



LIMSI – CNRS
Groupe Cognition Perception Usage



Institut de Recherche et de
Coordination Acoustique / Musique

ETUDE DES REPRESENTATIONS SPATIALES EN ENVIRONNEMENT SONORE

Master ATIAM

Acoustique, Traitement du signal et Informatique Appliqués à la Musique
Année 2009-2010

Lourdes García

sous la direction de Malika Auvray

Résumé

Ce travail, réalisé dans le cadre du projet scientifique NAVIG (Navigation Assistée par Vision artificielle et GNSS), a pour objectif l'étude des représentations spatiales mentales formées à partir de navigation en environnement virtuel sonore. Il expose un modèle expérimental visant à tester l'éventuel bénéfice de la présence d'un indice distal sonore dans l'environnement sur la stratégie employée pour appréhender l'espace (cadre de référence égocentrique vs. allocentrique).

Remerciements

Je tiens à remercier Malika Auvray, Mathieu Gallay et Gaëtan Parseihian pour m'avoir encadré, orienté et conseillé pendant ces quatre mois de stage.

Particulièrement merci à Gaëtan, pour m'avoir appris à utiliser le logiciel et me'avoir conseillé sur la partie de sonification de l'environnement ; et à Mathieu, pour la rigueur scientifique, les conseils sur la théorie et pour avoir corrigé toutes mes fautes d'orthographe. Merci de votre patience.

A Lucie et les autres stagiaires pour la bonne ambiance de travail.

Table des matières

1. Introduction	5
2. Etat de l'art sur la représentation mentale spatial	7
2.1. Représentation spatiale et carte cognitive. Carte vs. route -----	7
2.2. Représentation de l'espace chez les non-voyants -----	9
2.3. Caractéristiques de la représentation de l'espace chez les non-voyants -----	10
3. Etat de l'art sur la localisation auditive et la cognition spatiale	14
3.1. Le rôle de l'audition dans la connaissance spatiale -----	14
3.2. Localisation spatiale -----	16
3.3. Son binaural -----	19
3.4. Localisation auditive dans un environnement virtuel -----	20
4. Expérience sur la représentation mentale spatiale en environnement virtuel sonore	23
4.1. Hypothèse -----	23
4.2. Conception de et mise en œuvre d'une plateforme de navigation interactive	24
4.2.1. Description de l'environnement -----	24
4.2.2. Types de sons -----	26
4.2.3. Description du programme développé -----	27
4.2.4. Choix des sons de l'environnement -----	31
4.3. Procédure pré-expérimentale -----	33
4.3.1. Pré-test 1 -----	33
4.3.2. Pré-test 2 -----	35
4.4. Expérience -----	38
4.4.1. Participants -----	38
4.4.2. Matériaux -----	38
4.4.3. Procédure -----	38
4.4.4. Description de la tâche -----	41
4.4.5. Mesures effectuées -----	41
4.4.6. Hypothèse opérationnelle -----	44
4.4.7. Résultats -----	44
4.4.8. Discussion -----	47
5. Conclusion	49
6. Annexe 1	50

1. Introduction

L'étude qui suit s'inclut dans le cadre du projet scientifique NAVIG (Navigation Assistée par Vision artificielle et GNSS), qui a pour objectif d'augmenter l'autonomie des non-voyants et malvoyants dans le processus de navigation (<http://navig.irit.fr/>). Ce projet multidisciplinaire et innovant vise à mettre au point un dispositif de substitution sensorielle permettant aux non-voyants de se déplacer vers une destination voulue de façon fiable et de localiser et saisir des objets sans nécessité de les pré-équiper avec un composant électronique ou magnétique. Cet objectif sera atteint grâce à:

- Un module de géolocalisation par satellite du piéton.
- Un module de reconnaissance de formes: il s'agit d'un système de vision artificielle embarquée inspiré de la vision biologique. Ce système permet une reconnaissance robuste et la localisation précise de cibles (à partir de modèles préétablis) dans un temps de l'ordre de 10 ms. L'entrée du dispositif sera directement connectée au dispositif d'acquisition vidéo temps réel équipé de 2 caméras intégrées aux lunettes des utilisateurs.
- Un module de fusion de données provenant de la géolocalisation par satellite et du système de reconnaissance visuelle.
- Une interface utilisateur: L'interface de sortie du système NAVIG est purement auditive. Il s'agit d'un système de synthèse binaurale, qui augmentera la réalité avec des informations auditives permettant de localiser les cibles et d'obtenir diverses informations les concernant.

La conception d'un tel dispositif implique une connaissance a priori des processus de représentation de l'espace chez les personnes non-voyantes. Un grand nombre des recherches ont été consacrées à ce sujet depuis très longtemps. Dès le 17^{ème} siècle, des philosophes comme Molyneux, Locke et Diderot avaient déjà traité le problème de la spatialisation et le thème est toujours d'actualité de nos jours. Dans un premier temps, l'idée d'une représentation spatiale existant en l'absence de vision paraissait peu probable. Cependant, l'observation des activités réalisées par les aveugles dans leur vie quotidienne (ils arrivent à reconnaître la forme et la taille des objets, à réaliser des jugements de ressemblances et différences et à marcher et s'orienter dans des espaces plus ou moins grands) permet de remarquer qu'ils possèdent nécessairement une certaine connaissance de l'espace.

La manière dont sont encodées les informations spatiales par les voyants et les non-voyants vont faire naître des différences dans la représentation mentale dont elles feront l'objet, et ce sont ces différences qui vont principalement nous intéresser. Par exemple, alors que les personnes voyantes utilisent principalement un cadre de référence externe (allocentrique) pour construire leur représentation de l'espace, les non-voyants se basent plutôt sur un système de coordonnées intégrant les axes de leur propre corps (egocentrique).

La première partie de ce travail vise à fournir un aperçu des résultats plus importants obtenus jusqu'à présent sur la représentation mentale de l'espace, plus particulièrement en ce qui concerne les différences entre voyants et non-voyants (aveugles congénitaux ou dont la cécité est d'apparition plus tardive). La deuxième partie sera consacrée à la description de la méthodologie employée pour la réalisation de notre expérimentation. Enfin, la dernière partie concernera les résultats de l'expérimentation et leur implication dans le cadre du développement d'un système d'aide à la navigation pour aveugles.

2. Etat de l'art sur la représentation mentale spatial

2.1 Représentation spatiale et carte cognitive. Carte vs. route

La perception visuelle fournit les compétences et les ressources de base nécessaires pour permettre un fonctionnement autonome, efficace et sûr dans l'environnement physique et social. Elle nous permet de prendre conscience de notre position et d'actualiser les informations provenant de l'environnement. En effet, d'après de nombreux auteurs (Gärling, 1995 ; Gärling, Book et Lindberg, 1984 ; Russell et Ward, 1982), la connaissance de l'environnement acquise et mémorisée, joue un rôle important en permettant aux personnes de planifier et d'effectuer leurs déplacements. Ainsi, il est opportun de signaler l'importance de la vision dans les processus de navigation dans l'espace.

Les représentations cognitives de l'espace, qui résultent de nombreux processus mentaux et dynamiques, sont construites par l'individu en fonction de ses interactions avec l'environnement. L'étude de ces représentations cognitives spatiales conduit à appréhender les pratiques et les comportements spatiaux. Un nombre important de recherches ont été menées sur les aspects spatiaux des modèles mentaux. En conséquence de ces recherches, il a été montré que les individus sont capables de construire un « modèle mental spatial qui contient non seulement des repères, mais aussi des informations sur les propriétés et les relations géométriques de cet espace » (Doré et Mercier, 1992), ce que l'on appelle une « carte cognitive » (Tolman, 1948). Ce terme évoque, dans le sens où Tolman l'avait défini, une représentation interne de l'environnement.

Il est nécessaire de s'interroger sur cette information spatiale traitée par l'individu et organisée sous forme de représentation spatiale interne : Comment cette information est-elle acquise ? (à partir d'un modèle visuel, verbal ou autre) Quel type de perspective peut être construite ? (type route ou carte).

Ces deux types de représentations distinguant représentations d'ensemble (type carte), et représentations de trajets (type route) démontrent qu'il existe des différentes stratégies

de connaissance de l'espace. Les descriptions de type carte et de type route ont des caractéristiques spécifiques dont les origines se trouvent dans les différents cadres de référence employés : égocentrique (type route) vs. allocentrique (type carte). Dans un cadre de référence égocentrique, les entités sont représentées dans l'espace en un système de coordonnées intégrant les axes du corps du navigateur, alors que dans un cadre de référence allocentrique le système de coordonnées sous-jacent est externe, et intègre les axes géographiques terrestres.

Dans le cas d'une représentation de type route, la connaissance de l'espace se fait à partir de l'enregistrement en série de points de repère à mesure que l'individu navigue dans l'espace. Ce type de connaissance est séquentielle et dite sans plasticité, dans le sens où elle ne peut pas être réorganisée (si un point de repère a disparu, ou que le navigateur a un peu dévié, la structure est interrompue). Elle semble être suffisante quand on se déplace dans des lieux familiers en absence d'obstacles. Au contraire, la représentation de type carte est basée sur une connaissance des propriétés topographiques de l'environnement, telles que la localisation d'objets par rapport à un système allocentrique de coordonnées fixes. C'est donc une représentation indépendante de la position de l'individu et son rôle est crucial dans la capacité du navigateur à déterminer la direction des lieux situés en dehors de son champ de vision ou à établir des relations spatiales entre des lieux dont les liens n'ont pas déjà été explorés physiquement. L'organisation d'une description de type carte est souvent de type hiérarchique, c'est-à-dire, l'espace est divisé en différentes zones et chaque zone est décrite l'une après l'autre. Le type carte, contrairement à la représentation type route, permet une réorganisation des infos (détours, raccourcis, etc) (Conceptualisation d'O'Keefe et Nadel, 1978). D'ailleurs, la supériorité de la représentation de type carte sur la représentation de type route se confirme à un niveau expérimental.

Des nombreuses recherches ont été destinées à trouver si les représentations spatiales construites sur les bases de la navigation (par exemple à partir d'une perspective de type route) ont les mêmes propriétés métriques que les représentations à l'aide d'une carte symbolique (perspective de type carte). A cet effet, Noordzij et Postma (2005) ont mis en place une tâche consistant à comparer des distances à vol d'oiseau. Cette expérience obligeait aux participants à adopter une perspective aérienne en accord avec la description type carte et pas en accord avec la description type route. Les résultats montrent la supériorité d'une description type carte sur la description type route.

Cette méthodologie a été réutilisée en recherches ultérieures. Par exemple, Péruch, Chabanne, Nesa et Thinus-Blanc (2004) ont mis en place une expérience dans laquelle les participants sont invités à comparer mentalement les distances qui séparent les différents objets. Dans une première expérience, les deux distances à comparer avaient

le même point de départ, lorsque dans une deuxième expérience les points de départ étaient différents. Dans les deux cas, un plus grand nombre de bonnes réponses et un temps de réponse plus court sont observés à partir d'une perspective type carte. D'après ces résultats il a été signalé que l'estimation des distances est plus facile quand les représentations spatiales ont été construites à partir d'une description type carte qu'à partir d'une description type route d'un espace complexe.

2.2 Représentation de l'espace chez les non-voyants

Les recherches citées précédemment sur la représentation de l'espace sont basées sur la capacité humaine d'interaction avec l'environnement et sa capacité de navigation, qui permettent la mise à jour des informations spatiales reçues par le navigateur. Dans le cas des personnes voyantes cette interaction s'appuie fondamentalement sur l'organe de la vision. Quelles seront les conséquences de son absence sur la représentation de l'espace?

Face à une absence totale ou partielle de la vision, le processus de navigation et d'orientation devient particulièrement complexe. L'organe de la vision fournit des informations spatiales multiples, simultanées et dans un laps de temps très bref. D'ailleurs il a été montré que sans vision, les informations provenant de sources externes ne sont pas fiables, ni invariants, ni forcément corrélés avec les déplacements : la vision stabilise l'environnement alors que l'on s'y déplace (Thinus-Blanc et Gaunet ; 1997).

L'étude de la représentation de l'espace chez les non-voyants a été objet des nombreux et hétérogènes travaux de recherche. Pour classer ces études il est possible de réaliser une première distinction entre ce qu'on appelle champ proche et champ lointain. Les travaux réalisés en champ proche sont ciblés particulièrement sur les problèmes de reconnaissance et de représentation spatiale de la zone à portée de main des non-voyants. Dans ce cas les informations externes sont reçues à partir du sens tactile. Les recherches en champ lointain analysent les relations spatiales dans une zone plus large, seulement accessible, en absence de vision, à partir des autres modalités sensoriales, comme l'audition, le mouvement et la proprioception.

Par conséquent, les relations entre le champ proche et le champ lointain sont forcément plus complexes pour les non-voyants. Lorsque pour les voyants ces deux types de relations spatiales impliquent un système sensoriel : la vision, pour les non-voyants la façon d'accéder à ces deux espaces implique l'utilisation des autres modalités sensorielles : cénesthésique, haptique, auditive et proprioceptive. Ainsi, le transfert des

fonctions exercées par la vision sur ces autres modalités sensorielles oblige à une exploration séquentielle, plus lente et plus coûteuse de l'environnement. Une moins bonne performance des tâches spatiales pourrait être imaginée dans un premier temps.

Cependant, de nombreuses recherches ont montré que, à partir du feedback provenant de différents types de sources d'information externe, les personnes non-voyantes sont capables, comme les personnes voyantes, de construire une carte cognitive ou représentation mentale de l'espace (Casey, 1978 ; Rieser, Lockman et Pick, 1980; Dodds, Howarth et Carter, 1982 ; Hollyfiel et Foulke, 1983).

Les modalités sensorielles impliquées dans la représentation de l'espace chez les aveugles font encore aujourd'hui l'objet de nombreuses recherches. Les aveugles développent-ils, comme on le croit souvent, des capacités sensorielles tactiles et supérieures ? Dans un premier temps, aucune différence n'avait été trouvée ni dans les seuils tactiles de détection et de discrimination ni dans les seuils différentiels auditifs (Starlinger & Niemeyer, 1981). C'était donc plutôt sur l'orientation de l'attention et les modes de traitement que l'utilisation intensive des autres modalités pouvait agir, comme le montre la supériorité des aveugles dans l'identification des voix, la détection du nombre de voix dans un chœur ou la mémoire des chiffres (Miller, 1992).

Dans ce contexte, Noordzij, Zuidhoek et Postma (2007) ont montré que la stratégie basée sur l'expérience haptique ou acoustique peut être aussi juste que la stratégie basée sur l'expérience visuelle dans la réalisation d'un grand nombre de tâches cognitives et que les personnes non-voyantes sont probablement capables de compenser l'absence de vision en renforçant leurs capacités auditives et tactiles.

Cette hypothèse a été confirmée récemment. Une étude neuropsychologique de L'Université de Montreal a montré que le cerveau des non-voyants fait l'objet d'une restructuration qui lui permet d'utiliser une partie du cortex visuel pour améliorer la capacité auditive. Cette réutilisation du cortex visuel fait que les aveugles sont plus performants que les voyants pour des tâches de reconnaissance de notes, d'octaves et de détection de la provenance de sons. Ainsi, il a été constaté que les aveugles utilisent à la fois le cortex auditif et le visuel pour entendre. Ceci est possible grâce à la plasticité du cerveau, qui permet aux aveugles de mieux affronter leur condition sensorielle désavantageuse (Gougoux et Lepore, 2009).

2.3 Caractéristiques de la représentation de l'espace chez les non-voyants

Quand un aveugle se déplace dans un environnement inconnu, son sentiment de sécurité est moindre car il reçoit moins d'informations de l'extérieur qu'un navigateur voyant, et il n'a presque pas d'anticipation perceptive. Etant donné qu'il n'a pas accès à certains éléments qui serviraient de repères pour les voyants, il est plus compliqué pour lui de garder une orientation correcte. Même avec un apprentissage au préalable des stratégies de mobilité, il a été signalé que les non-voyants éprouvent des difficultés à avoir des déplacements appropriés, c'est-à-dire, indépendants, en confiance et confortables comme celui des voyants (Foulke, 1982). Les difficultés rencontrées viennent, selon ces auteurs, d'une quantité d'information insuffisante reçue de l'environnement d'un manque d'expérience ou encore d'une absence des stratégies nécessaires au transfert de l'information à d'autres modalités sensorielles plus efficaces.

Toutefois, tout ceci n'implique pas que les aveugles ne sont pas capables de se déplacer efficacement dans l'espace. Ainsi, des études ultérieures ont mis en évidence que les non-voyants peuvent réaliser des tâches présumées comporter des compétences visuelles avec une performance similaire à celle des voyants (Hollyfield et Foulke, 1983). Cela est cohérent avec la *théorie de la différence* (cité dans Hatwell, 2003), selon laquelle les aveugles, même congénitaux, ont les mêmes capacités spatiales et ils peuvent organiser les informations spatiales de la même façon que les voyants. Cependant, étant donnés les caractéristiques des systèmes perceptifs autres que la vision, les tâches d'acquisition, codage, stockage et utilisation de l'information spatiale exigent plus de temps et un plus grand effort cognitif chez les non-voyants.

Dans la mesure où l'absence totale de vision réduit la quantité d'informations de l'environnement disponible, il a été suggéré que la connaissance spatiale des aveugles reposerait plutôt sur des informations proches. Ces informations proches sont intégrées de manière séquentielle (ex : marcher sur une pelouse, puis sur une route, puis sur des pavés) qui vont être organisées de façon égocentrée, alors que les voyants peuvent utiliser la vision lointaine pour intégrer simultanément toutes ces informations. Selon Millar (1994), ce déficit de perception des informations distales va donc entraîner, chez l'aveugle de naissance, une tendance à se reposer sur des stratégies égocentriques plutôt qu'allocentriques dans les tâches spatiales. Ceci expliquerait pourquoi les aveugles de naissance sont plus enclins à organiser l'information spatiale sous forme de route plutôt que sous forme de carte. Noordzij, Zuidhoek et Postma (2007) ont testé les qualités de la représentation mentale d'un environnement décrit verbalement, à la manière d'une carte ou à la manière d'une route, chez des individus voyants et non-voyants. Les deux types de description (carte ou route) donnent lieu chez les non-voyants (de naissance et

tardifs) aux mêmes effets que ceux classiquement constatés chez les voyants. Dans la tâche de reconnaissance avec amorçage, on présente aux participants deux noms d'éléments. Ces éléments ont la propriété d'être soit présents, soit absents de la description verbale initiale. La tâche des participants consiste à dire le plus rapidement possible si le deuxième élément cité fait partie ou non de l'environnement décrit auparavant. Comme chez les voyants, quand les deux éléments font bien partie de la description, on observe un effet d'amorçage : plus les deux éléments sont géographiquement proches, plus les participants sont prompts à répondre. De plus, et toujours comme chez les voyants, cet effet disparaît quand la proximité entre les deux éléments existe dans le texte descriptif mais pas dans l'espace décrit. Les auteurs en concluent donc que les non-voyants (tardifs et de naissance) possèdent bien une représentation mentale spatiale (et non verbale) de l'environnement décrit verbalement, et que cette représentation spatiale contient même des informations métriques. La deuxième tâche à laquelle les participants sont soumis est une tâche de comparaison de distances entre éléments de l'environnement. Encore une fois, on observe chez les trois groupes de participants le même effet : plus la différence entre les deux distances à comparer est grande, plus les temps de réponse sont courts. Sur l'organisation de la représentation spatiale de l'environnement, les trois groupes de participants sont donc comparables. En revanche, l'étude de l'influence du type de description révèle que les participants non-voyants (tardifs et de naissance) présentent une représentation mentale de l'espace plus robuste quand elle est réalisée à partir d'une description de type route, alors que les voyants semblent plus à l'aise à partir de la description carte, ce qui confirme bien les dires de Millar (1994) sur le cadre de référence adopté.

Une solide représentation mentale de l'espace étant plus à même d'être mise en place à partir d'un cadre de référence allocentrique, on doit s'interroger sur la capacité des non-voyants à adopter ce type de stratégie, malgré une tendance naturelle à se baser sur les informations égocentriques. Tinti, Tamietto et Cornoldi (2006) ont conduit une expérience pour comparer le niveau de performance des non-voyants congénitaux par rapport aux voyants dans la réalisation de tâches basées sur une représentation de type carte. Les participants sont amenés à réaliser un parcours fait de virages et de lignes droites dans un espace de 10m x 10m et les auteurs testent la qualité de la représentation cartographique mise en place suite aux déplacements. Même si ces parcours sont réalisés dans des environnements relativement restreints, les résultats signalent que l'absence d'expérience visuelle ne semble pas empêcher la mise en place des représentations type carte. Les auteurs signalent même dans certaines conditions, de meilleures performances pour les non-voyants que pour leurs participants doués de vision.

Il semble donc, que dans certaines circonstances, les aveugles sont capables de mettre en place un cadre de référence allocentrique. Des différences sont susceptibles d'exister à ce sujet entre les aveugles congénitaux et les aveugles tardifs. En effet, l'âge d'apparition

de la cécité a notamment une influence sur les mécanismes d'interaction avec l'environnement mais aussi sur la perception des éléments distaux. Les aveugles tardifs sont censés avoir déjà appréhendé l'espace à la manière d'un voyant avant de perdre la vue, et devraient donc logiquement bénéficier d'un avantage spatial sur les aveugles congénitaux. Mais un manque de convergence est observé dans les recherches réalisées au ce sujet. Par exemple, Loomis, Klatzy, Golledge, Cicinelli, Pellegrino et Fry (1993) ont réalisé une expérience basés sur la navigation avec des non-voyants congénitaux, des non-voyants tardives et des voyants avec les yeux bandés. L'expérience était composée de plusieurs tâches : réaliser des inférences complexes, refaire de trajets multi-segments en sens inverse et pointer vers des cibles après la locomotion. Les résultats ont montré une performance similaire entre les trois groupes.

Cette diversité de résultats en ce qui concerne l'influence de l'âge d'apparition de la cécité est due à de nombreux facteurs. Des différences quant à la complexité de la tâche à réaliser, le temps et le type d'entraînement, le niveau de mobilité du sujet ou même le niveau de familiarité de l'individu avec son environnement peuvent faire varier les interprétations des résultats trouvés. Un des facteurs les plus importants qui peut faire diverger les résultats est la capacité des non-voyants de naviguer dans l'espace de façon indépendante.

En bref, si des différences en la réalisation de tâches spatiales sont observées en défaveur des aveugles précoces, elles sont dues à la difficulté d'apprentissage des techniques de traitement de l'espace et à un moindre accès aux informations externes. De même, si ces différences n'existent pas ou si celles-ci sont en faveur des aveugles, elles résultent de l'usage intensif des autres modalités sensorielles compensatoires et de la mise en œuvre d'autres modes de traitement moins dépendants des données perceptives (Hatwell, 2003).

3. Etat de l'art sur la localisation auditive et cognition spatiale

3.1 Le rôle de l'audition dans la connaissance spatiale

Dans ce paragraphe, nous allons aborder le rôle de l'audition dans la connaissance spatiale. Nous allons voir comment, à partir de la perception des sons dans l'espace et des déplacements dans cet espace, un sujet voyant ou non-voyant est capable de s'en faire une représentation mentale.

Chez les sujets voyants, dès qu'un son inattendu se fait entendre, l'ouïe guide le mouvement de la tête dans la direction du son pour que la vision puisse entreprendre le repérage final et l'examen de la source sonore. Chez les sujets privés de vision, le repérage auditif des objets sonores à distance devient très utile pour guider les tentatives d'appropriation manuelle de ces objets ou pour les informer sur la présence éventuelle d'obstacles au cours de leurs déplacements. L'audition pourrait jouer, donc, un rôle décisif pour la « conscience de situation » d'un individu.

Cependant, la perception spatiale ne se réduit pas à la simple analyse de notre environnement. Elle participe surtout à l'expérience vécue sur laquelle se fonde notre cerveau pour construire des représentations internes de l'espace. Notre cerveau est ainsi capable de réactualiser de façon très précise une représentation mentale durant la locomotion. La preuve en est le fait que, malgré l'absence visuelle, nous montrons une certaine habileté à nous déplacer tout en évitant les obstacles. Cet exemple illustre parfaitement les capacités d'actualisation spatiale, dont Loomis et Klatzy, entre autres, font l'étude. Ils proposent un modèle d'étapes, comme celui montré dans la figure 3.1, commun aux modalités visuelle, auditif et linguistique de connaissance de l'espace.

La phase d'encodage fait appelle à la phase durant laquelle l'image spatiale est créée. Elle est suivie d'une phase d'apprentissage, durant laquelle la représentation se stabilise en mémoire. Ensuite, grâce au déplacement, le cerveau réalise une phase d'actualisation spatiale mentale. En bref, la représentation spatiale peut fonctionner comme une carte géographique sur laquelle le cerveau réactualise la position. Les stratégies

d'actualisation de l'espace, egocentrique vs. allocentrique (ou combinaison des deux), dépendront en grande partie des facteurs individuels et de l'environnement considéré.



FIG. 3.1 - Les deux étapes du modèle de l'encodage et de l'actualisation spatiale d'après Loomis et al.

Plusieurs recherches ont été basées sur le rôle de l'audition dans la connaissance spatiale. Dans une expérience réalisée par Loomis et Al (1998), le but était de comparer, dans un premier temps, les capacités d'actualisation respectives de la vision et de l'audition. Les principaux résultats sont les suivants :

1.- L'individu est capable d'actualisation dans les deux conditions (visuelle, auditive), avec des performances similaires, -bien que l'on observe des différences en termes d'encodage et d'apprentissage.

2.- Les propriétés (e.g. la précision ou la prégnance) de la représentation spatiale sont quasiment les mêmes pour ces deux modalités sensorielles. La vue et l'ouïe produisent directement des indices spatiaux, au contraire de la modalité linguistique (les descriptions verbales doivent être converties afin de permettre une représentation mentale manipulable).

3.- L'expérience visuelle n'est pas une condition nécessaire au développement des capacités d'actualisation puisqu'il y a une équivalence fonctionnelle entre les modalités de la vision et l'audition.

Cependant, il n'est pas suffisant de montrer que l'audition joue un rôle important dans la connaissance de l'environnement, au moins pour les mécanismes de substitution sensoriel auxquels on s'intéresse. Il est nécessaire de déterminer si les hypothèses citées ci-dessus peuvent être vérifiées dans le cas des tâches spatiales plus proches de l'expérience quotidienne. Plusieurs études (Loomis et Al., 2002 ; Meijer, 1992) ont été consacrées à comparer la performance des modalités sensorielles dans des tâches nécessitant un grand effort cognitif, comme une tâche d'actualisation spatiale¹ et de détermination de direction depuis un point de vue allocentrique. C'est dans ce type de tâches qu'est attendue une performance équivalente entre les deux modalités sensorielles à comparer.

¹Le terme actualisation spatiale fait référence à la capacité à encoder la position de différents éléments de l'espace et à mettre à jour ces informations spatiales en fonction des déplacements effectués.

En effet, le résultat des recherches ont montré une équivalence fonctionnelle entre la vision et l'audition en tâches d'actualisation spatiale. En outre, le fait que le langage spatial et la perception spatiale mènent à une performance équivalente pour la réalisation de tâches d'actualisation spatiale et de détermination de direction allocentrique, suggère que les représentations spatiales ne dépendent de la modalité sensorielle (Gonot, 2008).

Maintenant que nous avons fait le point sur les performances du système auditif pour la cognition spatiale, il nous faut prendre en considération le processus mis en œuvre par le système auditif pour la localisation de sources sonores.

3.2 Localisation auditive

Au début du 20^è siècle, Lord Rayleigh a été le premier à expliquer comment l'être humain localise les sons. Le paramètre le plus important pour déterminer la position d'une source sonore est la différence entre les signaux sonores parvenant aux oreilles. Les mécanismes de localisation directionnelle d'une source sonore sont relativement bien compris de nos jours.

La localisation d'une source sonore, ainsi que celle d'une cible visuelle, est tridimensionnelle. Elle est spécifiée en utilisant des coordonnées sphériques : azimuth, élévation (pour exprimer la direction) et distance (voir figure 3.2).

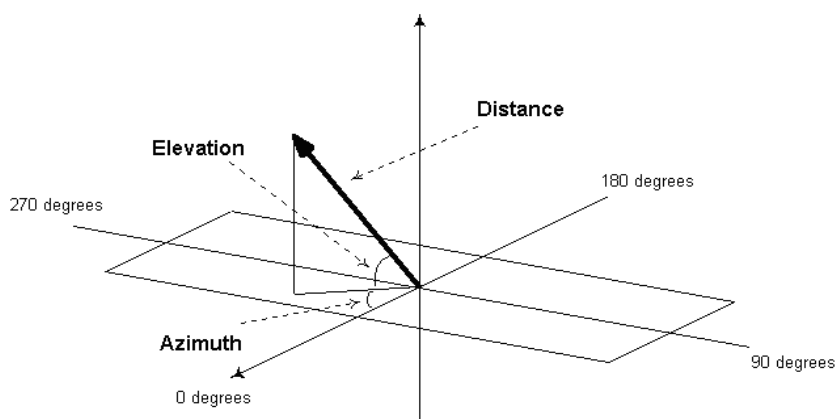


FIG. 3.2- Vecteur de position d'une source sonore par rapport au centre de la tête de l'auditeur exprimé en azimuth, élévation et distance.

Ainsi, le processus de localisation est basé fondamentalement sur la définition des deux indices, appelés indices binauraux ILD (*Interaural Level Difference*) et ITD (*Interaural Time Difference*) - et la fonction de transfert de la tête (*Head Related Transfer Functions*). Le son est transmis depuis la source jusqu'à l'auditeur à travers des chemins directs et indirects. La réflexion, absorption et diffraction du son produite par la tête sphérique fait qu'une plus petite quantité d'énergie acoustique arrive à l'autre cote de la tête (ce qu'on appelle "*acoustic head shadow*"). Cette caractéristique permet de définir l'ILD, qui correspond à une mesure de la différence du niveau du son qui parvient à chaque oreille.

L'autre indice nous permettant de localiser les sources est l'ITD. Il se produit lorsque la distance de la source du son aux deux oreilles est différente, ce qui entraîne des différences dans les temps d'arrivée du son sur les deux oreilles. Lorsque la source sonore est directement en face de l'auditeur, l'ITD vaut zéro, et son valeur augmente avec l'angle que forme la source avec le plan médian (voir figure 3.3).

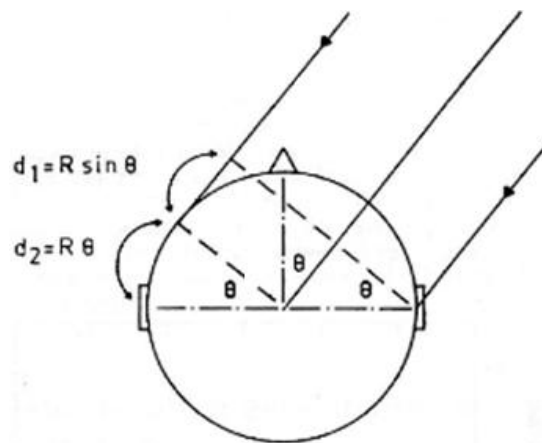


FIG 3.3 - Evaluation de la différence de marche Δd aux oreilles pour une source à l'infini d'azimut θ . La construction géométrique montre que : $\Delta d = d_1 + d_2 = n(\theta + \sin \theta)$. La différence interaurale des temps d'arrivée correspondante est donc : $\Delta t = t/c(\theta + \sin \theta)$

L'ITD et l'ILD donnent lieu à une détection latérale des événements sonores (azimut). Ils sont cependant insuffisants pour discriminer complètement deux sources de provenances différentes. Ils sont tous les deux constants pour des endroits différents ayant le même angle latéral, qui correspond à l'angle entre l'axe inter-aural et la direction de la source. Par contre, il existe une infinité de directions qui produisent une ITD constante. L'ensemble de ces directions forme ce que l'on appelle le «cône de confusion» (voir figure 3.4). Il tient son nom du fait qu'il est difficile pour un auditeur de discerner deux points diamétralement opposés sur ce cône. Ce phénomène est appelé *confusion avant-arrière*. Il est du, donc, à l'ambiguïté des différences inter-aurales qui sont symétriques par rapport à l'axe inter-aural, défini par le canal auditif des deux oreilles.

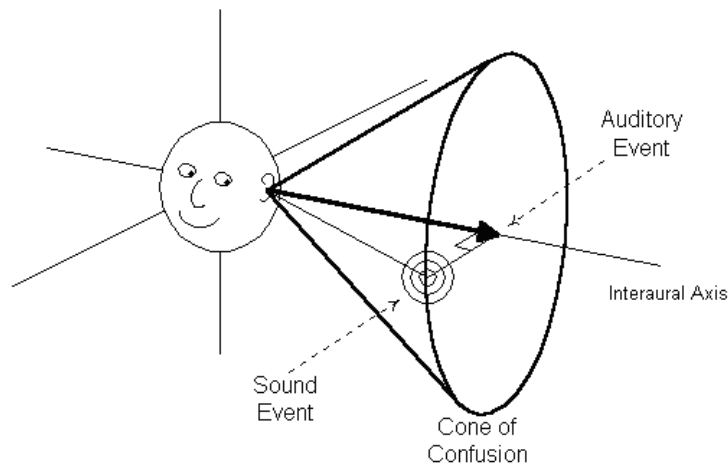


FIG. 3.4 - Le cône de confusion se produit uniquement quand l'LD ou ITD sont présents. Ici, la latéralisation à gauche de la personne se produirait afin de faire correspondre la projection du vecteur de son de la source à l'axe inter-aural.

Pour décrire complètement la manière dont le son qui arrive aux canaux auditifs est modifié par la tête et les épaules de l'auditeur, il faut définir des indices spectraux caractérisés par des fonctions appelées HRTF. Ces indices spectraux correspondent aux modifications de l'amplitude et de la phase du spectre de la source sonore, dues aux multiples réflexions, diffractions et autres obstructions, provoquées par les épaules, la tête et surtout l'oreille externe. Ces phénomènes sont décrits par un couple de HRTF : l'une entre la source sonore et l'oreille gauche, et l'autre entre la source sonore et l'oreille droite.

Ainsi, les HRTF apportent des indices spectraux qui complètent la localisation directionnelle dans le plan médian, affinent la position verticale et résolvent notamment les confusions avant -arrière.

En ce qui concerne la détermination de la distance à la source, le premier indice à observer est l'intensité du son. Dans une chambre anéchoïque, l'intensité du son d'une source à intensité constante, décroît de 6 dB quand on double la distance. Mais l'intensité de son n'est pas un indice absolu pour mesurer la distance, elle nous donne seulement information relative sur le changement de distance de la source. Un autre indice pour déterminer la distance relative est le changement spectral, du à l'absorption sélective des hautes fréquences à travers l'air. Finalement, si on veut pour déterminer la distance à la source de façon absolue il faudra avoir en compte le rapport champ direct-champ réverbérant. Ce rapport diminue à mesure que la distance à la source augmente. Par contre, il est difficile pour le système auditif d'identifier la distance d'une source sonore dans l'absolu. Une évaluation relative (« plus proche » ou « plus loin ») semble mieux maîtrisée. (Loomis et Al, 1999).

3.3 Synthèse binaurale

La synthèse binaurale est une technique de restitution audio permettant de spatialiser des sources virtuelles à l'aide d'un casque présenté aux deux oreilles. Le raisonnement est simple : si nous avons deux capteurs (nos deux oreilles), pour reproduire un espace sonore il suffit de deux sources sonores (une pour chaque oreille) présentant directement aux oreilles les informations sonores qu'elles auraient reçues si le sujet avait pris place dans l'espace sonore naturel.

Le principe de la technique binaurale est de filtrer une source sonore monophonique par le biais des fonctions de transfert acoustique mesurées entre une source sonore émettant dans une direction donnée et les canaux auditifs d'un individu (*HRTF*) (voir figure 3.5). Cette fonction de transfert binaural contient tous les indices acoustiques utilisés par le système auditif pour inférer la localisation spatiale d'un objet et peut donc permettre de donner à une source monophonique la direction 3D désirée.

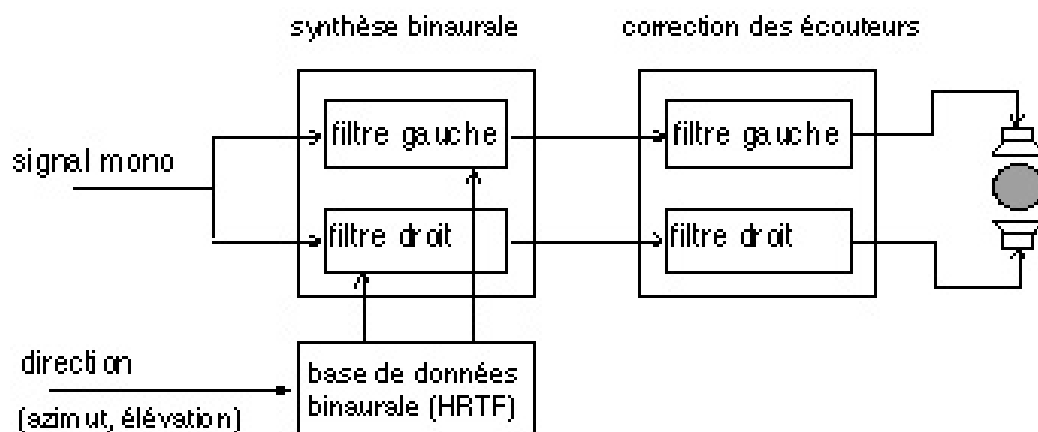


FIG 3.5 - Principe de la synthèse binaurale : La direction de la source sonore est contrôlée à l'aide de deux filtres audio numériques déduits des HRTF

Une grande quantité de mesures de fonction de transfert sont prises pour constituer une base de données de filtres acoustiques. Ces filtres peuvent être utilisés dans une application en temps réel pour synthétiser de manière dynamique la localisation des différentes sources composant une scène sonore virtuelle (Nguyen, Suied, Viaud-Delmon & Warusfel, 2009).

Le processus de localisation auditive par synthèse binaural pourrait entraîner quelques problèmes, parmi lesquels on peut signaler la confusion avant-arrière (une source

spatialisée en face de l'auditeur est souvent perçue comme provenant de derrière), commentée dans la partie précédente. Le système auditif peut résoudre ces ambiguïtés avec le mouvement latéral de la tête : l'ITD d'une source frontale et l'ITD d'une source située à l'arrière varient en sens inverse lorsque l'on tourne la tête de gauche à droite.

Une autre question d'importance concerne le fait que les HRTFs sont reflet de la morphologie humaine, donc une synthèse binaurale optimale devrait être individualisée en fonction de l'utilisateur pour une meilleure localisation. La solution adoptée consiste normalement à sélectionner, dans la base de données de filtres acoustiques, l' HRTF optimale, c'est-à-dire celle qui offre le positionnement de la source le plus réaliste. (Afonso, Katz, Blum, Jacquemin & Denis ; 2005).

Cependant, l'écoute binaurale offre des nombreux avantages. En plus de permettre la localisation des sources sonores dans l'environnement, les indices binauraux permettent à l'auditeur de sélectionner une source provenant d'un endroit précis pour lui apporter une attention plus particulière (attention sélective). Cette capacité est extrêmement importante quand il ya plusieurs sources sonores simultanées dans l'environnement. Quand bien même les indices spectraux n'amélioreraient pas la précision de la localisation proprement dite, de nombreuses études montrent qu'ils facilitent la ségrégation des sources dans une scène sonore. En effet l'utilisation de HRTF améliore de façon significative les performances lors d'une tâche d'attention sélective.

3.4 Localisation auditive dans un environnement virtuel

La simulation de sources sonores virtuelles ou auralisation est un facteur clé dans la production d'environnements sonores virtuels. Les techniques de spatialisation du son permettent aujourd'hui de simuler ces sources sonores virtuelles placées dans l'espace autour de l'auditeur. En particulier, les effets sonores spatialisés, dont la direction de provenance est fidèlement reproduite aux oreilles de l'auditeur, sont particulièrement importants pour localiser les objets, séparer des signaux sonores simultanés et donner des indices sur les caractéristiques spatiales de l'environnement (taille, matériaux, etc.).

L'auralisation, appelée aussi rendu sonore, est un vaste sujet qui regroupe trois problématiques principales : synthèse et contrôle de sons, simulation des effets de propagation du son dans l'environnement et enfin, perception et restitution aux oreilles de l'auditeur.

Pour la première problématique le but est de produire les sons résultant de l'interaction de l'utilisateur avec l'environnement. Suivant les effets voulus, on pourra travailler avec

des sons enregistrés ou purement synthétiques. Une fois synthétisés ou enregistrés, les signaux sonores sont ensuite traités de manière à reproduire les effets de propagation du son dans l'environnement de synthèse (occlusions par les obstacles, réflexions sur les parois et réverbération, effet Doppler pour les sources en mouvement). Là encore, ces effets dépendent fortement de la géométrie de l'environnement, de la position des sources et de l'auditeur. Cette partie peut être négligée dans les applications où la précision physique plus que le « réalisme » est suffisante, comme c'est le cas de notre expérience. En fin, la dernière partie du processus est la restitution du son aux oreilles de l'auditeur de manière la plus fidèle possible. L'acoustique virtuelle s'intéresse tout particulièrement à la fidélité de la restitution spatiale du son, la spatialisation. La spatialisation du son est un traitement spécifique qui vise à reproduire aux oreilles de l'auditeur la sensation correcte de la direction de provenance des sons virtuels. C'est l'un des composants les plus importants dans la production de son pour la réalité virtuelle.

La spatialisation auditive peut être simulée à partir d'un casque audio ou des haut-parleurs. Les techniques de synthèse et enregistrements binauraux sont vouées à une écoute au casque, qui permet de contrôler efficacement les signaux envoyés à chacun des tympanes de l'auditeur. En situation d'écoute par haut-parleurs, l'information sonore destinée à chaque oreille est perturbée par la présence de trajets croisés, qui bouleverse l'organisation spatiale et entraîne un repli de l'ensemble de la scène dans la zone délimitée par les deux haut-parleurs (voir figure 3.6).

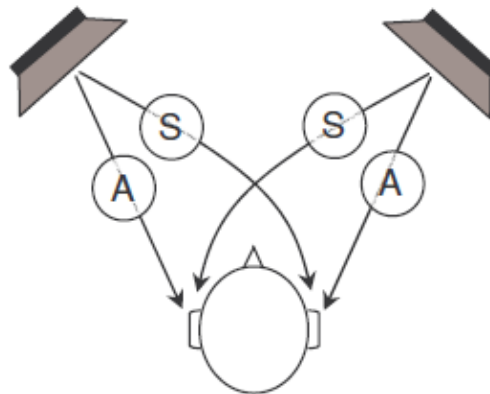


FIG. 3.6 - Trajets directs (A) et croisés des haut-parleurs aux oreilles en restitution stéréophonique standard

En général, le casque permet un contrôle plus précis des indices spatiaux, d'une part parce que les signaux qui arrivent aux oreilles de l'auditeur peuvent être contrôlés indépendamment et d'autre part parce qu'il n'y a pas de sons indirects qui arrivent aux oreilles (pas de écho ou réverbération).

Malgré l'évidence rien n'indique qu'en pratique (i.e. durant une tâche de navigation), la synthèse binaurale améliore significativement les performances de localisation d'une cible unique par rapport à la stéréophonie mixte classique. Cependant, il a été montré que l'utilisation de HRTF améliore de façon significative les performances lors d'une tâche d'attention sélective («effet cocktail party»). Une telle habileté peut faire nettement la différence lorsque l'individu doit se focaliser sur une cible sonore et s'orienter vers elle malgré la présence de multiples sources concurrentes. (Gonot, 2008).

Loomis (1985) avait proposé que les sons perçus comme provenant des points de l'espace (dits sons virtuels) pourraient être utilisés pour guider les non-voyants dans la navigation. Dans un premier temps cette idée apparaît comme une alternative à l'information de guidage fourni par la parole synthétisée. Dans un souci d'étudier l'allocation de ressources attentionnelles, Klatsky, Marston, Giudice, Golledge et Loomis (2006) ont mis en lumière ce qui semble être un avantage considérable pour les sons virtuels par rapport au langage. Alors que dans des conditions de navigation « normales » (la seule tâche est de se rendre d'un point A à un point B), un système d'informations verbales (« gauche, droite, tout droit ») révèle des performances égales à celles produites avec un système de sons spatialisés, ce dernier s'avère supérieur au verbal quand la navigation est parasitée par une tâche distractive. En effet, les temps et les distances de parcours sont dans ce cas plus courts avec les sons spatialisés, et les résultats à la tâche de comptage sont meilleurs. Il apparaît donc que les sons spatialisés sont étant traités à un niveau perceptif alors que le langage nécessite une médiation cognitive supplémentaire. Dans le premier cas, cela permet de libérer une certaine charge de travail qui va pouvoir être allouée à une autre tâche, comme par exemple parler avec quelqu'un pendant le trajet, garder en tête la liste des courses ou préparer un rendez-vous.

4. Expérience sur la représentation mentale spatiale en environnement virtuel auditive

4.1 Hypothèse

Bien que les non-voyants soient naturellement orientés à utiliser des représentations égocentriques de l'environnement, il a été montré la supériorité d'une perspective de type carte par rapport à une perspective de type route dans la réalisation des tâches spatiales, comme l'orientation dans l'espace ou l'inférence de distances entre objets. De plus, et malgré des résultats parfois contradictoires sur l'effet de l'âge d'apparition de la cécité, les non-voyants -précoces et tardifs- semblent capables sous certaines circonstances, d'adopter un point de vue allocentrique dans des tâches spatiales.

Le but de la présente étude est de tester l'éventuel bénéfice apporté par un indice distal dans un environnement purement sonore. Puisqu'un environnement sonore est censé être dépourvu d'indices distaux fiables (la grande majorité des sources sonores étant en mouvement et ne pouvant donc pas constituer un indice stable), nous nous demandons si un tel type d'information distale pourrait aider les participants, voyants ou non, à adopter une perspective allocentrique dans une tâche spatiale.

La capacité des non-voyants à élaborer des représentations mentales de l'espace à partir de leur interaction avec un environnement auditif a été étudiée précédemment. En 2006, Ohuchi, Iwaya, Suzuki, & Munekata ont construit un dispositif de navigation en environnement sonore. Le système est composé d'un capteur du mouvement de la tête qui envoie les données à l'ordinateur pour sélectionner la HRTF adéquate. L'environnement virtuel est créé par sources sonores qui vont convoluer avec les HRTFs (*Head-related transfer functions*) pour contrôler la localisation des sons. L'objectif du dispositif était d'encourager l'utilisateur à s'entraîner ' construire des cartes cognitive et à les utiliser plus efficacement. La tâche consistait dans un premier temps à se familiariser avec l'environnement grâce à une carte tactile présentant la forme du labyrinthe et la position des points de repère sonores. Ensuite, les participants devaient se déplacer dans l'environnement virtuel à l'aide d'une manette pour retrouver les

points de repères et mémoriser leur position. Enfin, ils devaient reconstruire l'environnement exploré et y placer les points de repères. Les résultats ont montré que le système favorise la formation des cartes spatiales cognitives et pourrait donc aider à améliorer l'orientation et la mobilité de l'utilisateur dans l'espace.

Dans le cadre de notre expérience, les participants pourront naviguer dans un espace virtuel purement auditif et recevoir des informations sonores tant en champ proche qu'en champ lointain. Grâce à un système de synthèse binaurale, l'utilisateur entendra un son provenir de l'endroit précis de l'espace où se trouve l'objet distal. Il pourra donc utiliser cet objet comme point de repère et ainsi mieux se situer d'ans l'environnement. Si l'hypothèse proposée est réelle, les utilisateurs ayant accès à ce type d'information auront une meilleure performance dans la réalisation des tâches spatiales proposées.

L'expérience mis en place à cet effet va consister à faire naviguer l'utilisateur dans un zoo virtuel constitué de chemins et des cages occupés par des animaux qu'il pourra entendre. Sa tâche sera, dans un premier temps, de faire des parcours définis en se laissant guider par l'expérimentateur et, ensuite, réaliser des inférences spatiaux à partir des informations de l'environnement qu'il avait perçu et retenu dans la réalisation de ces premiers parcours.

4.2 Conception et mise en œuvre d'une plateforme de navigation interactive

4.2.1 Description de l'environnement

La passation de l'expérimentation s'est faite dans une salle calme du laboratoire, elle a nécessité l'utilisation d'une manette sans fil, d'un capteur d'orientation spatiale, d'une chaise rotative ainsi que de l'ordinateur gérant la présentation du programme et des sons.

Le choix de l'environnement physique dans lequel réaliser l'expérience a été guidé par des contraintes pratiques et matérielles. Nous ne disposions pas de capteur de tête sans fil, ce qui nous a imposé de réaliser l'expérimentation dans un environnement restreint, ainsi qu'à obliger les participants à rester immobiles. C'est pour pallier à cette contrainte d'immobilité que nous avons choisi de forcer les participants à utiliser une manette pour se déplacer et à tourner physiquement sur une chaise rotative, de manière à adopter une

perspective égocentrique sur leur déplacements et ainsi garder un certain degré d'immersion. De plus, la tâche imposée étant à la fois complexe et inhabituelle, nous avons décidé de simplifier l'environnement sonore de manière à ce que les seuls virages à effectuer soient des virages à 90 ou 180°. A cet effet, nous avons installé au sol quatre repères podo-tactiles autour de la chaise, auxquels les sujets ont l'instruction de se référer pour effectuer les virages.

L'environnement virtuel mis en œuvre pour la réalisation de notre expérience consiste en une scène purement auditive représentant un zoo. Ce zoo est composé uniquement de chemins qui définissent des cages occupées par des animaux que l'utilisateur pourra entendre grâce à un casque audio. Tous les scénarios ne sont constitués que de lignes droites et de virages à angle droit. Afin d'augmenter la sensation de déplacement, l'utilisateur pourra entendre le son de ses pas. Il y aura toujours un animal par cage, mais les cages sont grandes et il n'est possible d'entendre l'animal si on se situe du bon côté de la cage, si bien que la zone dans laquelle le participant sera susceptible d'entendre l'animal correspond à une zone de 2 carrés.

Au milieu exact de l'environnement il y aura, pour la moitié des sujets, une balise sonore. Cette balise donne trois indices à l'utilisateur pour lui permettre de s'orienter par rapport à elle : elle s'actualise en fonction de la position et de l'orientation de sa tête (son 3D), le son s'intensifie et le rythme s'accélère au fur et à mesure que l'utilisateur s'approche.

Le son des animaux, de la balise sonore et des pas sont spatialisés.

4.2.2 Types de sons

Type du son	Signification
Carrefour	Son stéréophonique indiquant qu'il est possible de tourner à gauche (son dans l'oreille gauche) ou à droite (son dans l'oreille droite).
Mur	Son de contact entre le personnage virtuel et un mur, indiquant au participant qu'il lui est impossible d'aller plus loin dans cette direction (un son pour mur en marche avant et un son pour mur en marche arrière)
Pas	Son indiquant que l'utilisateur est en mouvement (en avant ou en arrière). Il est proportionnel à la vitesse de déplacement. Le son des pas est un son stéréo pour différencier le son du pied gauche du son du pied droit.
Animaux	Son spatialisé simulant les animaux à l'intérieur des cages. Le rayon d'écoute des animaux est modifiable (paramètre de réglage).
Point de repère (PR)	Son spatialisé simulant la balise sonore située au milieu de l'environnement virtuel. Elle est entendue périodiquement dans tout l'environnement. La fréquence de répétition et l'intensité sont dépendantes de la distance entre le navigateur et la balise (plus le personnage virtuel se rapproche, plus le rythme est rapide et plus l'intensité sonore augmente).

TAB. 4.1 – Description des différents types de sons qui composent l'environnement virtuel

4.2.3 Description du programme développé

Cette partie a pour objectif de décrire le fonctionnement du programme développé en MAX/MSP pour la conception du zoo virtuel. La création du programme n'a pas fait partie de mon travail de stage, il m'a fallu par contre comprendre le fonctionnement de cet environnement pour la réalisation de modifications. Mon principal apport a été, la création des différents scénarios, le choix des types de sons, le paramétrage initial du système et la résolution du problème de perception de la distance.

L'interaction entre l'utilisateur et le zoo virtuel s'effectue grâce à deux dispositifs périphériques :

- Un capteur d'orientation, mouvement et position, situé sur le casque audio de l'utilisateur. Le processeur interne du capteur fournit l'orientation de la tête dans les 3 axes : *yaw*, *pitch* et *roll* (voir figure 4.1). Ces 3 paramètres sont directement envoyés au logiciel MAX/MSP pour actualiser l'orientation de l'utilisateur dans l'environnement virtuel.

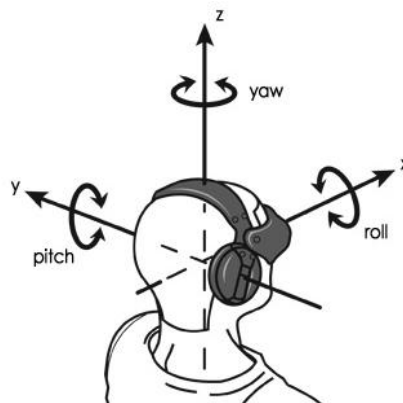


FIG. 4.1 – Le capteur de position de la tête fournit l'orientation de la tête dans les 3 axes représentés : *yaw*, *pitch* et *roll*

- Une manette sans fil, qui transmet au logiciel l'information sur le déplacement de l'utilisateur dans l'environnement virtuel. Les trois types de commandes possibles sont montrés dans le tableau 4.2.

Effet	Touche
Avancer	Touche haute de la croix de direction
Reculer	Touche bas de la croix de direction
Revenir au point de départ	L1

TAB. 4.1 – Description des types de commandes possibles à choisir avec la manette

Un exemple de la configuration réel est montré dans la figure 4.2.



FIG. 4.2 – Photographie du montage réel

Pendant la réalisation de l'expérience l'utilisateur est assis sur une chaise tournante. Pour tourner il lui suffit de tourner la chaise de 90° dans la direction souhaitée. Tous les scénarios dans lesquels l'utilisateur naviguera ne sont constitués que de lignes droites et de virages à angle droit. Les seules directions possibles seront donc les directions correspondantes aux 2 axes cartésiens perpendiculaires.

La fonction principale qui gère le fonctionnement du programme est la fonction `zoo.js`, implémentée en langage JavaScript. Cette fonction reçoit deux types d'**entrées** :

- Informations de l'extérieur relatives à :
 - L'orientation de la tête (information provenant du capteur).
 - La touche de la manette appuyée par l'utilisateur (avance/recul).

- Paramètres de réglage, à changer directement dans le patch MAX. Ces sont :
 - Position initiale de l'utilisateur dans le zoo virtuel.
 - Rayon d'écoute des animaux.
 - Rayon d'écoute du point de repère (balise central).
 - Vitesse de déplacement.

Comme **sortie**, elle transmettra deux types de sons directement au casque audio de l'utilisateur :

- Sons spatialisé, correspondant aux animaux situés à l'intérieur de chaque cage et au point de repère (balise central), s'il est présent.
- Son non-spatialisé, correspondant au son des pas, son des murs et son indiquant une voie libre à gauche ou à droite.

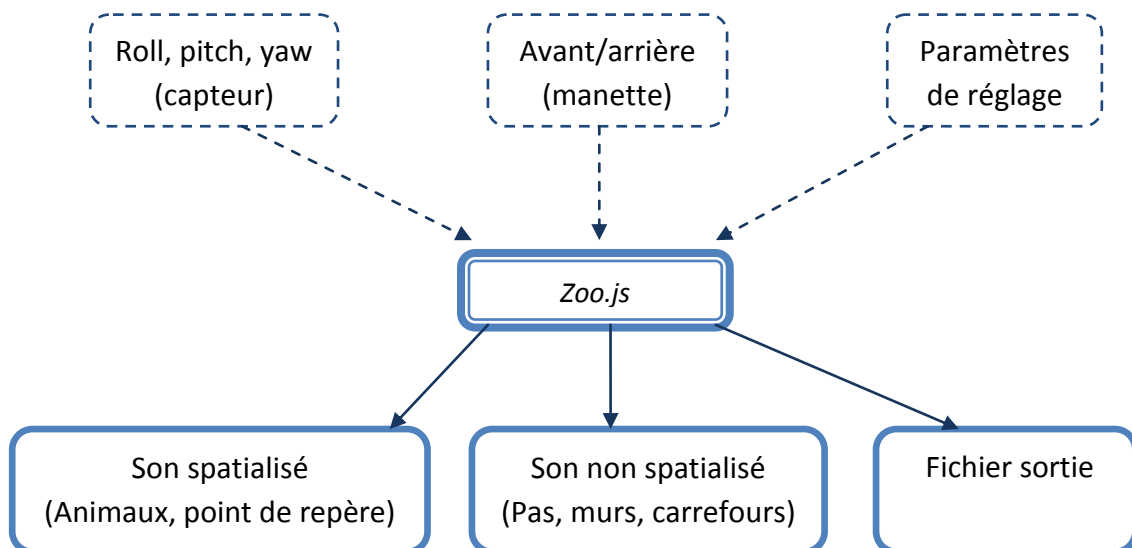


FIG. 4.3 – Schéma général du programme développé en MAX/MSP pour la réalisation de l'expérience. En ligne discontinue, les entrées de la fonction principale *Zoo.js*.

En ligne continue, les sorties.

Génération du fichier sortie

Le fichier sortie correspond à un fichier .txt contenant la position de l'utilisateur dans le zoo virtuel tous les 100 ms. La position est exprimée en deux dimensions (y, x), étant le centre de l'environnement virtuel la position (0, 0) (voir figure 4.4). Un exemple de fichier de sortie est montré dans la figure 4.5.

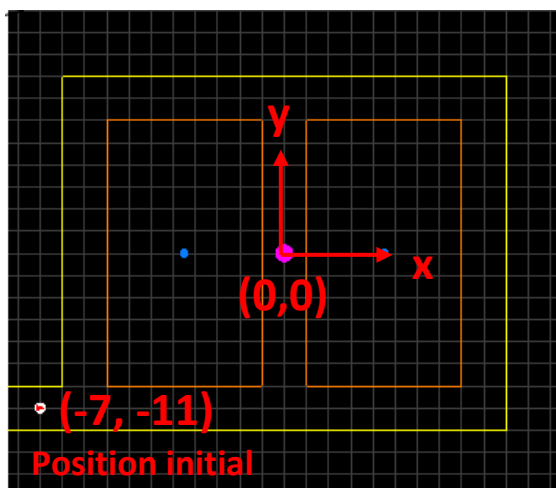


FIG. 4.4 – Scenario de l’environnement virtuel au avec les coordonnées du point de départ et du point central (en rouge)

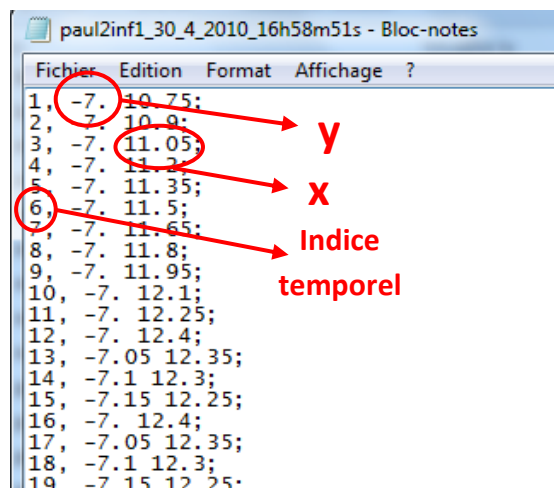


FIG. 4.5 – Fragment de fichier de sortie généré cours de la réalisation de l’expérience

Génération des sons spatialisés

Pour générer les sons spatialisés, nous avons utilisé le moteur de spatialisation binaurale développé au LIMSI dans l’équipe Son et Espace de Brian Katz.

Le processus implémenté en MAX est divisé en deux étapes fondamentales :

1.- *Changement de repère « corps-tête »* : Les paramètres *roll*, *pitch* et *yaw* transmis au logiciel pour le capteur, de même que les positions des animaux et du point de repère, sont exprimés par rapport au corps de l'utilisateur. Avec la fonction *Changement de repère* ses paramètres seront exprimés par rapport à la tête de l'utilisateur, de façon à pouvoir implémenter le processus de filtrage binaural. La sortie de cette fonction sera donc constituée des paramètres *roll*, *pitch* et *yaw* et la position des cibles par rapport à la tête.

2.- *Spatialisation du son*: divisée en 2 étapes :

- Choix de l'HRTF* : On sélectionne la fonction HRTF optimale depuis la base de données de filtres acoustiques à disposition, selon la position de la cible sonore à spatialiser.
- Filtrage binaural* : Convolution du son à spatialiser avec la réponse impulsionnelle correspondant à la position de l'espace du son à spatialiser.

Génération des sons non spatialisés

Ce sont les sons correspondant à :

- Son des pas : Il s'agit d'un son stéréophonique (pas avec pied gauche : plus intense dans l'oreille gauche, pas avec pied droit : plus intense dans l'oreille droite).
- Son des murs (ou cages) : Il s'agit d'un son monophonique.
- Son carrefour : Il s'agit d'un son stéréophonique (voie libre à gauche : son dans la oreille gauche, voie libre à droite : son dans la oreille droite).

Un schéma général de fonctionnement du programme développé en MAX/MSP pour la réalisation de l'expérience est montré dans la figure 4.6.

4.2.4 Choix des sons de l'environnement

L'objectif de cette partie est de décrire le procédé et les problèmes rencontrés pendant la phase de sonification de l'environnement.

Premièrement le but était de mettre en place un environnement strictement auditif et sans aucune information verbale (pour cause de charge cognitive liée au traitement du langage, Klatzky et al. 2006), mais contenant des informations de distance et de direction égocentrées, notamment, la position des points de navigation (animaux, carrefours, murs) ou des points de repère en dehors de la route (balise centrale).

Ensuite nous allons détailler les types de sons choisis et leur finalité.

- *Son des pas* : Après la réalisation de plusieurs tests nous avons ajouté cette information auditive pour augmenter la notion de déplacement et de distance parcourue. Le son choisi correspond à des vrais sons de pas.
- *Son des carrefours* : Son objectif est de signaler à l'utilisateur qu'il est possible d'emprunter la voie correspondant à l'oreille dans laquelle il a entendu le son. Il apparaît juste au moment de tourner. Dans le cas des carrefours où il est possible de tourner des deux sens, le son apparaît avec un décalage d'un demi seconde entre l'oreille gauche et l'oreille droite. Il s'agit d'un son électronique, pareil pour les deux oreilles. Plusieurs pré-tests ont été réalisés pour le choix de ce type de son (voir section 4.3.1).

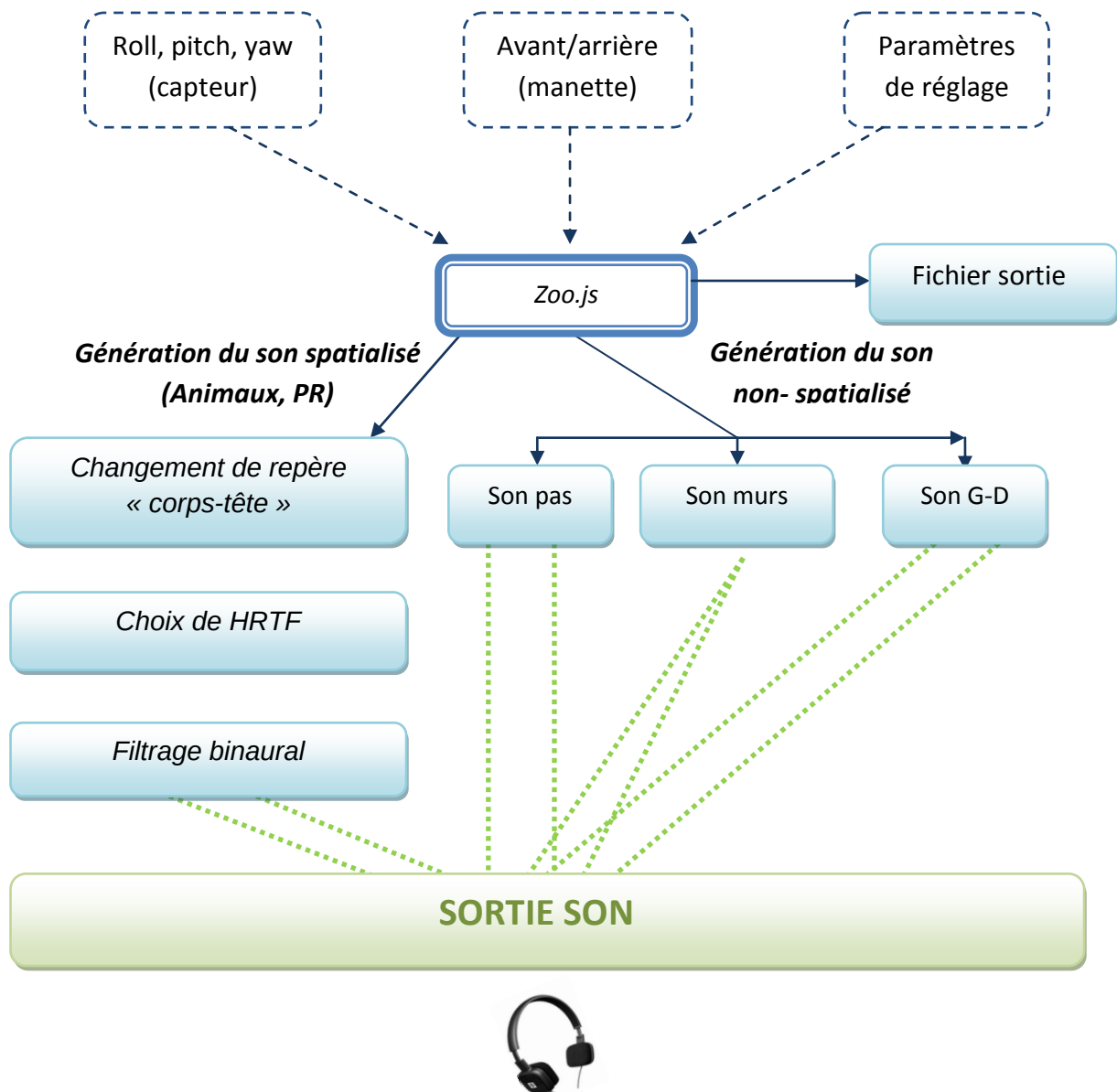


FIG. 4.6 - Schéma général de fonctionnement du programme développé en MAX/MSP pour la réalisation de l'expérience. En ligne discontinue, les entrées de la fonction principale *Zoo.js*. En ligne continue, les sorties. Les sorties correspondent à la génération des différents types de son (spatialisés et non spatialisés), qui vont direct à la carte son de l'ordinateur et puis au casque de l'utilisateur

- *Son des murs* : Son objectif est de signaler à l'utilisateur qu'il ne peut plus avancer. Il est répété le temps que l'individu regard vers le mur et la touche haut (ou bas, si le mur est derrière) de la manette est appuyée. Il s'agit d'un son électronique.
- *Sons des animaux* : Ils font partie de la tâche à réaliser. La spatialisation permet à l'utilisateur de localiser la position des animaux par rapport à son corps quand il passe à côté. Il s'agit des sons enregistrés, bien discernables et reconnaissables.
- *Son de la balise centrale* : Il fait partie de la tâche à réaliser. L'actualisation du son en termes de fréquence, intensité et spatialisation permet d'augmenter la sensation de navigation dans l'environnement et donne des informations de distance par rapport au centre de l'environnement. Puisque le son de la balise est entendu périodiquement dans tout l'environnement et sa position doit être repérée facilement, il a été choisi un son bien défini, avec un timbre et une fréquence le moins gênant possible. Il s'agit d'un son enregistré.

4.3 Procédure pré-expérimentale

Avant de tester l'effet de la présence de l'indice sonore distal sur la capacité des sujets à adopter des stratégies allocentriques, il était nécessaire d'effectuer des pré-tests.

4.3.1 Pré-test 1

L'objectif de ce premier pré-test est de choisir le type de son indiquant qu'il est possible de tourner à gauche ou à droite. Les deux options proposées étaient les suivants :

- *Option son 1* : Pendant que l'utilisateur navigue dans l'environnement il entend toujours le son spatialisé simulant les 2 murs du couloir dans lequel il passe. Au moment où l'un des deux sons disparaît d'une des oreilles, cela indique que la voie est libre et le navigateur emprunter ce chemin.

- *Option son 2* : Quand l'utilisateur arrive à une intersection, un son apparaîtra dans une de ses oreilles pour lui indiquer quelles directions il est possible d'emprunter.

Participants

Huit personnes voyantes (et sans problème connu d'audition) de sexe masculin, avec des âges compris entre 23 et 29 ans ont participé à cette expérience.

Matériel

Chaise rotative, repères tactiles disposés au sol autour de la chaise, ordinateur contrôlant l'environnement virtuel auditif, casque audio, capteur de position de la tête placée sur le casque, manette sans fil, bandeaux pour cacher les yeux des participants.

Procédure

La tâche va consister à faire un parcours défini dans deux scénarios virtuels différents, l'un très simple (scenario A), l'autre, plus complexe (scenario B). (Voir figure 4.7).

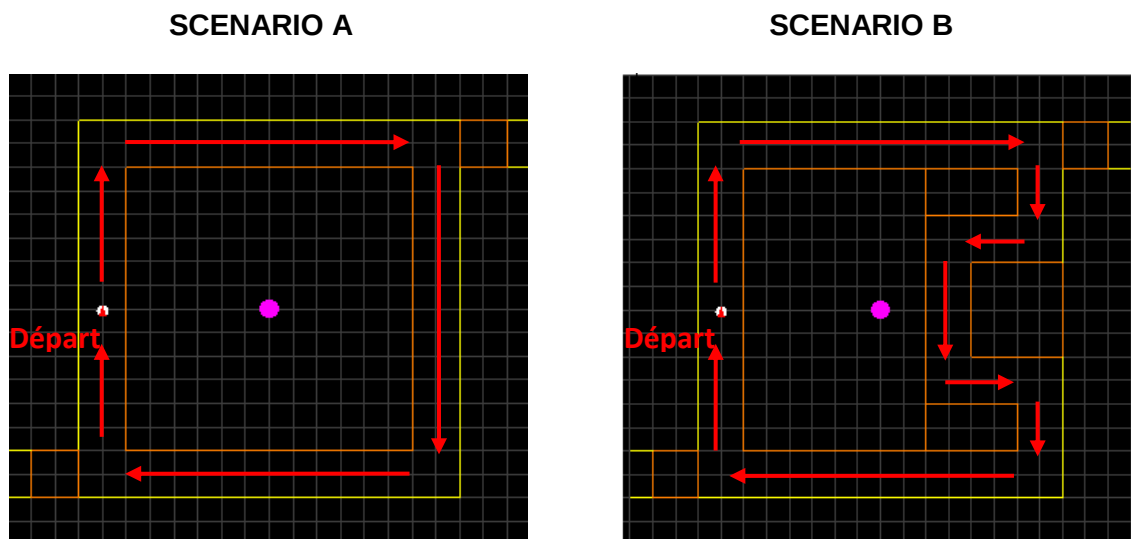


FIG. 4.7 – Scénarios conçus pour la réalisation du premier pré-test.

Chaque sujet doit réaliser 3 tours de parcours dans les deux scénarios. Tous les sujets seront confrontés aux deux types de sons (ordre de passation contrebalancé). La consigne sont les suivantes « faire le tour du parcours trois fois le plus vite possible » Une fois la tâche commencée, l'utilisateur ne recevra aucune information additionnelle. Nous mesurons le temps de réalisation de chaque tour et le temps de réalisation totale du parcours.

Après la navigation dans l'environnement virtuel, l'utilisateur répond à trois questions concernant ses préférences par rapport au deux types de son : Quel type de guidage sonore vous paraît plus simple ? Et plus efficace ? Lequel vous préférez ?

Résultats

Une ANOVA (2x2) avec le type d'environnement (simple vs. complexe) et le type de son (son1 vs. son 2) sur les 8 participants montre un effet significatif du type de labyrinthe : ($F(1, 7)=17,654, p<.005$). En revanche l'effet principal de son ($F(1, 7)=,50755$) n'est pas significatif. L'interaction entre son et labyrinthe est significative: $F(1, 7)=5,9152, p<.05$.

Un test post-hoc de Duncan visant à décomposer cette interaction montre que pour le scénario B les résultats sont significativement différents en fonction du type de son ($p>.02$). Les participants sont plus rapides avec l'option son 2.

Les résultats du questionnaire post-expérimental sont en accord avec ceux trouvés pour le temps de réalisation, c'est-à-dire que les sujets ont une préférence pour l'option son 2 (6/8 participants ont choisi l'option son 2 comme la plus simple, la plus efficace et la préféré).

L'expérience sera donc réalisée avec un son qui apparaîtra dans les oreilles de l'utilisateur, la gauche si le chemin libre est à gauche et inversement.

4.3.2 Pré-test 2

L'objectif de ce deuxième pré-test est de s'assurer que les participants sont cognitivement capables de traiter l'information sonore qui leur est fournie par la balise. Il s'agit ici de vérifier si les individus sont capables de déterminer la position d'un chemin par rapport à la balise.

Participants

Cinq personnes voyantes (quatre hommes et une femme) avec des âges compris entre 22 et 29 ans ont participé à cette expérience.

Matériel

Chaise rotative, repères tactiles disposés au sol autour de la chaise, ordinateur contrôlant l'environnement virtuel auditif, casque audio, bandeaux pour cacher les yeux des participants.

Procédure

La tâche va consister à parcourir un segment droit de l'environnement virtuel avec un chemin dont la position est variable.

Les indices sonores que l'utilisateur aura pour réaliser la tâche sont : les sons de ses pas, le son de la balise et le son de carrefour indiquant que le chemin est à sa gauche.

Nous testons 7 différences de distance entre la position du chemin et de la balise, différences comprises entre +3 et -3 (voir exemples figure 4.8).

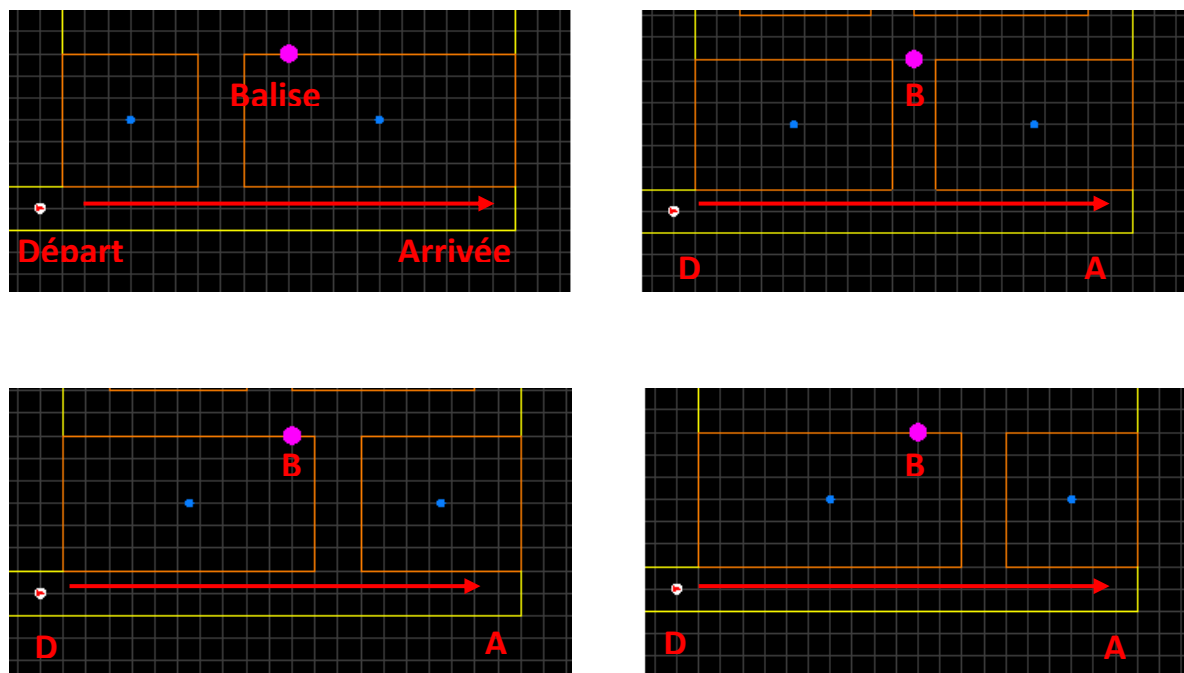


FIG. 4.8 – Exemples de configurations conçus pour la réalisation du deuxième pré-test. De gauche à droite et de haut en bas les distances entre le chemin et la balise sont : -3, 0, 1, 3.

Chaque utilisateur doit réaliser 5 fois le parcours pour chacune des 7 configurations, en total 35 parcours. L'ordre de présentation des configurations est aléatoire. La tâche du navigateur consiste à repérer la position du chemin situé à sa gauche par rapport au

centre de l'environnement (la balise). Ainsi, à la fin de chaque parcours il devra indiquer si, d'après ce qu'il a perçu, le chemin est situé avant ou après la balise.

Résultats

La courbe montrée dans la figure 4.9 représente la proportion de réponses correctes en fonction de la distance entre la balise et le chemin.

Comme prévu, le nombre d'erreurs de jugement diminue quand la distance augmente. La distance 0 n'est pas représentée dans le graphique puisqu'il n'y a pas de « bonne réponse » à la question posée pour cette distance. L'observation du graphique permet de se rendre compte que la courbe n'est pas centrée sur zéro. En effet, le plus grand nombre de mauvaises réponses est obtenu pour la distance 1. Ceci est dû au moment d'apparition du son matérialisant le chemin dans les oreilles du navigateur. En effet, les chemins font deux carrés de largeur, et le son du carrefour est envoyé à l'utilisateur quand il passe dans

le premier carré, pour lui permettre d'anticiper le virage. Pour la distance +1, l'utilisateur perçoit donc le son comme étant presque confondu avec le milieu de l'environnement. Pour la même raison, quand le chemin est situé juste au milieu (distance 0), les utilisateurs montrent une tendance à le placer avant.

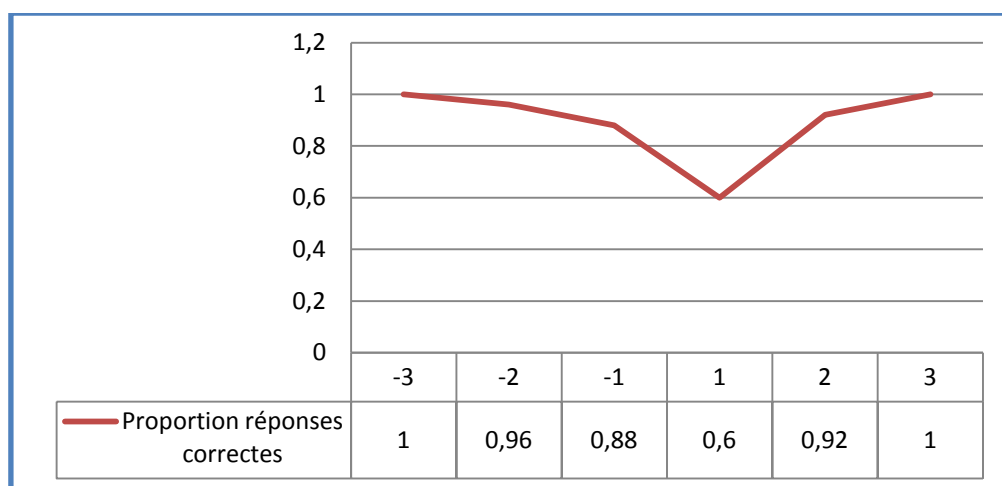


FIG. 4.9 - Résultats du pré-test 2. La courbe montre la proportion de réponses correctes en fonction de la distance entre la balise et le chemin.

Quoi qu'il en soit, ce pré-test permet de montrer que les participants sont bien capables d'utiliser les indices sonores fournis par la balise pour déterminer la position du chemin

avec une précision certaine. L'hypothèse étant vérifiée, nous pensons que les participants seront à même de prendre en compte les informations sonores fournies par un point distal de l'environnement, et pourront ainsi s'en aider dans la réalisation de l'expérience proposée.

4.4 Expérience

4.4.1 Participants

Dix personnes voyantes (huit hommes et deux femmes) avec des âges compris entre 22 et 29 ans ont participé à cette expérience. La moitié des sujets ont réalisé l'expérience avec la balise sonore présente au milieu de l'environnement. Les deux groupes de sujets ont été sélectionnés de manière à équilibrer la proportion hommes/femmes dans chaque groupe, ainsi que dans le souci d'obtenir des groupes équilibrés en termes de capacités d'imagerie et de représentation visuo-spatiale (voir annexe 1).

4.4.2 Matériaux

Le matériel utilisé est le même que pour le premier pré-test. Nous utilisons en plus un test de capacité visuo-spatiales et un test d'imagerie.

4.4.3 Procédure

L'utilisateur a les yeux bandés dès son entrée dans la salle et les gardera ainsi tout au long de l'expérience, de manière à ce que tous les sujets, voyants et non-voyants, soient dans les mêmes conditions pour la réalisation de la tâche.

La tâche consistant à effectuer des parcours définis, puis à effectuer des inférences spatiales dans l'environnement est précédée d'une phase de familiarisation en 3 étapes.

Première familiarisation

Avant d'expérimenter le zoo virtuel, l'utilisateur va évoluer dans un environnement très simple, composé de deux grandes cages (voir figure 4.10), afin de se familiariser avec la méthode de déplacement et de découvrir les sons qui composent l'environnement. Il

pourra explorer manuellement l'environnement simple dans lequel il va évoluer grâce aux briques de lego mises à sa disposition.

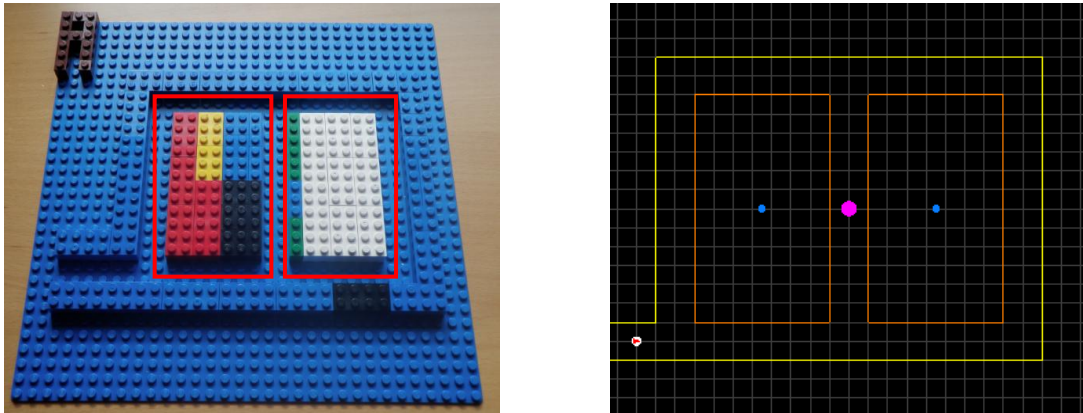


FIG. 4.10 – Environnement conçu en lego et en MAX/MSP pour la première familiarisation de l'expérience.

Une fois que l'utilisateur aura touché l'environnement et marché à l'intérieur pour entendre tous les types de sons possibles (son de pas, son des carrefours, sons des murs), sa tâche va consister à naviguer à sa guise dans l'environnement. Après le troisième virage l'expérimentateur stoppera le déplacement libre et il devra répondre les questions : « où pensez vous vous situer ? » et « dans quelle direction pointez vous ? ».

Deuxième familiarisation

La deuxième familiarisation consiste à faire rentrer l'utilisateur dans un univers un peu plus complexe, composé de quatre cages, qu'il pourra aussi toucher grâce aux lego (voir figure 4.11).

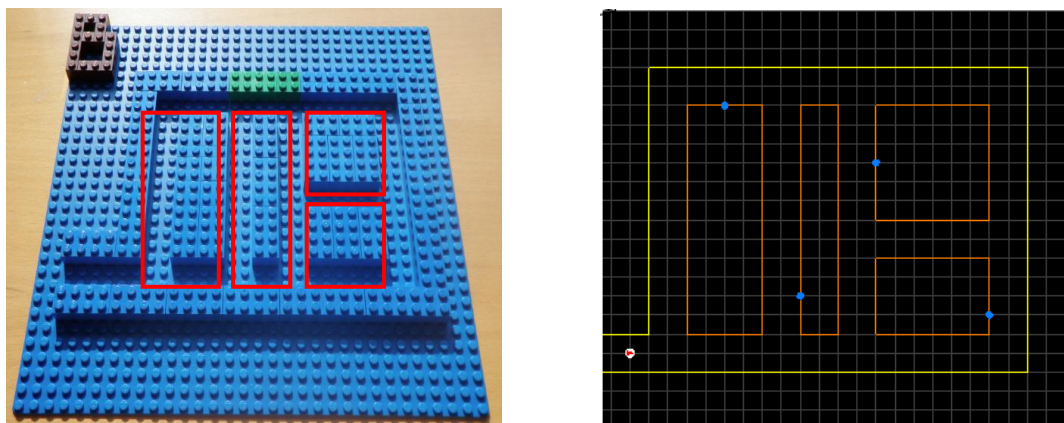


FIG. 4.11 – Environnement conçu en lego et en MAX/MSP pour la deuxième familiarisation de l'expérience.

Une fois que l'utilisateur aura touché l'environnement sa tâche sera de naviguer à l'intérieur et de chercher les 4 animaux qui s'y trouvent. Il devra ensuite indiquer sur le lego à quel emplacement se situent les animaux.

Troisième familiarisation

L'objectif de cette familiarisation est de permettre à l'utilisateur de s'entraîner à reconstruire un espace qu'il n'a jamais touché (voir figure 4.12). Dans cet espace, il va devoir faire le tour des cages par l'allée extérieure deux fois, puis, il devra utiliser les briques lego mises à sa disposition pour le reconstruire. Il a pour consigne de se servir du premier tour pour compter le nombre de chemins et du deuxième tour pour essayer d'identifier la position précise des chemins.

Ensuite, il devra reconstruire l'espace grâce aux lego grâce aux lego mis à son disposition (2 types de briques : 1x6 et 1x4).

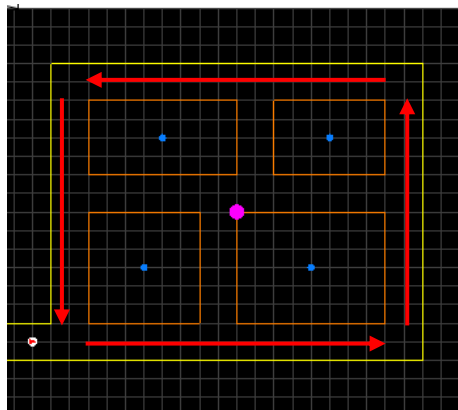


FIG. 4.12 – Environnement conçu en MAX/MSP pour la troisième familiarisation de l'expérience.

➤ Groupe avec balise sonore

Les sujets du groupe « balise sonore » bénéficient de la présence de l'indice distal sonore situé exactement au milieu de l'environnement pour les aider dans la réalisation de la tâche.

La balise donne trois indices à l'utilisateur pour lui permettre de s'orienter par rapport à elle : elle s'actualise en fonction de la position et de l'orientation de sa tête (son 3D), le son s'intensifie et le rythme s'accélère au fur et à mesure que l'utilisateur s'approche de la balise.

Ainsi, l'utilisateur pourra se servir d'elle pour savoir s'il a, ou non, déjà dépassé le milieu, et en déduire la position des routes (voir section 4.3.2).

4.4. 4 Description de la tâche

Une fois que l'utilisateur a appris à se déplacer, à identifier les cages et la position des animaux et à reconstruire un environnement en lego, il sera emmené dans un zoo virtuel un peu plus complexe, contenant 6 cages et 6 animaux (un seul animal par cage).

La tâche de l'utilisateur va consister dans un premier temps à se laisser guider par l'expérimentateur du point de départ jusqu'à un premier animal, puis jusqu'à un deuxième. En fin, au signal de l'expérimentateur, il devra se rendre le plus vite possible de l'animal 2 à l'animal 1 (exécution de l'inférence).

La tâche sera répétée 3 fois dans chaque scenario (jusqu'à trouver les 6 animaux), et il y aura 3 scenarios (zoos) différents (voir figure 4.13). Au total la tâche consistera donc en la réalisation de neuf inférences.

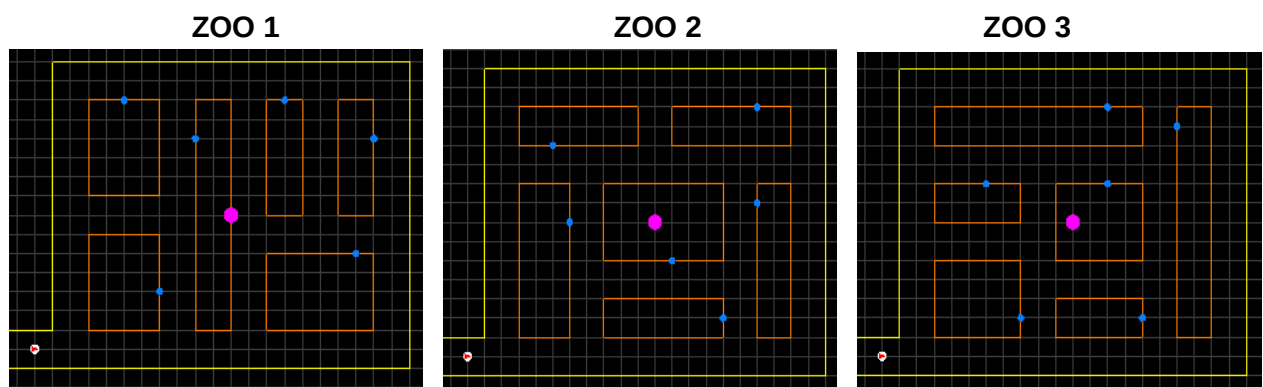


FIG. 4.13 – Les trois zoos virtuels conçus en MAX/MSP pour la réalisation de l'expérience.

4.4. 5 Mesures effectuées:

Les données relevées au cours de la réalisation de l'expérience sont les suivantes :

- Test d'imagerie et de capacités visuo-spatiales: divisé en trois parties. Un test de représentation mentale (test du cube), neuf items d'auto évaluation des habitudes d'imagerie mentale et deux questions sur les stratégies utilisées par l'individu lorsqu'il doit retenir ou expliquer un chemin (voir annexe 1).

Question 1 : Une personne vous donne des instructions pour trouver votre chemin, préférez-vous vous former une image du chemin en question ou vous remémorer les instructions verbales ?).

Question 2 : Un ami vous demande de lui expliquer un chemin que vous connaissez. Vous imaginez vous vous-même en train de marcher le long du chemin ou essayez vous plutôt de créer une carte mentale représentant le chemin ?

- Fichier de sortie généré par le logiciel : Il enregistre la position des participants dans le zoo chaque 100 ms.
- Réussite de l'inférence: On considère que l'inférence est réussie si le participant arrive du deuxième au premier animal. On considère que l'inférence n'est pas réussie quand le participant est perdu et se voit obligé d'appuyer sur la touche « reset » pour revenir au point de départ.
- Stratégie utilisée pour réaliser l'inférence : On considère que le participant a utilisé une stratégie egocentrique si, à partir du premier animal, il revient sur ses pas jusqu'au point de départ puis va jusqu'au deuxième animal. On considère que le participant a utilisé une stratégie allocentrique s'il emprunte un chemin alternatif jusqu'au deuxième animal. Nous codons 0 quand le participant utilise une stratégie égocentrique et 1 quand il utilise une stratégie allocentrique.
- Précision de la reconstruction des environnements en lego.

A partir de ces données nous avons calculés les valeurs des variables dépendantes nécessaires pour l'analyse statistique ultérieure. Ces calculs sont détaillés à la suite.

- *Note test cube* : Elle est calculé sur 5, avec des coefficients en fonction de la difficulté des questions.
- *Note items* : Elle correspond à la moyenne des neuf items sur une échelle de 1 à 5 avec 1= pas du tout d'accord, 3= neutre et 5 = tout à fait d'accord.
- *Question 1 et question 2* : on note 1 si la réponse implique une stratégie allocentrique, 0 pour une stratégie egocentrique et 0,5 si la stratégie impliquée est mixte.

- *Distance parcourue* : Elle est obtenue à partir du fichier de sortie généré par le logiciel. C'est la distance en nombre de carrés parcourus dans l'environnement, entre le moment où l'utilisateur arrête d'entendre le deuxième animal jusqu'au moment où il entend à nouveau le premier (en sachant que le rayon d'écoute des animaux est de deux carrés). L'erreur de distance parcourue correspond donc à la différence entre la distance réelle parcourue par le participant et la distance minimale théorique nécessaire pour réaliser l'inférence.
- *Temps de réalisation de l'inférence* : Il est obtenu à partir du fichier de sortie généré par le logiciel. Le temps est mesuré en secondes, depuis le moment où l'utilisateur arrête d'entendre le deuxième animal jusqu'au moment où il entend à nouveau le premier.
- *Pourcentage de reconstruction des zoos* : Il est obtenu à partir des photos des environnements lego réalisés par les participants. Elles sont comparées avec les environnements réels, en considérant que chaque « pion » lego correspond à un carré de l'environnement. Au total, les trois scenarios ont une surface de 320 carrés de surface (16x20). Chaque « pion » bien placé par l'utilisateur vaut un point. De cette note sur 320 point, on dérive un pourcentage de reconstruction (voir figure 4.14).

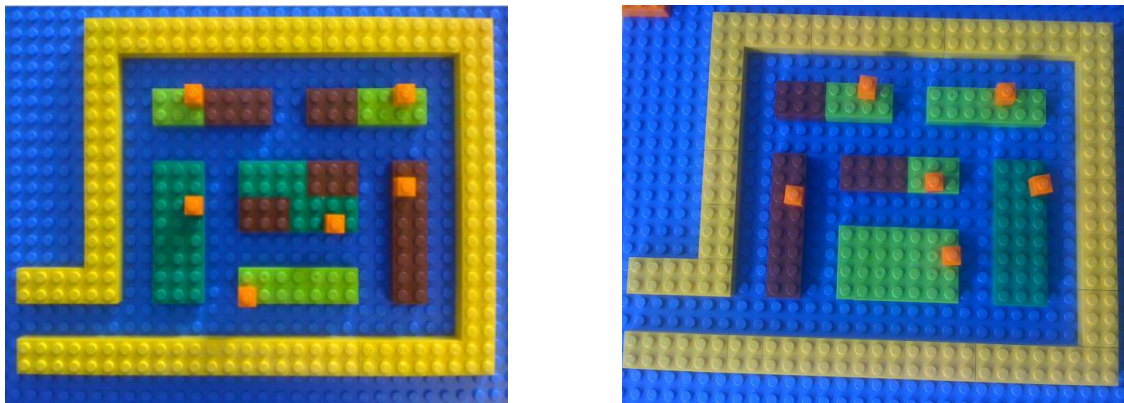


FIG. 4.14. Zoo 2 en Lego (à gauche) et exemple de reconstruction du zoo 2 (à droite). Les « pions » orange représentent la position des animaux.

4.4. 6 Hypothèse opérationnelle

Les inférences peuvent être réalisées de deux façons : en passant par le point de départ, ce qu'implique l'adoption d'une stratégie egocentrique, ou en utilisant les informations de l'espace acquises pendant la navigation pour retrouver un chemin alternatif, ce qui implique l'utilisation d'une stratégie allocentrique.

Selon notre première hypothèse, la présence de la balise devrait aider les participants à ne pas retourner sur leurs pas mais plutôt adopter une perspective allocentrique au moment de réaliser l'inférence. Le navigateur disposant de la balise aura en effet toujours une notion de sa position et de son éloignement par rapport au centre de l'environnement. Ainsi, quand il passera à côté d'un animal il pourra plus facilement retenir sa position par rapport au point central et l'aidera à « placer le point sur la carte ».

De plus, une stratégie allocentrique réussie (c'est-à-dire que le navigateur trouve le deuxième animal en passant par le chemin le plus court) impliquerait un temps de réalisation plus court et moins de distance parcourue qu'une stratégie égocentrique réussie. A partir de cette remarque et en conséquence de notre première hypothèse, nous proposons que la présence de la balise devrait donner lieu à des temps et à des distances parcourues moindres.

Enfin, la présence de la balise devrait aider les participants à se former une représentation de l'espace plus solide, et les participants en bénéficiant devraient donc présenter de meilleurs résultats en reconstruction de l'environnement.

4.4. 7 Résultats

Les différences de moyennes en fonction de la présence de la balise ont été analysées avec le t de *Student* (voir tableau 4.2). Le seul effet significatif est observé pour la variable Stratégie ($t_{88}=-2,6$; $p<.02$). En présence de la balise, les participants ont tendance à moins utiliser la stratégie egocentrique pour la réalisation de l'inférence.

Même si les moyennes diffèrent pour les autres variables, aucune de ces différences n'atteint le seuil de significativité. On observe seulement une tendance pour la variable temps ($t_{88}=1,86$, $p=.066$).

BALISE	Réussite	Stratégie	Temps	DP-DR
Sans	0,80	0,73	38,51	14,94
	(0,40)	(0,45)	(33,86)	(19,43)
Avec	0,91	0,93	25,95	8,88
	(0,29)	(0,25)	(25,20)	(15,37)

TAB. 4.2 - Les moyennes obtenues pour les différentes variables dépendantes, en fonction de la balise (entre parenthèse les écart-types)

Légende Table

Stratégie	Moyenne stratégie utilisé. Valeur entre 0 et 1. 1 = Allocentrique, 0 = Egocentrique.
Réussite	Moyenne réussite de l'inférence. Valeur entre 0 et 1. 1 = Réussite, 0 = Echec.
Temps	Moyenne en secondes des temps de réalisations des inférences.
DP-DR	Distance Parcourue – Distance Réelle. Moyenne en nombre de carrés parcourus pour la réalisation des inférences.
% Zoo	Pourcentage de reconstruction des zoos.
Cube	Moyenne note test cube.
Items	Moyenne note test items.
Question 1	Moyenne des réponses à la question 1.
Question 2	Moyenne des réponses à la question 2.

Le pourcentage de reconstruction des environnements en Lego est élevé dans les deux cas et ne diffère pas d'un groupe à l'autre. Si l'effet de la balise est négligeable dans ce cas, on peut cependant remarquer que les environnements sont bien reconstruits par les sujets (83,3% en moyenne).

Des calculs de coefficients de corrélation sont effectués dans le but d'étudier les liens entre les diverses variables mises en jeu, et plus précisément les liens entre capacités d'imagerie et visuo-spatiale des sujets et leur comportement dans l'environnement virtuel (voir tableau 4.3).

	Réussite	Stratégie	DP-DR	% Zoo	Cube	Items	Question1	Question2
Temps	-0,08	-0,14	0,87	0,32	0,28	-0,31	-0,72	-0,25
Réussite	1,00	-0,04	0,14	0,43	0,42	0,37	-0,13	-0,31
Stratégie		1,00	-0,45	-0,45	0,27	0,03	0,16	0,00
DP-DR			1,00	0,52	0,11	-0,21	-0,72	-0,13
% Zoo				1,00	0,49	-0,19	-0,36	-0,05
Cube					1,00	0,04	-0,27	-0,05
Items						1,00	0,51	0,42
Question1							1,00	0,48
Question2								1,00

TAB. 4.3 - Coefficients de Corrélations de Rangs de Spearman. Corrélations significatives marquées à $p < .05$ (en rouge).

En dehors de la corrélation significative observée entre le temps de réalisation de l'inférence et la distance parcourue observe une autre corrélation significative entre le temps et la réponse à la question 1. Cela signifie que les participants qui préfèrent utiliser, à priori, une stratégie allocentrique (former une image mentale du chemin en question) pour naviguer dans l'espace, sont aussi ceux qui ont mis le moins de temps à réaliser les inférences.

Malgré l'effet que la corrélation n'est pas significative, une forte tendance entre le pourcentage de reconstruction du zoo et la distance parcourue pour réaliser l'inférence est observée. Cela est logique aussi puisque les participants qui ont navigué plus dans l'environnement ont reçu plus des indices sonores provenant de l'environnement, ce qui facilite la tâche de reconstruction.

Un résultat important est l'absence de corrélation entre les tests d'imagerie et de capacité visuo-spatiale et la stratégie utilisée pour la réalisation des inférences. Le choix d'une stratégie de réalisation de l'inférence ne semble donc pas être lié aux capacités intrinsèques des sujets, ce qui laisse penser que l'utilisation d'une stratégie allocentrique est bien due à la présence de la balise.

4.4. 8 Discussion

L'objectif de cette expérience était de tester l'effet de la présence d'un objet distal sur la stratégie de déplacement adoptée (allocentrique vs. egocentrique), les temps et distances de parcours, ainsi que sur la qualité de la reconstruction d'un environnement virtuel sonore.

L'hypothèse proposée selon laquelle le groupe avec balise devrait présenter de meilleurs résultats en reconstruction de l'environnement n'est pas validée. Sans doute le traitement de la balise induit un coût cognitif supplémentaire, coût qui a une influence négative quand il s'agit de mémoriser et de reconstruire le parcours. Dans le cadre de notre expérience, nous nous situons sur une construction spatiale de l'environnement en mémoire à court terme (mémoriser et reconstruire dans un temps de 10 minutes maximum) alors que la construction spatiale solide d'un environnement doit à un moment ou à un autre faire l'objet d'un stockage à long terme. Il est probable que sur une durée de test plus longue (étude longitudinale, avec un environnement plus complexe), cet effet pourrait ressortir plus clairement. Toutefois les pourcentages de reconstruction des deux groupes sont bons et indiquent que les individus sont capables de se former une représentation spatiale uniquement basée sur des informations non-verbales.

Nos résultats sont cohérents avec ceux trouvés par Ohuchi, Iwaya, Suzuki et Munekata (2006) (voir section 4.1). Ils avaient trouvés que la majorité des participants étaient capables de créer une carte cognitive de l'environnement après la navigation en environnement virtuel purement sonore.

La deuxième hypothèse, selon laquelle la présence de la balise devrait donner lieu à des temps et à des distances parcourues moindres, n'est pas validée non plus. Cela est sans doute dû aux tentatives de stratégie allocentrique ratées des participants. En effet, les participants n'ayant pas physiquement expérimenté les relations spatiales entre les 2 animaux, ils sont plus susceptibles de se « perdre » dans le zoo en recherchant l'animal par une stratégie allocentrique, qu'en utilisant une stratégie plus longue mais plus sûre, consistant à repasser par le point de départ, en n'utilisant que des chemins physiquement explorés au préalable.

En ce qui concerne la stratégie utilisée pour la réalisation de l'inférence notre hypothèse est validée : en présence de la balise, le nombre de départs égocentriques diminue et le nombre de départs allocentriques augmentent. De plus, il n'y a pas de corrélation significative entre la stratégie et les résultats aux tests (*Cube + Items*). Cela montre donc que, en permettant aux participants de rattacher leurs parcours à un objet distal, la

balise les aide à adopter un point de vue plus allocentrique que les sujets qui ne l'ont pas.

Pour pouvoir confirmer ces résultats et pour mieux comprendre le phénomène étudié, les travaux suivants sont envisagés :

1. Réaliser l'expérience avec des participants non-voyants.

Une fois que l'effet de présence d'un point distal sur la stratégie d'actualisation de l'espace est validé pour les participants voyants, l'idée est de répliquer cet effet chez les non-voyants. On s'attend même à ce que l'effet soit plus important chez les sujets-non-voyants, étant donnée leur tendance naturelle à adopter des stratégies égocentrées, pour intégrer les résultats dans le dispositif d'aide à la navigation NAVIG.

2. Réaliser une étude plus écologique

Les conditions de réalisation de cette étude ont été en partie déterminées par des contraintes matérielles. Avec des moyens plus importants, Il serait intéressant de simuler l'environnement virtuel dans une grande chambre anéchoïque, dans laquelle les participants pourraient se déplacer à l'aide d'un casque et d'un capteur de mouvement sans fil. L'idée étant de fournir à l'utilisateur le plus haut degré d'immersion possible, on peut même pousser l'idée jusqu'à réaliser une expérimentation en milieu urbain réel et complexe.

3. Réaliser l'étude sur un plan longitudinal

Les enjeux de la construction mentale d'un environnement spatial tel qu'une ville se tiennent plus sur le long terme. Il serait donc intéressant de réaliser une expérience similaire avec des mesures de construction espacées dans le temps, afin de se rendre compte de l'effet à long terme de la présence des indices distaux sonores.

5. Conclusion

L'étude réalisée apporte des résultats satisfaisants à petite échelle. En effet, la présence d'un élément distal aide les participants à adopter un point de vue plus allocentrique dans la réalisation d'une tâche d'orientation spatiale. Il serait nécessaire de continuer l'étude en prenant en compte les considérations citées précédemment pour observer les résultats à plus grande échelle. Puisque les non-voyants sont habitués à utiliser des modalités sensorielles autres que la vision et, en plus, sa capacité auditive est supérieure en général à celle des voyants (Gougoux et Lepore, 2009), ils pourraient se servir mieux des informations sonores provenant d'un élément distal de l'environnement.

Si dans le cas d'une étude à plus grande échelle les résultats sont aussi satisfaisants, alors l'intégration de cette option (présence de points distaux de l'environnement) dans un système d'aide à la navigation pour des personnes non-voyantes pourrait être envisagée. Cela permettrait aux non-voyants d'attacher des sons à des endroits de l'environnement qu'ils choisissent comme point de repère (domicile, lieu de travail, mairie, etc...). Ils pourraient ainsi activer les indices distaux à la demande et déterminer leur position par rapport à ces indices grâce à la spatialisation des sons et aux mécanismes de localisation auditive (indices binauraux et HRTF) (Voir section 3.2). A mesure qu'ils actualisent leurs déplacements quotidiens dans la ville avec ces indices distaux, ils pourraient lier les différents trajets qu'ils connaissent déjà d'un point de vue égocentrique en une représentation allocentrique et ainsi solidifier leur représentation de leur environnement urbain.

6. Annexe 1 : Test d'imagerie et de capacités visuo-spatiales

Nom, Prénom:

Date :

Date de naissance :

Condition visuelle :

Main Dominante :

Précision(s) :

TEST CUBES

Représentez-vous mentalement un cube lui-même constitué de 27 petits cubes (3x3x3).

Imaginez que toutes les faces extérieures des 27 petits cubes sont coloriées en rouge, et que toutes leurs faces intérieures, leurs faces cachées, sont coloriées en blanc. Imaginez maintenant que ce grand cube soit éclaté et que les 27 petits cubes le composant soient accessibles à la vision.

Question	Réponse	Temps
Combien de cubes possèdent 0 faces rouges ?		
Combien de cubes possèdent 1 face rouges ?		
Combien de cubes possèdent 2 faces rouges ?		
Combien de cubes possèdent 3 faces rouges ?		
Combien de cubes possèdent 4 faces rouges ?		

ITEMS

(Cocher la case correspondante, 1= pas du tout d'accord, 5= tout à fait d'accord)

-Lorsqu'une personne me raconte des choses qui lui sont arrivées, il ne me vient pas habituellement à l'esprit l'image visuelle des évènements qu'elle me décrit

1 ☐ ☐ ☐ ☐ ☐ 5

-Lorsque je lis un roman, je forme en général une image visuelle des scènes et des lieux qui sont décrits.

1 ☐ ☐ ☐ ☐ ☐ 5

-J'utilise fréquemment des images visuelles pour trouver la solution d'un problème

1 ☐ ☐ ☐ ☐ ☐ 5

-En fermant les yeux, il m'est facile de me représenter visuellement une scène que j'ai vécue.

1 ☐ ☐ ☐ ☐ ☐ 5

-Lorsque quelqu'un me raconte quelque chose qui lui est arrivé, je me représente les événements par l'imagination de manière très vive.

1 ☐ ☐ ☐ ☐ ☐ 5

-J'ai des difficultés à me représenter l'image mentale visuelle du moindre objet.

1 ☐ ☐ ☐ ☐ ☐ 5

-Mes pensées sont souvent constituées d'images visuelles.

1 ☐ ☐ ☐ ☐ ☐ 5

-J'apprécie souvent de retrouver un souvenir grâce à des images mentales.

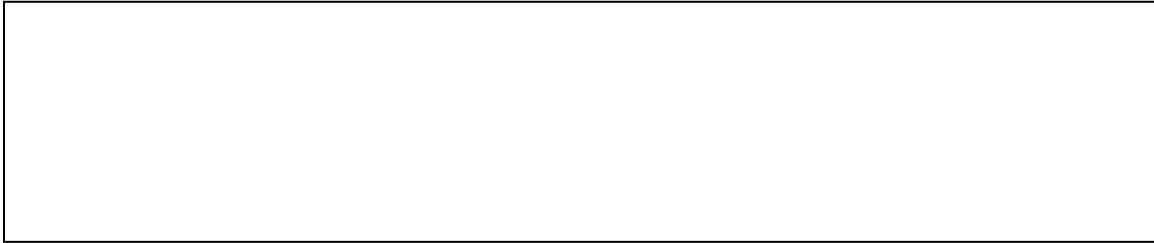
1 ☐ ☐ ☐ ☐ ☐ 5

-Juste au moment de m'endormir, je me remémore en image les événements que j'ai vécus dans la journée.

1 ☐ ☐ ☐ ☐ ☐ 5

QUESTIONS SUBSIDIAIRES

-Une personne vous donne des instructions pour trouver votre chemin, préférez-vous vous former une image du chemin en question ou vous remémorer les instructions verbales.

A large, empty rectangular box with a thin black border, intended for the user to write their answer to the first question.

-Un ami vous demande de lui expliquer un chemin que vous connaissez. Vous imaginez vous-même en train de marcher le long du chemin ou essayez vous plutôt de créer une carte mentale représentant le chemin ?

A large, empty rectangular box with a thin black border, intended for the user to write their answer to the second question.

Bibliographie

- AFONSO, A., KATZ, B., BLUM, A., JACQUEMIN, C. & DENIS, M. (2005). "A Study of Spatial Cognition in an Immersive Virtual Audio Environment: Comparing Blind and Blindfolded Individuals". Proceedings of ICAD 05.
- CASEY, S. M. (1978). "Cognitive mapping by the blind". *Visual Impairment and Blindness*.
- DODDS, A. G., HOWARTH, C.I., & CARTER, D.C (1982). "The mental maps of the blind: the role of previous visual experience". *Journal of Visual Impairment and Blindness*.
- DORÉ F. & MERCIER P. (1992). "Les fondements de l'apprentissage et de la cognition ", Lille, Presses Universitaires de Lille, Gaëtan Morin (éd.).
- FOULKE, E. (1982). "Perception, cognition and mobility of blind pedestrian". *Spatial abilities*.
- GÄRLING T., BOOK A. & LINDBERG E. (1984). "Cognitive Mapping of large scale environments: the interrelationships of action plans, acquisition and orientation", *Environment and Behavior*, Vol. 16, pp. 3-34.
- GÄRLING T. (1995). "How do urban residents acquire, mentally represent, and use knowledge of spatial layout?", in GÄRLING T. (Ed.), *Readings in environmental Psychology: urban cognition*, Academic Press, London, pp.1-12.
- GONOT, A. (2008). "Conception et évaluation d'interfaces de navigation dans les environnements sonores 3D". Thèse de doctorat.
- GOUGOUX, F. & LEPORE, F. (2009). "Voice perception in blind persons: A functional magnetic resonance imaging study". *Neuropsychologia*.
- HATWELL, Y. (2003). "Le développement perceptivo-moteur de l'enfant aveugle". *Presses Universitaires de France*.
- HOLLYFIELD, R.L. & FOULKE, E. (1983). "The spatial cognition of blind pedestrian". *Journal of Visual Impairment and Blindness*.
- KLATZKY, R., MARSTON, J., GIUDICE, N., GOLLEDGE, R. & LOOMIS, J. (2006). "Cognitive load of navigating without vision when guided by virtual sound versus spatial language". *Journal of Experimental Psychology: Applied*, 12, 223-232.
- LOOMIS, J.M. (1985). "Digital map and navigation system for the visually impaired". Unpublished manuscript, Department of Psychology, University of California.

- LOOMIS, J.M., KLATZKY, L., GOLLEDGE, R., CICINELLI, J., PELLEGRINO, J. , & FRY, F. (1993). "Nonvisual Navigation by Blind and Sighted: Assessment of Path Integration Ability". *Journal of Experimental Psychology*.
- LOOMIS, J.M., KLATZKY, L., PHILBECK, J.W. & GOLLEDGE, R.G. (1998). "Assessing auditory distance perception using perceptually directed action". *Perception & Psychophysics*, 60.
- LOOMIS, J.M., KLATZKY, L. & GOLLEDGE, R.G. (1999). "Auditory Distance Perception in Real, Virtual, and Mixed Environments". *Mixed Reality: Merging Real And Virtual Worlds*, Chap. 11.
- LOOMIS, J.M. LIPPA, Y., KLATZKY, R.L., & GOLLEDGE, R.G. (2002). "Spatial updating of locations specified by 3-D sound and spatial language". *Journal of Experimental Psychology : Learning, Memory, and Cognition*.
- MEIJER, P. (1992). "An experimental system for auditory image representations". *IEEE Transactions on Biomedical Engineering*, 39.
- MILLAR, S. (1994). "Understanding and representing space. Theory and evidence from studies with blind and sighted children". Oxford, UK: *Clarendon*.
- MILLER, L. (1992). "Diderot reconsidered: Visual impairment and auditory compensation". *Journal of Visual Impairment and Blindness*.
- NGUYEN, K., SUED, C., VIAUD-DELMON, I. & WARUSFEL, O. (2009). "Spatial audition in a static virtual environment : the role of auditory-visual interaction". *Journal of Virtual Reality and Broadcasting*, Volume 6, no. 5.
- NOORDZIJ, M. L., & POSTMA, A. (2005). "Categorical and metric distance information in mental representations derived from route and survey descriptions". *Psychological Research- Psychologische Forschung*.
- NOORDZIJ M.L., ZUIDHOEK S. & POSTMA A. (2007). "The influence of visual experience on visual and spatial Imagery". *Perception & Psychophysics*.
- O'KEEFE, J. & NADEL, L. (1976). "The Hippocampus as a Cognitive Map".
- OHUCHI, M., IWAYA, Y., SUZUKI, Y. & MUNEKATA, T. (2006). "Cognitive-map formation of blind persons in a virtual sound environment". Proceedings of the 12th International Conference on Auditory Display, London.
- PÉRUCH, P., CHABANNE, V., NESA, M. & THINUS-BLANC, C. (2004). "Mental scanning of images constructed from visual experience or verbal descriptions: the impact of survey versus route perspective". *Imagination, Cognition and Personality*.

- RIESER, J.J., LOCKMAN, J. & PICK, H.L. (1980). "The role of visual experience in knowledge of spatial layout". *Perception & Psychophysics*.
- RUSSELL J.A. & WARD L.M. (1982). "Environmental psychology". *Annual Review of psychology*, n° 33, pp. 651- 688.
- STARLINGER, I. & NIEMEYER, W. (1981). "Do the blind hear better ? Investigations on auditory processing in congenital or early acquired blindness. I. Peripheral functions. *Audiology*.
- THINUS-BLANC, C., & GAUNET, F. (1997). "Representation of space in blind persons: Vision as a spatial sense?". *Psychological Bulletin*.
- TOLMAN, E.C, (1948). "Cognitive maps in rats and men". *Psychological Review*, 55.
- TINTI C., TAMIETTO M. & CORNOLDI C. (2006), "Visual experience is not necessary for efficient survey spatial cognition: Evidence from blindness". *Psychology Press*.