonnage est fixée à 48 kHz. L'application Unity est réglée pour garantir une latence minimale. La taille des trames audio traitées par la fonction de "callback" est de 256 échantillons. Le retard induit par le système est quasiment imperceptible. Aucun sujet n'a reporté une quelconque gêne sur ce point. Aucune réverbération n'a été ajoutée et l'amplitude de la source sonore ne varie pas avec la distance. Le niveau sonore pour le stimuli joué en position frontale est réglé à 61 dBA pour chacun des écouteurs du casque. La différence de niveau entre les écouteurs est nulle.

Chaque jeu de HRTF est égalisé champs diffus. L'implémentation des fonctions de transfert utilise des filtres à phase minimale associés à un retard de groupe qui est estimé à partir des HRTF mesurées. Un lissage fréquentiel par douzième d'octave est appliqué. Afin d'éviter la perception de différences de latéralisation entre les jeux pour les directions simulées dans le plan médian, les fonctions de transfert sont symétrisées. La taille du filtre à phase minimale est de 128 échantillons. Le retard maximal estimé sur l'ensemble des HRTF de tous les jeux est de

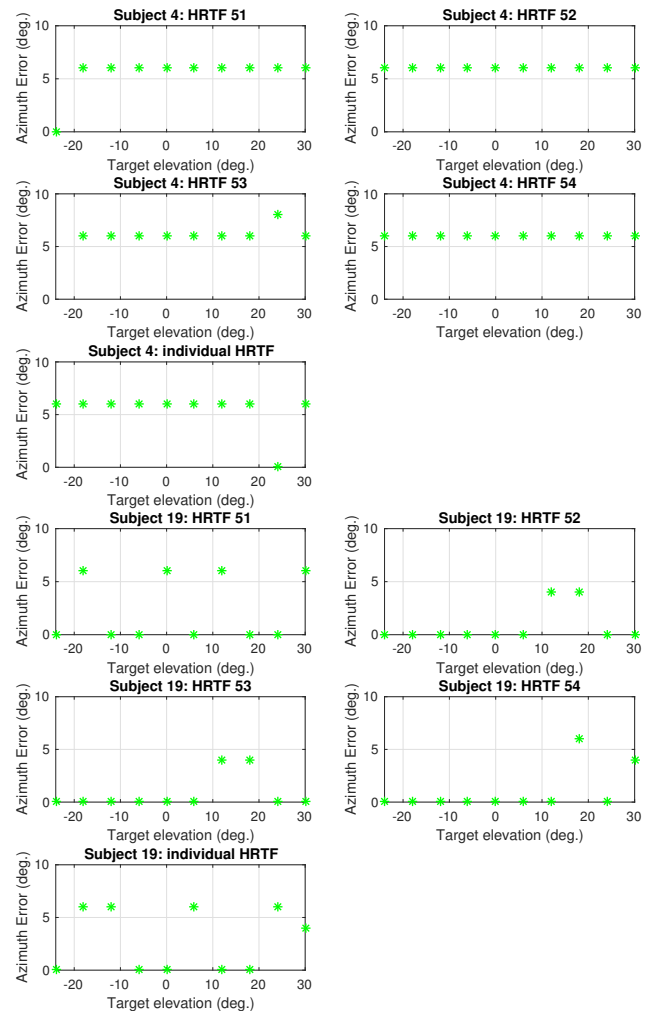


FIGURE 2 – Erreur de localisation en azimut en fonction de l'élévation cible : Sujets 4 et 19.

35 échantillons. La longueur totale des filtres combinant la composante à phase minimale et le retard varie donc entre 128 et 163 échantillons selon la direction.

3.5 Résultats

Pour l'ensemble des sujets, les performances de localisation en azimut sont excellentes. L'erreur est même nulle pour toutes les conditions pour cinq sujets. Pour six autres sujets, on ne relève pas plus de deux erreurs non nulles. Pour 21 sujets, l'erreur en valeur absolue reste inférieure ou égale à 6 degrés pour tous les jeux et toutes les directions. Aucune inversion avant/arrière n'est observée, ce qui n'est pas surprenant en raison du "head-tracking" et de la présence des indices visuels. Les résultats du sujet 19 (Fig. 2) donnent un exemple des performances les moins bonnes avec 30% d'erreurs non nulles toujours inférieures ou égales à 6 degrés. Trois sujets se démarquent sensiblement de ces observations. Pour deux sujets, on relève un biais systématique de 6 degrés, comme l'illustrent par exemple les réponses du sujet 4 (Fig. 2). L'un de ces sujets a mentionné qu'il avait eu une petite otite sur une oreille quelques jours avant l'expérience. Ce biais pourrait aussi venir d'un mauvais positionnement du casque, introduisant une fuite acoustique plus importante sur une des voies. Un troisième sujet présente une anomalie avec une erreur de 180 degrés

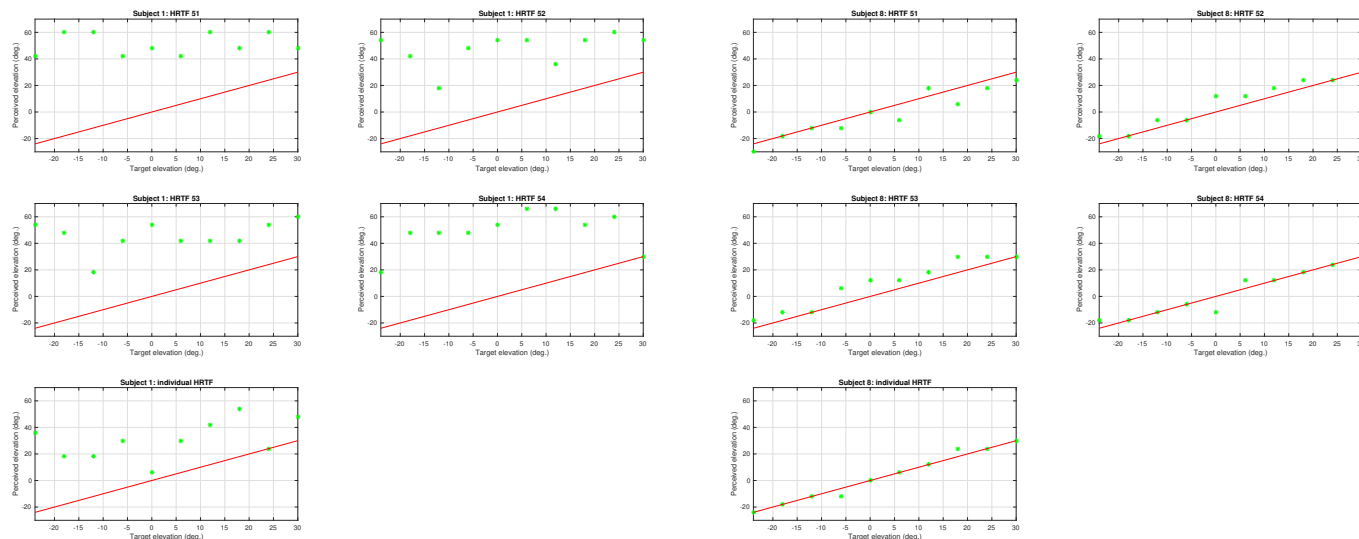


FIGURE 3 – Performances de localisation en élévation du sujet 1. L'angle perçu est représenté en fonction de la valeur cible (* en vert). La droite en rouge correspond aux réponses idéales sans erreur.

pour une réponse. Il ne s'agit pas d'une confusion avant/arrière, mais probablement d'une faute d'inattention.

On remarque sur la figure 2 que les performances de localisation en azimut ne varient pas significativement en fonction du jeu de HRTF. L'erreur de localisation en azimut n'apparaît pas discriminante pour identifier le jeu de HRTF adapté à un auditeur. L'identification du meilleur jeu va donc se concentrer sur les erreurs de localisation en élévation.

La figure 3 reproduit les élévations perçues par le sujet 1 pour l'ensemble des conditions. On constate que ce sujet présente de très fortes erreurs de localisation. Il localise l'ensemble des sons entre 40 et 60 degrés d'élévation quel que soit l'angle cible et quel que soit le jeu de HRTF. Ses performances en azimut ne sont pas très bonnes non plus. On peut se demander si ses performances de localisation n'ont pas été perturbées par l'environnement virtuel. Ce sujet a été exclu de la suite de l'analyse qui portera donc sur les 23 sujets restants.

La moyenne de l'erreur de localisation en élévation calculée sur les 10 directions évaluées pour chaque jeu est représentée sur la figure 5. Contrairement à l'erreur en azimut, on observe cette fois des différences prononcées entre les différents jeux pour la plupart des sujets. On note par exemple le cas du sujet 9 qui présente de fortes variations d'erreur selon le jeu. Pour le sujet 4, l'erreur obtenue avec le jeu 51 est nettement plus élevée (de l'ordre de 16 degrés) que pour les 3 autres (de l'ordre de 5 degrés). Les profils de réponse s'avèrent très différents d'un sujet à l'autre. Seulement sept sujets présentent l'erreur minimale pour leur jeu de HRTF individuelles. Les autres semblent lui préférer un jeu de la sélection proposée. Les performances de localisation en élévation paraissent donc être un critère pertinent pour discriminer les jeux de HRTF. La figure 4 reproduit les réponses de deux sujets pour chaque condition. Pour le sujet 8, on note des erreurs de localisation très faibles pour son jeu de HRTF individuelles. Il présente également une erreur faible pour le jeu 54 qui semble lui convenir. Le sujet 9, en revanche, obtient des erreurs importantes avec son jeu de HRTF individuelles. On observe un biais

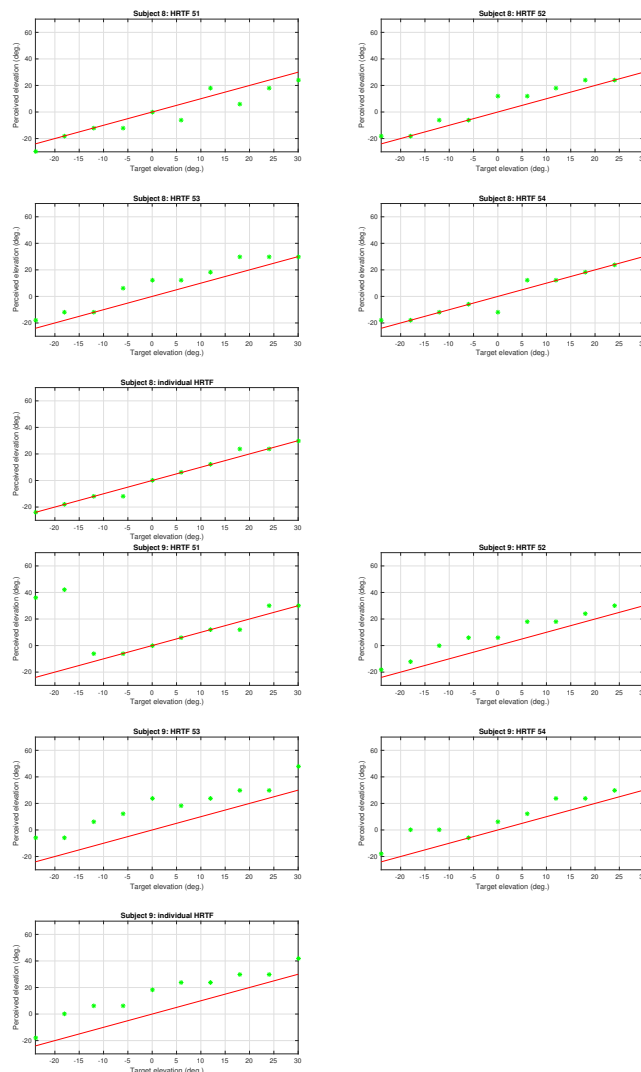


FIGURE 4 – Performances de localisation en élévation des sujets 8 et 9. L'angle perçu est représenté en fonction de la valeur cible (* en vert). La droite en rouge correspond aux réponses idéales sans erreur.

systématique allant jusqu'à 20 degrés. L'erreur n'est pas sensiblement plus faible avec aucun des jeux proposés dans la sélection.

Pour répondre à la question qui a motivé cette étude, à savoir s'il est possible d'identifier un jeu adapté à chaque auditeur dans la sélection, il faut se doter d'un critère permettant de valider un jeu au sens de la fidélité spatiale en termes d'élévation. On pourrait se baser sur l'erreur moyenne de localisation en élévation et fixer un seuil maximal (par exemple 6 degrés). Le(s) jeu(x) satisfaisant ce critère seraient alors identifiés comme convenant à l'auditeur. Cependant une analyse fine des performances de localisation (Fig. 4) suggère que l'erreur moyenne ne reflète pas les différences de profil observées (biais systématique, erreur importante isolée, ...).

4 Discussion

Afin de comparer les jeux, trois critères ont été définis :

- **MaxErrAll** : Le maximum de l'erreur en élévation pour toutes les directions évaluées est inférieur ou égal à 6

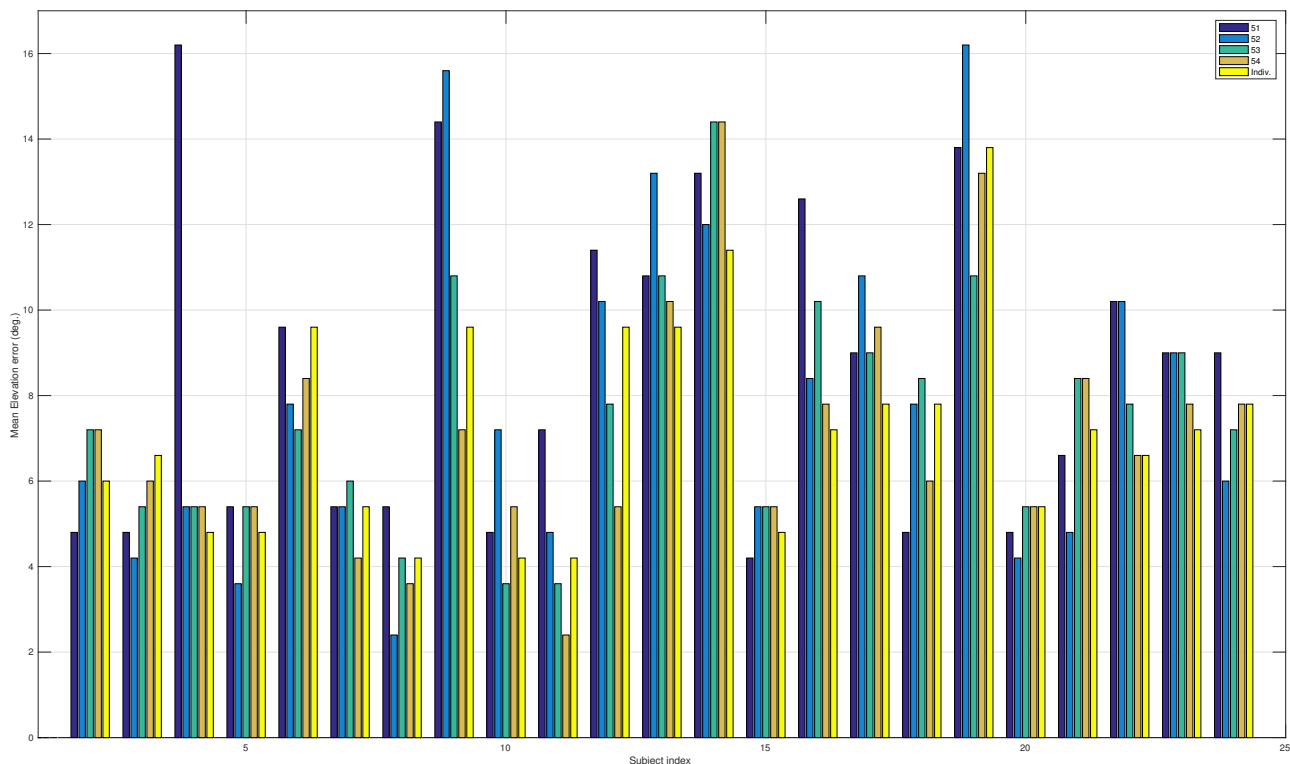
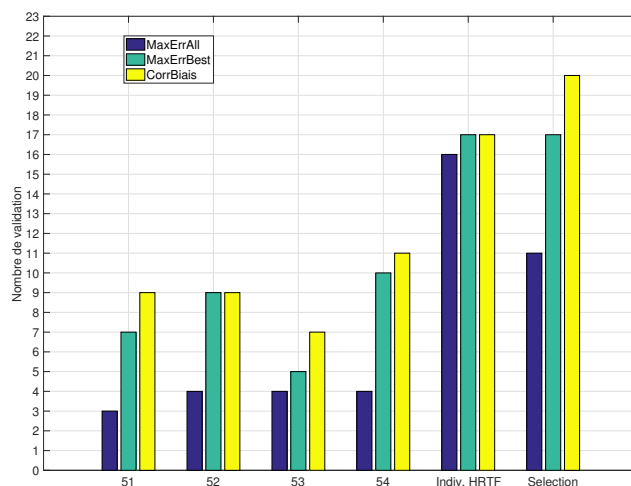


FIGURE 5 – Erreur moyenne de localisation en élévation en fonction du jeu de HRTF pour chaque sujet.

degrés.

- **MaxErrBest** : On reprend le critère précédent mais en tolérant une seule erreur de localisation supérieure à 6 degrés, ce qui permet de ne pas pénaliser un jeu pour une seule mauvaise réponse.
- **CorrBiais** : On mesure la corrélation entre les réponses du sujet et les valeurs cibles. Une corrélation supérieure à 0.9 indique que le sujet discrimine et ordonne bien les différentes élévations. Cependant une corrélation élevée peut être associée à un biais systématique et important. Une estimation de ce biais est donnée par l'erreur moyenne de localisation en élévation. On considère donc que le jeu est adapté si la corrélation est strictement supérieure à 0.9 et si l'erreur moyenne (biais) est strictement inférieure à 6 degrés.

Pour chacun des sujets, les résultats obtenus avec chaque jeu sont examinés à la lumière de ces 3 critères. On se rend compte que pour sept sujets (respectivement six) le jeu de HRTF individuelles n'est pas validé au sens du critère MaxErrAll (respectivement au sens de MaxErrBest et CorrBiais). En revanche, la majorité de ces sujets obtient des performances de localisation satisfaisantes avec un jeu de la sélection au sens d'au moins un critère. Au final, au moins un jeu de la sélection est validé pour onze (respectivement dix-sept et vingt) sujets au sens du critère MaxErrAll (respectivement au sens de MaxErrBest et CorrBiais). Seul un sujet ne valide aucun jeu (y compris son jeu de HRTF individuelles) au sens des trois critères. La figure 6 récapitule pour chaque jeu de HRTF le nombre de fois où il a été validé selon les différents critères. Ces résultats montrent que si on applique un critère trop exigeant (MaxErrAll), on ne peut contenter que la moitié des auditeurs. Il semble illusoire

FIGURE 6 – Nombre de fois où chaque jeu (S1, S2, S3, S4) de HRTF a été validé pour un sujet au sens des différents critères. *Individual HRTF* : Nombre de fois où le jeu de HRTF individuelles a été validé. *Selection* : Nombre de fois où l'un des jeux de la sélection a été validé.

de chercher un jeu garantissant une erreur de localisation en élévation inférieure à 6 degrés. On se rend compte d'ailleurs que la moitié des sujets n'atteint pas ce niveau de performances avec leur propre jeu de HRTF individuelles. Il convient donc de s'en tenir au critère CorrBiais qui semble le mieux adapté à notre objectif.

5 Conclusion

Une sélection de 4 jeux de HRTF "prêt-à-porter" a été proposée, puis évaluée en termes de performances de

localisation. La sélection proposée s'est avérée exploitable pour 20 sujets (soit 83%) au sens du critère CorrBiais. Ce taux de satisfaction est du même ordre de grandeur que celui mesuré pour les HRTF individuelles : 17 sujets (soit 71%) ont validé leur jeu de HRTF individuelles au sens du critère CorrBiais. Pour les sujets qui présentent des fortes erreurs de localisation quel que soit le jeu, se pose la question de leurs performances naturelles de localisation qu'il conviendrait de vérifier avec des sources réelles.

La sélection de HRTF que nous avons constituée semble très bien adaptée au déploiement d'applications et de services de réalité virtuelle ou augmentée embarquant une synthèse binaurale dynamique. On peut même aller plus loin : elle est essentielle. En effet, sur la figure 6, on observe qu'aucun des jeux de la sélection ne ressort comme jeu universel. Les choix des sujets sont répartis de manière quasiment uniforme entre les 4 jeux. Or les performances de localisation peuvent être très dégradées avec un jeu qui n'est pas adapté. Il est donc nécessaire de proposer aux auditeurs une sélection de plusieurs jeux de HRTF si on veut satisfaire un large panel d'utilisateurs. Il reste toujours la solution de calcul sur mesure de HRTF individuelles pour les utilisateurs qui ne trouvent pas de jeu adapté dans la sélection. Pour cela, il conviendrait de développer des méthodes de mesure allégée.

Références

- [1] R. Nicol, Binaural Technology, *Audio Engineering Society* (2010).
- [2] O. Warusfel, Listen HRTF Database, <http://recherche.ircam.fr/equipes/salles/listen/index.html> (2003).
- [3] D. Schönstein, B.F.G. Katz, HRTF selection for binaural synthesis from a database using morphological parameters, *20th International Congress on Acoustics, ICA* (2010).
- [4] R. Nicol, P.Y. Michaud, Multi dimensional scaling of perceived dissimilarities between non-individual HRTFs : Investigating the perceptual space of binaural synthesis, *BiLi Project Deliverable* (2015).
- [5] R. Nicol, M. Emerit, E. Roncière, H. Déjardin, How to make immersive audio available for mass-market listening, *Technical Review EBU* (2016).
- [6] H. Déjardin, Web HyperRadio RadioFrance, <http://hyperradio.radiofrance.fr> (2018).
- [7] B.F.G. Katz, G. Parseihian, Perceptually based head-related transfer function database optimization, *The Journal of the Acoustical Society of America* **131**(2), EL99-EL105 (2012).
- [8] H. Bahu, Th. Carpentier, M. Noisternig, O. Warusfel, Comparison of Different Egocentric Pointing Methods for 3D Sound Localization Experiments, *Acta Acustica united with Acustica* **102**(1), 107-118 (2016).
- [9] F. Rugeles Ospina, M. Emerit, J. Daniel, A fast measurement of high spatial resolution head-related transfer functions for the BiLi Project, *International Conference of Spatial Audio* (2015).