



**HAL**  
open science

# Conception et évaluation de techniques d'interaction non visuelle optimisées pour de la transmission d'information

Damien Appert

► **To cite this version:**

Damien Appert. Conception et évaluation de techniques d'interaction non visuelle optimisées pour de la transmission d'information. Informatique et langage [cs.CL]. Université Paul Sabatier - Toulouse III, 2016. Français. NNT: 2016TOU30095 . tel-01394411v2

**HAL Id: tel-01394411**

**<https://theses.hal.science/tel-01394411v2>**

Submitted on 14 Jun 2017

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

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



Université  
de Toulouse

# THÈSE

En vue de l'obtention du

## DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par l'Université Toulouse III – Paul Sabatier

Discipline ou spécialité : *Informatique*

---

Présentée et soutenue le 27 mai 2016 par

**DAMIEN APPERT**

**Titre :** *Conception et évaluation de techniques d'interaction non visuelle optimisées pour de la transmission d'information*

---

### JURY

Dubois Emmanuel – Directeur  
Jouffrais Christophe – Co-Directeur  
Pissaloux Edwige – *Rapporteuse*  
Poirier Franck – *Rapporteur*  
Truillet Philippe – Invité

Professeur à l'Université Toulouse III  
Chargé de Recherche CNRS  
Professeur à l'Université Rouen  
Professeur à l'Université Bretagne-Sud  
Maître de Conférences à l'Université Toulouse III

---

**École Doctorale et spécialité :** Mathématiques, Informatique et Télécommunications de Toulouse (MITT) – Image, Information, Hypermédia.

**Unité de Recherche :** Institut de Recherche en Informatique de Toulouse (IRIT) – UMR 5505.

**Directeurs de Thèse :** Emmanuel DUBOIS et Christophe JOUFFRAIS.

**Co-Encadrant :** Philippe TRUILLET.







# Remerciements

Je souhaiterais remercier mon directeur de thèse, Emmanuel DUBOIS, mon co-directeur, Christophe JOUFFRAIS et mon co-encadrant, Philippe TRUILLET. Je souhaite remercier également l'équipe ELIPSE, dans laquelle je suis arrivé cinq ans plus tôt, en stage de recherche sous la tutelle de Philippe TRUILLET, puis reparti avant de revenir pour un an sous la supervision d'Emmanuel DUBOIS en tant qu'assistant de recherche, puis reparti pour revenir pour cette thèse avec Manu et Philippe, avec également Christophe, dans l'équipe pédagogique, encadrante et dirigeante. Cette thèse était dans le thème de la transmission d'information, tout comme mes échanges avec eux, l'objectif étant de concevoir et d'évaluer des techniques d'interaction optimisée pour y parvenir. Je remercie également l'Université Paul Sabatier pour le financement de ma thèse et mes rapporteurs, les Professeurs Edwige PISSALOUX et Franck POIRIER.

Je remercie donc l'équipe ELIPSE pour son accueil chaleureux depuis cinq ans, où j'ai pu échanger et discuter avec les membres actuels, anciens ou invités de cette équipe : Rémi, Philippe B., Emmanuel, Christophe J., Marc, Mustapha, Bernard, Mathieu, Antonio, Marcos, Nathalie, Philippe T., Frédéric, Nadine, Eliane, Sandra, Grégoire, Julie, Lilian, Stéphanie, Jing, Guillaume L., Yassine, Gary, Philippe R., Houssein, Damien S., Boris, Victor, Rita, Anke, Damien C., Slim, Samuel, Adrien, Olivier, Louis-Pierre, Adrien, Laurent, Christophe B., Guillaume G. (vu 1 fois : ) , Cédric, Bénédicte et Chantale (nom attribué à une Yucca, une plante du bureau 270, après un vote des membres du bureau). J'ai particulièrement apprécié l'ambiance de franche camaraderie avec les autres doctorants, mais aussi les « séminaires d'équipe » qu'ils soient en ELIPSCIENCE ou avec un verre à la main en d'autres lieux. Parmi tous les séminaires effectués ou suivis, il y en a un autre qui mérite d'être cité ici : le séminaire sur la théorie des jeux, où plusieurs doctorants se réunissaient autour d'une table et parlaient sérieusement du sujet...

Je remercie aussi l'ED MITT pour ces offres de formations quasi-illimitées, ce qui m'a permis de suivre près de deux cent heures de formation, lesquelles m'ont permises de connaître nombre de personnes de qualité, extérieures à l'équipe. Je remercie particulièrement pour leur écoute, leur réactivité et leur efficacité : Agnès, Martine, Elodie, Jean-Michel.

Je remercie les personnels de l'IRIT, Jean Pierre, Françoise, Floriane, Annie, Anne-Marie, Arnaude, Véronique, Charlotte, Clémentine et la direction, pour leur efficacité et leur sourire.

Je remercie également tous les doctorants et/ou docteurs IRIT qui m'ont fait confiance en tant que représentant au conseil du laboratoire, mais aussi en tant que membre de la commission des doctorants. Je remercie également les membres de cette commission avec qui j'ai partagé de bons moments : Camille, Faustine, Adrien, Jonathan, Teodora, Arnaud, Laurent, Jean-Philippe, Régis, Jérôme, Jacob, Violaine, Clément, Imen, Pierrick, Arnaud, Adrien, Julie, Martin, Rémi, Nicolas, Julien, Racim, François, Armelle, Vincent, Chiraz, François-Xavier, Bérengère.

Je remercie aussi, les autres membres organisateurs du séminaire DocToMe (séminaire mensuel par et pour les doctorants en informatique de Toulouse). L'ancienne équipe : Laure, Raja, Victor, Phillippe, Laurent, Alexandre et Louis-pierre et la nouvelle : Remi, El Malki, Willy, Bilel, Julien et Nguyen.

Je souhaite remercier Laurence CADIEUX et Bertrand MONTHUBERT.

Je remercie pour leur soutien indéfectible :

-> Mes amis proches : Ludovic, Clément, Manon, Arnaud, Fanny et Simon (et, Emelyne et Caroline), Céline, Noëline et Thomas, Cindy et Thomas (et Alice), Adeline et Harold (et Alexis), Charline et Vincent, David, Aude et Marie, Valou et Mika, Fabien, Aude et Damien, Stéphanie, Antoine et Gaëlle, Rémy.

-> Mes parents, Laurence et Jacques,

-> Mes beaux-parents, Marie-Jeanne et Bernard,

-> Ma femme, Camille ;-)

# Conception et évaluation de techniques d'interaction non visuelle optimisées pour de la transmission d'information

par Damien APPERT

## Résumé :

Dans des situations où la perception visuelle est fortement contrainte ou déficiente, il est nécessaire de rendre perceptible l'information dans une modalité non visuelle, tout en prenant en compte des capacités sensorielles et mnésiques humaines. Par exemple, un non-voyant, souhaitant prendre connaissance d'un itinéraire, devra le parcourir de façon non visuelle et le mémoriser. Cependant, outre l'aspect matériel, la mise en œuvre de solutions alternatives (non visuelles) demeure confrontée aux capacités cognitives de l'utilisateur (compréhension, mémorisation, intégration de plusieurs informations, etc.).

L'objet de cette thèse est de contribuer à la conception de techniques d'interaction permettant d'optimiser la transmission non visuelle d'informations. A ces fins, j'ai exploré l'apport de la multimodalité comme moyen d'optimisation permettant d'outrepasser les limites de la mémorisation. Je me suis concentré sur l'étude des techniques d'interaction basées sur les modalités auditives et tactiles, en limitant au maximum l'utilisation de la parole, afin de concevoir des techniques pour des environnements différents (flexibilité), d'optimiser l'utilisation de canaux perceptifs (exploitation des propriétés du son dans des messages audio pour transmettre plus d'informations, par exemple), d'éviter de limiter mes techniques par la barrière de la langue ou de sa compréhension et enfin, pour explorer d'autres solutions que la synthèse vocale seule. Les travaux de ma thèse ont mené à la conception, à l'implémentation et à l'évaluation de techniques d'interaction multimodale non visuelle, en réponse à différents contextes, dont, en particulier, ceux de la transmission d'informations de type <valeur>, <position> (couple de coordonnées) et <itinéraire> (séquence de couples direction-distance).

Pour parvenir à concevoir mes interactions, j'ai, tout d'abord, effectué une revue de la littérature, afin d'en extraire les principaux facteurs de conception de techniques d'interaction dédiées à la transmission non visuelle d'information. Puis, j'ai organisé ces facteurs sous la forme d'un cadre d'analyse, sur lequel je me suis appuyé pour concevoir chacune de mes techniques. Trois expériences distinctes ont permis d'évaluer l'influence de facteurs de conception sur l'efficacité des interactions et la satisfaction des utilisateurs vis-à-vis des techniques. Je peux, notamment, citer l'implication des utilisateurs (actif ou passif), la présence d'aides explicites, la transmission de plusieurs informations en parallèle et la modalité principale utilisée et, le type de codage dans lequel est encodée l'information.

Concrètement, en m'appuyant sur cette première contribution théorique, j'ai conçu et évalué dix-neuf techniques d'interaction non visuelle dans une tâche d'estimation de valeur entre 1

et 100. Les résultats de cette première étude ont démontré l'influence bénéfique de deux facteurs de conception («active/passive» ; «AvecRéférences/SansRéférence») sur la précision de l'estimation et la satisfaction des utilisateurs. J'ai aussi mis en évidence l'équivalence des modalités auditives et tactiles. Dans un deuxième temps, je me suis confronté au problème spécifique de la mémorisation d'itinéraires comportant plus d'informations que le nombre «magique» de  $7 \pm 2$  mnèmes (Miller, 56). Les résultats de cette deuxième étude ont montré les limites de l'utilisation de la parole, mais aussi les avantages liés à des interactions multimodales synergiques (parallèles et complémentaires). Enfin, j'ai conçu et évalué, dans une troisième étude, quatre techniques basées sur la modalité vibrotactile, facilitant la localisation d'une petite cible dans un grand espace. La meilleure de ces quatre techniques a été adaptée et testée dans un cas d'assistance à la recherche de cibles pour des personnes déficientes visuelles et a permis de diminuer la durée de la recherche visuelle d'un facteur trois. Toutes ces techniques pourraient assister des utilisateurs voyants ou non.

**Mots-clés :** interaction non-visuelle, multimodalité, estimation de valeurs, technique d'interaction active, technique d'interaction avec références, technique d'assistance.

# Design and evaluation of techniques non-visual interaction optimized for the transmission of information

by Damien APPERT

## Sum up:

In situations where the visual perception is strongly constraint or deficient, it is necessary to make perceptible the information with a “not visual form” while taking into account human sensory and mnesic capacities. For example, a blind person wishing to acquaint an itinerary must read it under a non visual form and memorize it. However, besides the material aspect, the implementation of alternatives (non-visual) still faces to the cognitive abilities of the user (comprehension, memorization, integration of various information, etc.).

The purpose of this thesis is to contribute to the design of interaction techniques allowing to optimize the transmission not visual of the information. For these purposes, I explored the feature of multimodality as a means of optimization, allowing of exceeding the memorization limits. I focused on the study of interaction techniques based on auditory and tactile modalities and by minimizing the use of the speech, in order to develop techniques for different environments (flexibility), optimize the use of perceptual channels (operating the properties of sound in audio messages to transmit more information, for example), avoid limiting my techniques by the language barrier or understanding and finally, to explore alternatives to the synthesised voice alone. The works of my thesis led to the design, to the implementation and to the evaluation of interaction techniques "non-visual" and "multiform", in answer to different contexts, whom in particular those of the information transmission of type <value>, <position> (pair of coordinates) and <itinerary> (sequence of couples direction-distance).

To achieve design my interactions, I have made a review of literature in order to extract the main factors of design of interaction techniques dedicated to the transmission not visual of the information. Then, I have organized these factors in an analytical framework on which I have relied to design each of my techniques. Three separate experiments were led to evaluate the influence of design factors on the effectiveness of interactions and satisfaction towards users of technology. I can give some of them, the involvement of users (active or passive), the presence of explicit help, the transmission of several information in parallel, the main modality used and the type of coding in which is encoded the information.

Concretely, lean on the first theoretical contribution, I have designed and evaluated nineteen interaction techniques non-visual in a value estimation task from 1 to 100. The results of this first study demonstrated the beneficial influence of two design factors ("active / passive" "WithReference / WithoutReference") on the accuracy of the estimate and user satisfaction. I also demonstrated the

equivalence of auditory and tactile modalities. Secondly, I am faced with the specific problem of memorizing itineraries with more information than the "magic" number  $7 \pm 2$  chunk (Miller, 56). The results of this second study have shown the limits of the use of the speech, but also the benefits of synergistic multimodal interactions (parallel and complementary). Finally, I have designed and evaluated in a third study four techniques based on vibrotactile modality, facilitating the localization of a small target in a large space. The best of these four techniques had been adapted and tested in a case of assistance in targets search for the visually impaired and has reduced the duration of visual search by factor of three. All these techniques could assist blind users or sighted users.

**Keywords:** non-visual interaction techniques, multimodality, value estimation, technique with references, technique with references, technique of assistance.

**SOMMAIRE :**

<b><u>1</u></b>	<b><u>Introduction</u></b>	<b>15</b>
1.1	PLAN DE LECTURE	17
<b><u>2</u></b>	<b><u>Etat de l'art sur la transmission d'information basée sur les modalités tactiles et/ou auditives</u></b>	<b>23</b>
2.1	MODALITES TACTILES	23
2.2	MODALITES AUDITIVES	46
2.3	MULTIMODALITE	66
2.4	CONCLUSION DE L'ETAT DE L'ART	74
<b><u>3</u></b>	<b><u>Utilisation/Définition d'un cadre analytique</u></b>	<b>79</b>
3.1	MOTIVATION	79
3.2	PRESENTATION DES FACTEURS DE CONCEPTION DE NOTRE CADRE ANALYTIQUE	80
3.3	ORGANISATION DES FACTEURS DE CONCEPTION DANS NOTRE CADRE D'ANALYSE	90
3.4	ILLUSTRATION DE GENERATION DE TECHNIQUES AVEC NOTRE CADRE D'ANALYSE	95
3.5	CONCLUSION SUR L'APPORT DU CADRE D'ANALYSE POUR CONCEVOIR NOS TECHNIQUES	111
<b><u>4</u></b>	<b><u>Première étude : techniques d'interaction non visuelle pour estimer une valeur</u></b>	<b>115</b>
4.1	MOTIVATION	115
4.2	EXPERIENCE 1.1 : TECHNIQUES D'INTERACTION NON VISUELLE POUR ESTIMER UNE VALEUR	116
4.3	EXPERIENCE 1.2 : ETUDE SPECIFIQUE DES FACTEURS DE CONCEPTION A/P ET AR/SR	137
4.4	CONCLUSION DE LA PREMIERE ETUDE	156
<b><u>5</u></b>	<b><u>Deuxième étude : interactions non visuelles pour faciliter le rendu d'une séquence de couples direction-distance</u></b>	<b>161</b>
5.1	INTRODUCTION A LA DEUXIEME ETUDE	161
5.2	RAPPEL SUR L'ETAT DE L'ART : TRANSMISSION DE DIRECTION ET DE DISTANCE	163
5.3	PRE-ETUDES ET CHOIX DE DESIGN	163
5.4	DESCRIPTION DES QUATRE TECHNIQUES RETENUES	165
5.5	HYPOTHESES DE TRAVAIL	172
5.6	PROTOCOLE EXPERIMENTAL DE L'EXPERIENCE 2.1	172
5.7	ANALYSE DES RESULTATS DES SUJETS NON-DV	174
5.8	ANALYSE DES RESULTATS PRELIMINAIRES CONCERNANT LES SUJETS DV	179
5.9	ANALYSE SPECIFIQUE : MEMORISATION DES SEGMENTS SELON LEUR POSITION DANS L'ITINERAIRE	181
5.10	DISCUSSION ET CONCLUSION DE LA DEUXIEME ETUDE	182
<b><u>6</u></b>	<b><u>Troisième étude : assistance tactile à la localisation de cibles périphériques pour des personnes à vision tubulaire</u></b>	<b>187</b>

<b>6.1</b>	<b>ÉTAT DE L'ART SUR LES TECHNOLOGIES D'ASSISTANCE A LA RECHERCHE VISUELLE</b>	<b>188</b>
<b>6.2</b>	<b>DISPOSITIFS TACTILES POUR LA PERCEPTION SPATIALE</b>	<b>191</b>
<b>6.3</b>	<b>EXPERIENCE 3.1 : CONCEPTION ET COMPARAISON D'INDICES TACTILES PERMETTANT DE LOCALISER UNE CIBLE PERIPHERIQUE</b>	<b>193</b>
<b>6.4</b>	<b>EXPERIENCE 3.2 : RECHERCHE VISUELLE D'UNE PETITE CIBLE DANS UN SALON VIRTUEL ENCOMBRE</b>	<b>208</b>
<b>6.5</b>	<b>EXPERIENCE 3.3 : ETUDE DE CAS SUR UN SUJET AVEC UNE VISION TUBULAIRE</b>	<b>215</b>
<b>6.6</b>	<b>DISCUSSION ET CONCLUSION DE LA TROISIEME ETUDE</b>	<b>216</b>
<b>7</b>	<b>Résumé des contributions, conclusion générale et perspectives</b>	<b>223</b>
<b>7.1</b>	<b>RESUME DES RESULTATS DE LA PREMIERE ETUDE</b>	<b>223</b>
<b>7.2</b>	<b>RESUME DES RESULTATS DE LA DEUXIEME ETUDE</b>	<b>228</b>
<b>7.3</b>	<b>RESUME DES RESULTATS DE LA TROISIEME ETUDE</b>	<b>230</b>
<b>7.4</b>	<b>CONCLUSION GENERALE</b>	<b>231</b>
<b>7.5</b>	<b>PERSPECTIVES</b>	<b>235</b>
<b>8</b>	<b>Bibliographie</b>	<b>241</b>
<b>9</b>	<b>Annexes</b>	<b>253</b>
<b>9.1</b>	<b>SOMMAIRE ÉTENDU</b>	<b>255</b>
<b>9.2</b>	<b>TABLE DES FIGURES</b>	<b>258</b>
<b>9.3</b>	<b>TABLE DES TABLEAUX</b>	<b>264</b>
<b>9.4</b>	<b>CAHIER DES SUJETS DE L'EXP. 1.2</b>	<b>265</b>
<b>9.5</b>	<b>ARTICLES ET POSTERS</b>	<b>279</b>
9.5.1	EYES-FREE INTERACTION TECHNIQUES FOR DISTANCE ESTIMATION	280
9.5.2	ASSISTANCE TACTILE A LA LOCALISATION DE CIBLES PERIPHERIQUES POUR DES PERSONNES A VISION TUBULAIRE	300
9.5.3	INTERACTIONS NON-VISUELLES POUR FACILITER LE RENDU D'UNE SEQUENCE D'INFORMATIONS	320
9.5.4	INFLUENCE DE LA PRESENTATION DES MOTS POUR DES LECTEURS DYSLEXIQUES	332
9.5.5	AN MDE-BASED FRAMEWORK TO SUPPORT THE DEVELOPMENT OF MIXED INTERACTIVE SYSTEMS	340

# **Chapitre 1**

## **Introduction**



## 1 Introduction

Dans un contexte de transmission d'informations, les interactions entre l'Homme et la machine doivent être conçues en prenant en compte les capacités sensorielles et mnésiques de l'Homme. Dans des situations où la perception visuelle est fortement contrainte ou déficiente, il est nécessaire de rendre l'information perceptible par une modalité sensorielle non visuelle. Par exemple, une personne non-voyante souhaitant prendre connaissance d'un itinéraire pourra parcourir une séquence d'instructions (aller tout droit pendant 100 m puis tourner à droite ; continuer pendant 150 m puis tourner à gauche, etc.). Elle devra mémoriser cet itinéraire sur la base de cette séquence d'instructions. Pour réaliser cette tâche, la synthèse vocale et la plage braille sont les solutions les plus communément utilisées. Cependant, la mémorisation d'une séquence d'instructions verbales atteint rapidement ses limites. De plus, l'utilisation de certains dispositifs, comme la plage braille ou la synthèse vocale, peut conduire à des contextes d'usage non adaptés. Par exemple, la plage braille est difficile à utiliser en mobilité. La synthèse vocale est en concurrence avec l'environnement sonore et risque de masquer des dangers potentiels. On peut même identifier des contextes plus larges, non liés à la déficience, comme la barrière de la langue pour les touristes par exemple.

L'objectif général de cette thèse est de contribuer à la conception de techniques d'interaction non visuelle permettant d'optimiser la transmission d'informations. A ces fins, nous avons construit un cadre d'analyse permettant de concevoir des techniques d'interaction selon plusieurs paramètres de conception, puis nous avons exploré l'apport de la multimodalité pour améliorer la transmission et la mémorisation des informations. Nous nous sommes concentrés sur l'étude des techniques d'interaction basées sur les modalités sensorielles auditives et tactiles. L'objectif était de concevoir des techniques pour des environnements différents, comme un environnement bruyant par exemple. Mais aussi, de concevoir des techniques optimisées afin de transmettre plus d'informations sur les canaux perceptifs, comme exploiter plusieurs propriétés du son dans des messages audio ou répartir des informations dans des messages audio et tactile par exemple. Enfin, nous modèrerons au maximum l'utilisation de la parole dans nos techniques, afin d'éviter de limiter les techniques par la barrière de la langue, notamment par sa compréhension et, plus largement, afin d'explorer d'autres solutions que la synthèse vocale seule (langage oral) ou la plage braille seule (langage écrit non visuel).

**PROBLEMATIQUES ET CONTRIBUTIONS**

Les travaux conduits dans cette thèse de doctorat ont mené à la conception, à l'implémentation et à l'évaluation de techniques d'interaction multimodale non visuelle. Ces techniques répondent à différents contextes nécessitant la transmission d'un ou plusieurs types d'informations comme par exemple une <valeur> et une <position> (couple de coordonnées) ou un <itinéraire> (séquence de couples direction-distance).

Pour concevoir les techniques d'interaction, nous avons tout d'abord effectué une revue de la littérature afin d'en extraire les principaux facteurs de conception de techniques d'interaction dédiées à la transmission d'information sous une forme non visuelle. Pour organiser cet immense espace de conception et présenter de manière structurée les résultats expérimentaux, nous avons construit un cadre analytique, c'est-à-dire une méthode d'analyse destinée à comprendre les techniques d'interaction à travers des facteurs de conception communs. En nous appuyant sur cette première contribution théorique, nous avons conçu et évalué un ensemble de techniques d'interaction non visuelle, afin d'évaluer l'influence de certains facteurs de conception sur l'efficacité des interactions et la satisfaction des utilisateurs vis-à-vis de ces techniques. Ces évaluations ont été menées dans trois contextes différents. Nous avons notamment évalué : 1) l'implication des utilisateurs (technique active ou passive) en nous inspirant des travaux de (Gibson 1962) ; 2) l'impact de la présence de repères sensoriels, censée faciliter l'estimation des valeurs d'après les travaux de (Fraisse 1984) ; 3) la transmission de plusieurs informations en parallèle en suivant les travaux de (Nigay and Coutaz 1996) qui distinguent la modalité principale utilisée (dispositifs) et, le type de codage (langage) pour transmettre l'information.

Dans une première étude, nous avons confronté une vingtaine de techniques, dans une tâche d'estimation de valeur entre 1 et 100. Dans une deuxième étude, nous avons évaluées certaines techniques, dans le cadre spécifique de la mémorisation d'itinéraires composés d'une séquence de cinq couples direction-distance. Enfin, nous avons conçu et évalué dans une troisième étude (collaborative), quatre techniques basées sur la modalité vibrotactile, facilitant la localisation d'une petite cible dans un grand espace. La meilleure de ces quatre techniques a été adaptée et testée dans un cas d'assistance à la recherche de cibles pour des personnes déficientes visuelles. Ce dispositif pourrait être facilement intégré dans des lunettes intelligentes et détecter des cibles d'intérêt, soit sur demande soit de façon automatique (par exemple les obstacles potentiels), facilitant ainsi la recherche visuelle et la

perception spatiale de l'environnement. Toutes ces techniques pourraient assister des utilisateurs voyants ou non dans des tâches d'estimation de valeurs, de positions ou d'itinéraires.

En synthèse, mes contributions pour l'ensemble de ces trois études, sont triples. La première contribution est la construction d'un cadre d'analyse pour concevoir des techniques d'interaction non visuelle et l'évaluation de certains facteurs de conception, extraits de la littérature et adaptés à notre contexte de transmission d'informations par estimation de valeurs, tels que l'implication des utilisateurs et la présence de repères sensoriels. La deuxième contribution est la conception, l'évaluation et la comparaison d'une vingtaine de techniques d'interaction pour une tâche d'estimation de valeur entre 1 et 100, de quatre techniques pour une tâche d'estimation d'un itinéraire et de quatre techniques pour une tâche d'estimation d'une position. Ces trois études montrent qu'il est possible d'estimer efficacement une information avec des modalités non visuelles et que l'usage exclusif de la modalité parole ne facilite pas la mémorisation d'un grand nombre d'informations. Enfin la troisième contribution est l'apport de notre cadre d'analyse, soit pour optimiser des techniques d'interaction déjà existantes, soit pour concevoir de nouvelles techniques, afin de transmettre efficacement et ce, de manière satisfaisante pour les utilisateurs, des informations.

## 1.1 Plan de lecture

Cette thèse de doctorat est composée de sept chapitres, le premier étant l'introduction elle-même. Dans le chapitre 2, nous dressons un état de l'art succinct concernant la psychologie de la perception tactile humaine, qui inclut la perception cutanée (vibration et pression) et la perception kinesthésique (position spatio-temporelle des articulations lors de déplacements des membres). Ces deux perceptions sont alors rassemblées dans l'appellation plus large de perception haptique. A la suite de cette présentation, nous illustrons l'utilisation de la modalité tactile dans des situations interactives liées à notre problématique de transmission d'informations. Dans une seconde partie de ce chapitre 2, nous observons le même plan concernant la psychologie de la perception auditive humaine et nous illustrons la mise en œuvre de cette modalité auditive dans des situations interactives toujours liées à notre problématique. Enfin, nous présentons la multimodalité et ses propriétés (CARE et CASE, (Nigay and Coutaz 1996)) dans le domaine de l'Interaction Homme-Machine (IHM). Nous finissons ce chapitre par une synthèse menant à la

proposition de créer un cadre d'analyse afin d'apporter une aide conceptuelle à la réflexion, pour une conception maîtrisée de techniques d'interaction non visuelle.

Le chapitre 3 présente la première contribution théorique de cette thèse. Nous y recensons les principaux facteurs de conception de techniques d'interaction issues de l'état de l'art, puis nous les organisons sous la forme de paramètres dans un cadre d'analyse. Enfin, nous illustrons leur utilisation au travers d'exemples mais aussi, par la génération de nos propres techniques d'interaction. Nous concluons ce chapitre en rappelant les avantages de ce cadre d'analyse et, par une transition vers les chapitres suivants : transmission d'une seule information, transmission d'un couple d'information, transmission d'une séquence de couples d'information.

Le chapitre 4 présente notre première étude expérimentale visant à comparer de nombreuses techniques non visuelles permettant d'estimer une seule valeur entre 1 et 100. Cette étude comparative est une contribution en soi, car la majorité des techniques d'interaction liées à notre problématique (et présentées dans l'état de l'art) sont difficiles à comparer entre elles en termes d'efficacité et de satisfaction. En effet, ces techniques ont été évaluées dans des conditions différentes : la plage de valeur à estimer peut varier entre 0 et 1 ou peut être un ensemble de quelque valeurs ([4, 6, 8 et 10]) ; le type d'information peut varier également, il peut être soit une distance soit une direction (par ex.). La deuxième contribution de ce chapitre 4 est les techniques d'interaction elles-mêmes, basées sur les modalités tactiles et auditives. Enfin, notre troisième contribution concerne l'évaluation de l'impact de deux facteurs de conception sur l'estimation de la valeur. Ces deux facteurs sont l'implication des utilisateurs (actif ou passif) et la présence de références sensorielles explicites permettant d'évaluer la valeur cible.

Le chapitre 5 présente notre deuxième étude expérimentale visant à comparer quatre techniques dans une tâche d'estimation et de mémorisation d'un itinéraire composé d'une séquence de couples direction-distance. La première contribution de cette étude est l'évaluation des limites de l'utilisation de la parole seule, comparativement à nos techniques multimodales optimisées où la parole est utilisée en complément d'une autre modalité (tactile ou auditive) ainsi que notre technique multimodale basée sur la modalité «sonification» (auditive sans parole). La seconde contribution est la démonstration des avantages liés à l'utilisation d'interactions multimodales synergiques (parallèles et complémentaires) et alternées (séquentielles et complémentaires).

Le chapitre 6 présente la troisième étude expérimentale qui est appliquée à un contexte de recherche de cible dans un environnement large et encombré. Dans cette étude, nous avons tout d'abord comparé quatre techniques basées sur des stimuli vibrotactiles dans une tâche de localisation d'une cible dans l'espace, puis, nous avons évalué la meilleure de ces quatre techniques dans une tâche d'assistance à la recherche visuelle d'une petite cible dans une scène visuelle encombrée. Les contributions de cette troisième étude permettent la comparaison de deux systèmes de coordonnées (cartésien ou polaire), mais aussi la comparaison de deux types de codage vibrotactile, l'un basée sur la durée du stimulus, l'autre sur le nombre de vibrations (DUREE ou PULSE). Enfin nous avons montré que l'intégration de cette technique dans un dispositif embarqué au poignet peut réduire le temps de recherche d'une cible de manière significative.

Pour finir, le chapitre 7 présente notre conclusion et les perspectives que ces travaux peuvent générer.



# **Chapitre 2**

## **Etat de l'art sur la transmission d'information basée sur les modalités tactiles et/ou auditives**



## 2 Etat de l'art sur la transmission d'information basée sur les modalités tactiles et/ou auditives

La transmission d'information(s) entre un Homme et un système informatique nécessite des interfaces, pour recevoir et/ou émettre des stimuli, qui seront interprétés par le système et perçus par l'Homme (« les 5 sens »). Cependant, certaines situations, lorsque la perception visuelle est fortement contrainte ou déficiente, nécessitent de rendre perceptible, i.e. présenter, l'information sous une forme non visuelle. C'est justement dans un contexte de transmission non visuelle d'information que s'inscrivent nos travaux de thèse et plus spécifiquement sur des transmissions basées sur les modalités tactiles et/ou auditives. Dans ce chapitre, nous rappellerons les principales caractéristiques et limites des sens tactile et auditif de l'Homme, puis nous présenterons, avec des exemples choisis, le socle de connaissances sur lequel s'appuient nos travaux.

Ce chapitre se compose de trois grandes sections, présentant les modalités tactiles et auditives ainsi que les principes de la multimodalité, et d'une conclusion pointant vers les chapitres suivants.

### 2.1 Modalités tactiles

Ce nous appelons modalités tactiles au sens large regroupent les stimuli tactiles, les modes de communication nécessitant le sens tactile et les interactions tactiles. Nous présentons la grande famille des modalités tactiles dans les sections suivantes et dans cet ordre.

#### 2.1.1 Stimulus tactile

Un « stimulus tactile » sur la peau peut être, soit une vibration, soit une pression (dans nos travaux). Dans notre contexte, nous considérons qu'une vibration est un mouvement forcé (non naturel) d'oscillation autour d'une position avec une certaine durée, fréquence et amplitude<sup>1</sup>. Toujours dans notre contexte, nous considérons qu'une pression correspond à la force par unité de surface qu'exerce un solide ou un fluide (ex. eau/air) sur la peau.

Dans la section suivante nous nous intéressons à l'interprétation des stimuli tactiles par l'Homme.

<sup>1</sup> Why is Vibration Amplitude in G? : <http://www.precisionmicrodrives.com/tech-blog/2013/02/25/why-is-vibration-amplitude-in-g>

## 2.1.2 Perception chez l'Homme : modalité sensorielle tactile

La perception tactile est intégrée par le cortex somatosensoriel primaire S1 de notre cerveau. Celui-ci permet de localiser un stimulus sur la surface cutanée et d'en évaluer l'intensité, l'étendue et la durée. De plus, c'est grâce aux récepteurs de la peau que, la reconnaissance tactile des objets (formes et dimensions) et de la texture des surfaces (lisse, rugueuse, etc.) est possible. Nous allons décrire les principales capacités humaines liées au toucher et utilisables par l'Homme pour recevoir une information.

D'après le cours de (Raggenbass 2011) sur la somesthésie, il y a trois grands récepteurs (mécanorécepteurs) permettant de percevoir la vibration et la pression. Ces trois récepteurs ont leur spécialité : les corpuscules de Pacini sont responsables de la sensation de vibration (haute résolution temporelle) ; les corpuscules de Meissner sont responsables des (brèves) sensations tactiles fines (haute résolution spatiale) ; les terminaisons de Ruffini sont responsables de la sensation de pression constante.

### **VIBRATION**

La vibration est perçue par l'Homme dans une gamme de fréquence allant de 30 Hz à 1 500 Hz, notamment bien perçue entre 100 Hz et 320 Hz et, perçue de manière optimale à une fréquence de 250 Hz (Makous, Friedman, and Vierck 1995).

### **PRESSION**

La pression et le contact sont difficilement séparables du fait qu'elles sont liées l'une à l'autre. Ainsi nous parlerons uniquement de modalité sensorielle de pression. Cette perception est déclenchée par des appuis sur la surface cutanée. Par exemple, nous sommes capables de percevoir la présence d'une feuille de papier de cinq milligrammes et d'une surface de 64 mm<sup>2</sup> (densité de 80g/m<sup>2</sup>) posée sur la peau. Cette perception est fortement liée à la zone du corps sollicitée (seuils de sensibilité à l'intensité différents), aux nombres de points de pressions et à la distance entre ces points (résolutions spatiales différentes).

Après avoir vu les limites générales des stimuli vibrotactiles et de pression sur la peau, nous allons apporter dans la section suivante, des précisions sur les propriétés tactiles des différentes zones du corps.

### 2.1.2.1 Sensibilité, résolution spatiale et temporelle

Les perceptions tactiles proviennent de toute la surface du corps, mais elles n'ont pas toutes la même précision, ni la même nature.

**PARTIES DU CORPS ET LEURS PROPRIETES**

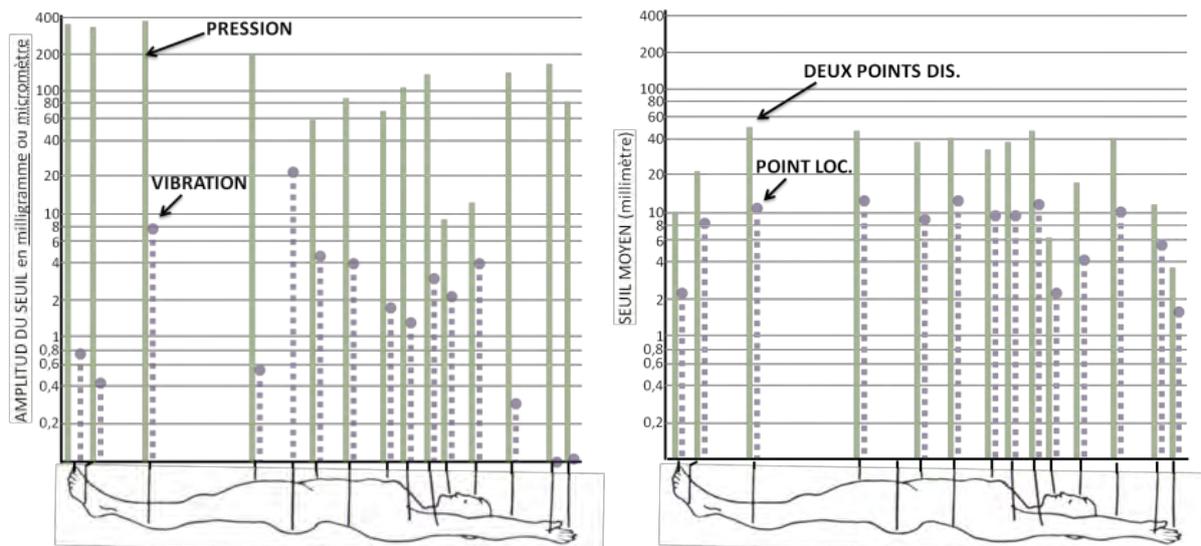
Différentes zones de la peau ont différentes propriétés, comme la « peau velue » qui peut jouer un rôle de capteur de pression semblable aux corpuscules de Meissner (sensations tactiles fines). De plus, chaque poil possède un muscle arrecteur<sup>2</sup> dont la fonction principale est d'hérissier le poil, suite à une sensation physique (ex. le froid) ou émotionnelle (ex. la peur). Ainsi, les peaux velues transmettent plus d'information que les peaux glabres. De plus, la peau n'est pas de même épaisseur selon les zones du corps. La peau a une épaisseur de 0,6 mm sur le corps, de 0,12 mm sur le visage ; elle est plus fine sur les lèvres et autour des yeux (0,3 mm sur les paupières) et plus épaisse sur les paumes de mains et la plante des pieds (1,2 mm à 4,7 mm). Cette différence d'épaisseur entraîne une différence de sensibilité et des seuils de perceptions différents.

**PARTIES DU CORPS ET LEURS SENSIBILITES**

Plusieurs travaux, dont ceux de (Johnson and Phillips 1981; Lederman and Klatzky 2009), ont dressé des cartes de sensibilité en fonction des parties du corps et de la nature du stimulus. Ainsi, la main et le poignet sont très sensibles aux vibrations (<0.1  $\mu$ m, cf. Figure 2-1 gauche) et, ont une bonne acuité spatiale, c'est-à-dire que ces zones sont capables de discriminer deux points de pression (localisés) très proches l'un de l'autre (<6 mm, cf. Figure 2-1 droite). Tandis que le visage est très sensible à la pression (<20 mg, cf. Figure 2-1 gauche) et les lèvres sont capables de discriminer une distance de 6 mm entre deux points de contact (cf. Figure 2-1 droite). Ces cartes de sensibilité peuvent être améliorées en prenant en compte la latéralisation des Hommes (gaucher/droitier). Par exemple, dans le domaine de l'interaction Homme-machine, certains travaux sont allés jusqu'à évaluer l'efficacité de techniques d'interaction en fonction de la latéralisation des participants (Wagner *et al.* 2013).

---

<sup>2</sup> Muscle arrecteur : petit muscle lisse à innervation sympathique segmentaire dont la contraction (sous l'effet du froid, de la peur, etc.) déclenche le redressement du poil (« chair de poule »).



**Figure 2-1 : Plus le seuil est bas, plus la partie du corps est « sensible » ; à gauche, seuil de perception de vibrations et de pressions ; à droite, seuil de discrimination d'une distance entre deux points de contact (« gap detection ») et seuil de discrimination d'un second point de contact autour d'un premier (« two-point discrimination ») ; source : (Johnson and Phillips 1981)**

Ces informations sont pertinentes pour concevoir des dispositifs adaptés à la sensibilité des différentes zones du corps. Par exemple, un dispositif « affichant » un motif dans le dos, pourrait être composé d'une matrice de « picots » où l'espace entre chaque picot serait au minimum de 12 millimètres et chaque picot exercerait une pression minimale de 81 milligrammes sur la peau du dos. Ainsi, ce dispositif serait une matrice de 20cm<sup>2</sup> composée de 256 picots (16x16) et permettrait (en théorie) d'afficher une icône de 16x16 pixels, perceptible via la matrice sur le dos.

Les sections suivantes traitent de la dégradation du sens tactile, posent plusieurs définitions liées au sens tactile (dont le sens haptique) et présentent une expérience mettant en avant les propriétés du sens haptique et la participation de l'utilisateur dans une tâche de reconnaissance de forme.

### 2.1.2.2 La dégradation du sens tactile

Le sens tactile peut être dégradé ou se dégrader : les seuils de perception peuvent être altérés. Les causes possibles sont des accidents (brûlures, nerf coupés, etc.), innées (pas de sensation de douleur notamment, cas de désafférentation par ex.), liées à l'âge (Jacquet, Chambert, and Remache 2013) ou à des pathologies neurologiques. Ainsi, la baisse de sensibilité vibratoire semble se produire vers l'âge de 50 ans et serait beaucoup plus significative au niveau des extrémités inférieures que supérieures<sup>3</sup>. Certaines études ont

<sup>3</sup> <http://papidoc.chic-cm.fr/O8sensstouch.html> (Dr Lucien Mias - mai 1992)

démonstré, qu'il n'y avait pas de perte de la perception vibrotactile du bout des doigts avec l'âge (Stuart *et al.* 2003).

### 2.1.2.3 Définitions et rôles du tactile

---

Avant de présenter des techniques d'interaction permettant de transmettre de l'information en utilisant la modalité sensorielle tactile, nous prendrons le temps de définir plusieurs termes liés à cette modalité.

Le **sens cutané**, ou du toucher, permet de percevoir les propriétés matérielles d'un objet : texture, dureté et température (modalité sensorielle cutanée).

Le **sens kinesthésique**, ou proprioception, fournit les informations de position et de mouvement des membres : géométrie d'un objet, position de la main, etc. (modalité sensorielle kinesthésique).

Le **sens haptique** intègre les informations somatosensorielles notamment pour la reconnaissance d'objet. Ce sens résulte de la combinaison des sens cutané et kinesthésique (Appelle 1991; Révész 1950) (modalité sensorielle haptique).

Le **sens tactile** regroupe les sens cutané, kinesthésique et haptique (sens cutané et kinesthésique). Le sens tactile est la modalité sensorielle tactile.

La **modalité d'interaction haptique** regroupe toutes les techniques d'interaction se basant sur les modalités sensorielles cutanées, kinesthésiques et haptiques (sens cutané et kinesthésique). Remarquons ici que le terme qu'il serait logique d'employer afin d'avoir une continuité dans les définitions est : modalité d'interaction tactile. Seulement, ce terme est déjà utilisé pour définir les techniques d'interaction avec des dispositifs tactiles, tel qu'un écran (tactile), un clavier (tactile), un bouton (tactile), etc. Nous choisissons donc, à contrecœur, de respecter la définition existante de la modalité d'interaction tactile et de prendre à la place le terme de « modalité d'interaction haptique ».

Notons que le terme modalité tactile peut se référer au mode de communication mettant en œuvre des modalités sensorielles ou d'interactions.

Le **tacton**, ou code, motif, modèle vibrant, a été clairement défini par (Brewster and Brown 2004) : « Les tactons, ou icônes tactiles, sont des messages structurés et abstraits qui peuvent être utilisés pour communiquer des messages non visuels. Une gamme de différents paramètres peut être utilisée pour la conception de tactons comme : la fréquence, l'amplitude, la durée d'une pulsation, ainsi que d'autres paramètres tels que le

rythme et la position du vibreur sur les zones du corps. Les tactons ont le potentiel d'améliorer les interactions dans plusieurs domaines, en particulier lorsque l'affichage visuel est surchargé, limité en taille ou non disponible, tels que les interfaces pour les personnes aveugles ou dans les appareils mobiles et portables ».

#### 2.1.2.4 L'« active touch »

Une caractéristique remarquable du sens tactile n'a pas encore été énoncée : l'« active touch ». Gibson (Gibson 1962) s'est posé la question des différences : Entre manipuler un objet et juste être en contact avec lui ? Entre bouger pour percevoir (toucher) et juste percevoir (être touché) ? Entre être actif et être passif ? Son expérience, connue sous le nom de « The great cookie-cutter experiment », compare deux conditions :

- Condition 1 : L'expérimentateur presse un moule de cookie dans la paume de l'utilisateur et celui-ci doit reconnaître la forme ; résultat : 49% d'identifications correctes.
- Condition 2 : l'expérimentateur donne le moule à l'utilisateur et celui-ci est libre de le manipuler, le toucher ; résultat : 95% d'identifications correctes.

Cette expérience démontre que l'exploration active (ici, manipulation), et donc le sens haptique est essentiel dans notre habilité à percevoir des objets dans le monde physique (pour aller plus loin voir (Gordon 1978; Lederman and Klatzky 2009)).

#### 2.1.3 Artéfacts technologiques pour produire des stimuli tactiles

Il existe des technologies capables de produire des stimuli nous faisant percevoir :

- **Des vibrations**, par l'utilisation de moteur vibrant, dont leur durée d'activation, leur fréquence, leur amplitude et bien sûr, leur placement sur le corps peuvent varier. Notons que la fréquence d'un moteur vibrant est liée au nombre de rotations par minute (RPM) du dispositif ; par exemple 12 000 RPM équivaut à une fréquence de 200 Hz, ce qui est proche de la fréquence vibratoire perceptible optimale (voir (Makous *et al.* 1995)). La Figure 2-2 présente plusieurs artéfacts technologiques pour produire des stimuli vibrotactiles.



Figure 2-2 : exemples de moteur vibrant ; de la gauche vers la droite : VPM2, Samsung A400 vibration motors, LilyPad Vibe Boards et le C2 tactor et Vibrating "Eyeronman" Vest.

- **Des pressions**, par l'utilisation d'actionneur mécanique (articulation, picot, gonfleur, etc.). Ces dispositifs peuvent avoir plusieurs degrés de liberté (nombre de dimensions possibles pour faire un mouvement), et requièrent un contact avec une partie de notre corps tel que la main, le dos etc. L'amplitude du stimulus peut varier, et si le dispositif dispose d'une matrice d'actionneurs, alors chaque actionneur est indépendant (cela permet de créer des illusions tactiles notamment). La Figure 2-3 présente plusieurs artéfacts technologiques pour produire des stimuli de pression.

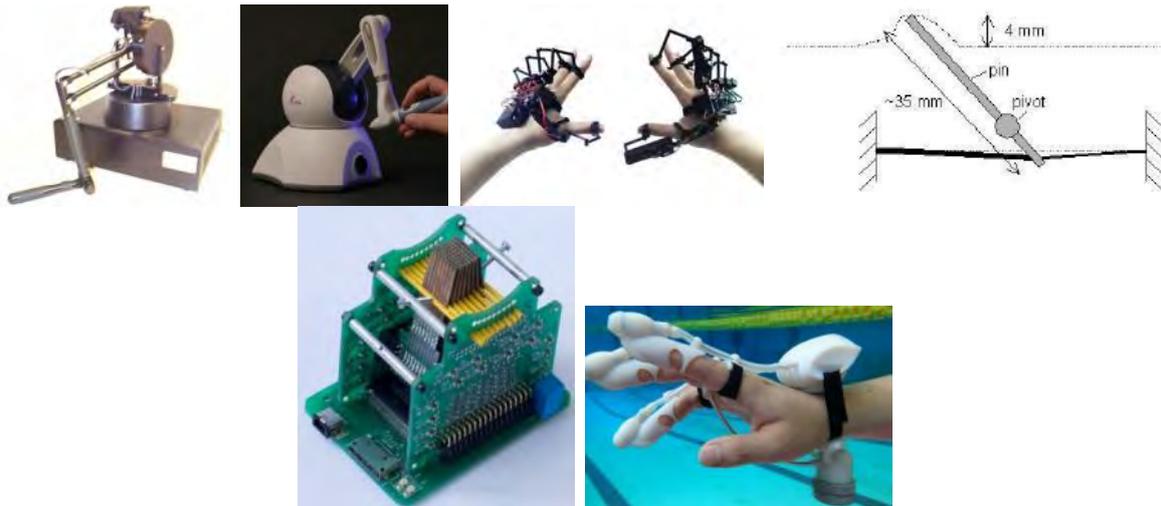


Figure 2-3 : exemple de dispositifs capables de produire un stimulus de pression ; de gauche vers la droite : PHANToM 6/6 DOF, PHANToM Omni 6/3 DOF, Dexmo F2, Mechanism of the Tactile Vest actuator (Nakamura and Jones 2003), « STReSS » (Pasquero and Hayward 2003) et un Haptic Glove.

#### 2.1.4 Mise en œuvre de la modalité sensorielle tactile dans des situations interactives

Nous structurons la mise en œuvre de la modalité sensorielle tactile autour de trois catégories de modalités d'interaction haptique :

- Les modalités (d'interaction) **haptiques** englobent toutes les modalités d'interaction se basant sur les modalités sensorielles tactiles : bouger pour percevoir (toucher) ou être perçu (faire un geste) ou juste percevoir (être touché). C'est mettre l'accent sur ce que l'utilisateur perçoit pendant la tâche ;
- Les modalités (d'interaction) **vibrotactiles**, qui sont des sous-modalités de l'haptique avec comme composante principale l'utilisation de tactons via des moteurs vibrants disposés en ligne ou répartis sur une surface ;
- Les modalités (d'interaction) **tangibles**, qui se différencient des autres modalités haptiques par le rôle primordial de l'objet dans l'interaction, mais aussi dans la prise en main de l'objet. De nombreuses techniques utilisent une interface physique (ici, tangible) pour interagir entre l'Homme et la machine (par exemple, une souris, un

clavier, un écran tactile). C'est mettre l'accent sur le fait de manipuler un objet pour interagir avec le système.

Dès lors qu'une technique d'interaction nécessite l'utilisation d'un objet physique, sa classification comme étant soit haptique, soit tangible résulte d'un choix du concepteur qui a tranché pour l'un des deux paradigmes : percevoir pour interpréter pendant l'interaction ou manipuler pour percevoir et interagir.

#### **2.1.4.1 Modalité d'interaction haptique : mises en œuvre réelles**

La modalité d'interaction haptique, rappelons-le, regroupe toutes les modalités se basant sur les modalités sensorielles cutanées, kinesthésiques et haptiques, et peut être décomposée en sous-modalités d'interaction telles que le retour de force, la pression, l'effort, la marche à l'aveugle (blind walking), l'atteinte à l'aveugle (blind reaching) et le relief. Nous traiterons les modalités vibrotactiles et tangibles dans des sections indépendantes. Cependant, nous ne traiterons pas les modalités gestuelles qui requièrent la modalité sensorielle kinesthésique seule (ex. kinect, gaze-tracker, etc.). Dans les sections suivantes nous définissons chacun de ces termes et les illustrons avec des situations interactives de transmission d'informations. Nous tâcherons de définir et d'appuyer ces définitions par des exemples choisis dans la littérature. La composante principale de cette modalité est de percevoir pendant l'interaction. Cependant, la composante secondaire de la modalité haptique est le mouvement d'une partie du corps pendant l'interaction. Cela fait référence, dans la classification de (Roudaut and Lecolinet 2007), aux modalités : mouvement en entrée (système) et kinesthésique en sortie (système).

##### **MODALITE D'INTERACTION « RETOUR DE FORCE »**

Le retour de force est une réponse mécanique du système qui peut modifier la résistance au mouvement de l'utilisateur, voire le mouvement lui-même. Cette modalité d'interaction est souvent utilisée afin de reproduire des comportements (ou des propriétés) physiques du monde réel encodés dans une simulation virtuelle. Intéressons-nous aux dispositifs à retour de force pour des situations de transmissions d'informations.

Les dispositifs à retour de force sont conçus pour accompagner les mouvements du corps qu'ils soient simples (translation ou rotation) ou complexes (translations et rotations multiples). Certains de ces dispositifs peuvent également être utilisés comme des capteurs de mouvement ou de force. Par exemple le PHANToM « 6/6 DOF », peut capter les mouvements de l'utilisateur en six dimensions (trois translations et trois rotations), ce type

de dispositif peut également transmettre à l'utilisateur un mouvement en six dimensions (d'où le « 6-in/6-out Degree Of Freedom (DOF) »). Notons que le degré de liberté (DOF) et la précision que ces dispositifs offrent, sont importants à la fois pour l'interaction elle-même mais aussi pour l'immersion (ex. jeux sur bornes d'arcades, consoles ou PC (Choi, Chang, and Kim 2004)).

Dans le contexte bien précis d'une exploration « virtuelle » avec une modalité haptique de retour de force, les travaux de (Nikolakis and Tzovaras 2004) ont porté sur la perception d'objets virtuels. Leur système est basé sur la combinaison d'un PHANToM™ « 6-in/3-out DOF » et d'un CyberGrasp™ (gant à retour de force), voir Figure 2-4. Ainsi équipé, l'utilisateur peut percevoir la forme d'un objet 3D en le « touchant » virtuellement avec des mouvements réels : ces dispositifs peuvent fournir un retour de force allant jusqu'à 1,75 newton sur l'ensemble du bras de l'utilisateur (main, poignet, avant-bras, coude, bras et épaule) avec le PHANToM™ et sur chacun de ses doigts (gant CyberGrasp™). Les dispositifs haptiques ont aussi fréquemment été utilisés dans le cadre des technologies d'assistance. Par exemple, (Schloerb *et al.* 2010) ont conçu un système basé sur un Desktop PHANToM® « 6-in/3-out DOF » afin d'améliorer et accélérer la compréhension d'un environnement inconnu pour un utilisateur. L'utilisateur manipule le dispositif à l'instar d'une canne blanche miniature et perçoit des vibrations correspondant à des obstacles potentiels. L'utilisateur se déplace dans l'environnement virtuel avec un clavier et lorsqu'un mur se dresse devant lui, sa « canne » virtuelle recule d'autant pour signaler l'obstacle.

En résumé, ces exemples illustrent le fait que le retour de force peut être utilisé pour transmettre des informations telles que des obstacles et la distance qui nous en sépare mais aussi, pour « explorer » une pièce ou « toucher » un objet virtuel.



Figure 2-4 : Illustration des dispositifs PHANToM™ (à gauche) et CyberGrasp™ (à droite).

A l'instar du retour de force qui utilise le sens kinesthésique, cette modalité d'interaction haptique utilise principalement le sens cutané. Des sensations usuelles, comme tenir un objet ou être en contact, peuvent être encodées dans une simulation virtuelle afin d'en améliorer l'immersion ou alors d'indiquer simplement une direction. Intéressons-nous à ces dispositifs pour des situations de transmission d'informations.

Un des avantages certains de dispositifs basés sur des modalités haptiques de pression, est qu'ils ne modifient pas le mouvement de l'utilisateur et fournissent des informations en continu. Néanmoins l'interaction fonctionne uniquement lorsque l'utilisateur est en contact avec le dispositif.

Prenons en exemple le gant à coussinets qui se gonflent (« Hands Omni haptic glove ») de l'Oshman Engineering Design Kitchen (Université de Rice, Houston, Texas, 2015)<sup>4</sup>. Ce gant fournit une pression sur le bout des doigts, capable de reproduire l'appui sur une touche ou la préhension d'un objet virtuel (en l'occurrence, une arme à feu, voir Figure 2-5-gauche). D'autres parties du corps ont fait l'objet d'études, comme le dos. Le dos offre une grande surface de contact et peut être utilisé comme interface sensorielle afin de transmettre de l'information. Le prototype « Tactile Vest » de (Nakamura and Jones 2003) a été conçu en ce sens. Cette veste est composée de presque quatre cents picots (matrice de 28x14) capables d'exercer une pression suffisante sur le dos afin qu'un utilisateur puisse percevoir et identifier chaque picot. Cette veste pourrait par exemple, transmettre des motifs en 2D ou générer un frisson dans le dos (immersion). Sous une autre forme, des effecteurs mécaniques sous une main, peuvent servir à indiquer directement une direction. Le prototype « HaptiQ », amélioré par les travaux de (Siegfried 2015), permet d'indiquer les quatre directions cardinales et leurs intermédiaires, soit huit directions (voir Figure 2-5-droite).

En résumé, la modalité d'interaction pression peut être utilisée pour transmettre des informations visant à renforcer l'immersion ou transmettre un message sous une forme codifiée (ex. avec HaptiQ). Nous avons choisi de ne pas parler de techniques basées sur l'électrotactile (Paul Bach-y-Rita *et al.* 1998), technologie clairement basée sur du contact avec la peau uniquement. Nous y ferons référence spécifiquement dans le chapitre 6, présentant notre troisième étude, dans un état de l'art spécifique à la substitution sensorielle.

---

<sup>4</sup><http://news.rice.edu/2015/04/22/gamers-feel-the-glove-from-rice-engineers-2/#sthash.wPQOA0XX.dpuf>



**Figure 2-5 : Dispositifs utilisant la modalité haptique de pression. A gauche, illustration d'un « Hands Omni haptic glove », à l'intérieur des coussinets se gonflent afin de reproduire par exemple, la prise en main d'un objet. A droite, illustration d'un HaptiQ, qui transmet une direction à l'aide de picots sous la main.**

### **MODALITE D'INTERACTION « EFFORT »**

L'effort est une modalité sensorielle haptique souvent utilisée dans la vie quotidienne : soupèsement d'objet (ex. « combien pèse ma valise de voyage ? ») ; pousser une porte ; jeter un objet avec force ; etc. Cependant, pouvons-nous utiliser la modalité haptique de l'effort afin de transmettre des informations de manière précise (efficace) ?

Dans la littérature, nous retrouvons de nombreux travaux sur l'estimation de valeurs à partir d'un effort. Par exemple, les travaux de (Proffitt *et al.* 2003) ont montré que porter du poids, en l'occurrence un « sac à dos lourd » ( $\frac{1}{5}$  du propre poids des participants), permet de mieux estimer une distance parcourue que sans poids additionnel. (Witt, Proffitt, and Epstein 2004), dans une expérience similaire, ont montré que l'estimation d'une distance en jetant une « balle lourde » était plus précise que celle d'un jet d'une « balle légère ». Néanmoins, les travaux de (Woods, Philbeck, and Danoff 2009) attirent notre attention sur le fait que la perception de l'effort puisse être une modalité permettant une estimation correcte de distances, mais soulignent le fait que la reproductibilité des expériences n'est pas « évidente » dans ce contexte, car de nombreux facteurs, non pris en compte dans les protocoles, pourraient impacter les résultats. De plus, nous remarquons que les exemples choisis ne sont pas à proprement parler des techniques d'interaction, faute de système informatique (pas de Machine, pas d'IHM), et relèvent plus du domaine des sciences cognitives.

En résumé, il existe un grand éventail de conceptions possibles pour des interactions capables de transmettre des informations avec la modalité haptique de l'effort. Cependant, il faut prendre en compte les capacités humaines dans la conception de ses techniques :

anticipation de l'effort à fournir pour effectuer un mouvement précis ; adaptation de l'effort à la tâche (lancer plus loin, porter un sac ou soulever un objet plus lourd). Et surtout, il faut concevoir des dispositifs interactifs capables de générer ou de recevoir des stimuli haptiques d'effort.

#### ***MODALITE D'INTERACTION MARCHE A L'AVEUGLE (BLIND WALKING)***

La marche à l'aveugle est une modalité haptique particulière. Elle fait appel au sens kinesthésique, mais aussi au sens vestibulaire (il existe une sorte de centrale inertielle dans l'oreille interne permettant de capter avec précision les mouvements de notre tête et de fait, de tous nos déplacements). De plus, lors du déplacement de son corps (par ex. la marche), le sujet peut compter le nombre de ses pas, mesurer la durée de son déplacement ou encore apprécier son effort fourni pour marcher. La modalité haptique de la marche à l'aveugle fait appel à un nombre impressionnant de nos sens et à leurs propriétés, mais quelle est la précision de notre perception de la marche ? Nous nous intéressons ici, uniquement à l'estimation de distance.

Les expériences, utilisant la modalité marche à l'aveugle, requièrent généralement, et comparativement aux autres modalités d'interactions haptiques, un matériel nécessitant une grande surface au sol. En effet, les sujets doivent généralement être debout et marcher. Toutefois, des jalons, un mètre et un équipement informatique simple, peuvent suffire à une expérimentation visant à évaluer des techniques basées sur la marche à l'aveugle. Cela rend cette modalité simple d'utilisation et écologique (dans le sens « au plus proche du réel ») pour transmettre de l'information.

Un exemple typique d'estimation de distance avec de la marche à l'aveugle nous est donné par les travaux de (Ellard and Shaughnessy 2003). En effet dans leur première expérience, les sujets devaient estimer une distance, en parcourant un trajet (aller-retour) en aveugle. Les résultats de cette étude ont montré qu'il était possible d'estimer correctement des distances (4, 6, 8 et 10 mètres) en utilisant juste la modalité de la marche à l'aveugle .

#### ***MODALITE D'INTERACTION ATTEINTE A L'AVEUGLE***

La modalité de l'atteinte à l'aveugle fait appel au sens kinesthésique et plus particulièrement à la « position » des articulations. Par exemple, la position d'une cheville sur un sol incliné ou la position d'un bras (épaule, coude, poignet, ...) pour atteindre une cible. Toutefois, nous nous posons la question : une transmission d'information de précision

est-elle possible avec des interactions basées sur la modalité haptique de l'atteinte à l'aveugle ?

Les articulations humaines sont identifiées et bien détaillées dans des ouvrages contemporains de médecine. Nous y apprenons, entre autres, qu'elles ne sont pas égales en termes de « degré de liberté », ni en amplitude de mouvement<sup>5</sup>. Par exemple, un coude a 160° d'amplitude, un genou a 170° et une cheville a 65°, et ces trois articulations permettent une rotation (flexion-extension) dans un seul plan (1D) ; un poignet permet d'effectuer des rotations en 2D (85° d'amplitude pour une flexion-extension et 45° pour une abduction-adduction) et en 3D pour une épaule (205° flexion-extension, 210° abduction-adduction et 140° rotation interne-externe).

De récents travaux se sont intéressés à l'estimation à l'aveugle d'une valeur avec une interaction basée sur la modalité de l'atteinte à l'aveugle. Commençons par les articulations du pied et de la main. Dans une série d'expériences, (Hajnal, Abdul-Malak, and Durgin 2011) ont utilisé un plan incliné (sept angles de 4° à 16°) où les participants ont en moyenne surestimé l'angle du plan (Angle estimé = Angle cible **x2,5**). Dans leur quatrième expérience, (Hajnal *et al.* 2011) ont utilisé un PHANTOM™ pour simuler l'inclinaison (seize angles, de 0° à 90°) d'un doigt sur une surface dure. Ainsi, les sujets avaient l'impression de toucher une table (par ex.) avec leur index dans un dé à coudre. Les résultats montrèrent que les participants surestimaient toujours de 6 degrés l'angle cible (Angle estimé = Angle cible x 1 +6). Cette modalité peut aussi être utilisée en « sortie utilisateur ». Par exemple, les travaux de (Altenhoff *et al.* 2012) portèrent sur la différence d'estimation entre une réponse verbale et une réponse par atteinte à l'aveugle, en l'occurrence la position du bras (droit) des participants. Leurs résultats montrèrent que les réponses des participants via une atteinte à l'aveugle (Distance estimée = Distance cible **x0.84 +2.3**) étaient plus précises que les réponses verbales (D-EsT = D-Cib **x1.1 -34**), car les réponses verbales étaient grandement sous estimées. Un dernier exemple pour illustrer cette modalité, nous vient des travaux de (Enokibori and Mase 2014), qui portent sur la précision d'une technologie basée sur du tissu « intelligent » permettant de distinguer neuf angles d'un coude pendant la flexion de celui-ci. Ce futur capteur pourrait à terme, être utilisé pour remplacer les réponses verbales des utilisateurs ou leur permettre de communiquer avec le système par une interface « haptique ».

<sup>5</sup> [https://fr.wikipedia.org/wiki/Articulation\\_synoviale](https://fr.wikipedia.org/wiki/Articulation_synoviale) (articulation mobile)

En résumé, la modalité haptique de l'atteinte à l'aveugle permet de transmettre des informations vers le système ou vers l'utilisateur et offre potentiellement une alternative à la réponse verbale dans des tâches de « stimulus-réponse » (« je perçois, je réponds »).

#### **MODALITE D'INTERACTION « RELIEF »**

Le relief, ou plutôt « surface non plate », donne l'illusion d'une troisième dimension. Ainsi une image en 2D peut être perçue via notre sens haptique. La transmission d'information via la modalité haptique du relief rend accessible et ce, de manière très simple, une information visuelle.

Nous connaissons déjà l'intérêt de toucher la texture d'un objet (perception du piquant, lisse, velouté, rugueux, etc.) et d'explorer les formes d'une surface (braille, sculpture, bas/haut-relief, etc.). Certains travaux sont allés plus loin dans l'utilisation de cette modalité, en proposant des cartes « tactiles » composées d'éléments simples mis en relief. Les cartes en relief sont de bons exemples, car elles rendent accessible des informations géographiques<sup>6</sup> comme les montagnes ou les fleuves, mais elles peuvent aussi *magnifier*<sup>7</sup> un élément du terrain, comme un point d'intérêt (Brock *et al.* 2015). Un problème subsiste néanmoins dans l'utilisation de la modalité relief afin de transmettre une information : les objets sont statiques ! En effet, prenons encore l'exemple de la carte tactile, celle-ci doit d'abord être « imprimée » avant toute utilisation. Ceci permet, par exemple, de présenter dans les lieux culturels de plus en plus d'informations en relief (braille, plan, symbole, etc.). Dans de tels contextes, les informations diffusées sont toujours et uniquement statiques. Cependant, ce type de solutions est totalement incompatible pour des interactions de transmission « dynamique » d'informations.

Même s'il n'existe pas (encore) de technologie capable de nous faire ressentir la caresse de l'herbe dans une prairie virtuelle ... des chercheurs canadiens (Pasquero and Hayward 2003) ont mis au point un dispositif, « STReSS », capable de produire dynamiquement des images tactiles et même des films tactiles d'une fréquence incroyable de 700 images par seconde (700 Hz). Pour l'instant leur dispositif est suffisant pour y placer le bout d'un doigt, mais leur technologie de « films tactiles » permet de transmettre des textures, comme si l'on explorait une surface avec son doigt. (Qi Wang and Hayward 2006) ont d'ailleurs conçus une version

<sup>6</sup> Par exemple : <http://wiki.openstreetmap.org/wiki/HaptoRender>.

<sup>7</sup> Magnifier : Rendre quelqu'un, quelque chose plus grands moralement ; élever, idéaliser : Ces souvenirs magnifiaient une réalité médiocre. (larousse.fr) ; La loupe (en anglais, magnifier) est un composant du système d'exploitation Windows de Microsoft.

paramétrable et portable de STReSS. Des équipes françaises se sont également intéressées à cette modalité haptique du relief pour transmettre des informations. Ainsi, en se basant sur une technologie différente, l'équipe française de chercheurs (Zeng *et al.* 2010) a conçu un dispositif permettant de transmettre des textures ou des formes, mais aussi, pour transmettre la description d'un obstacle détecté devant l'utilisateur (Zeng, Prescher, and Weber 2012) grâce à une matrice braille de 30 x 32 picots cadencée à 5 Hz. Une deuxième équipe française (Velazquez *et al.* 2007) a proposé une interface à stimulation tactile pour afficher des cartes géographiques dynamiques égocentrée.

En résumé, la modalité relief est actuellement utilisée pour un usage limité aux mains. Pourtant la technologie promet la création d'écran en relief dynamique, ce qui lèverait probablement le verrou technologique des « cartes en relief non-dynamiques » et permettrait également de nouvelles interactions comme par exemple, des jeux en « 3D tangibles » (voir projet InForm et <http://tangible.media.mit.edu/project/inform/>)

#### **CONCLUSION DE LA MODALITE HAPTIQUE**

**La modalité haptique se base sur les mouvements du corps (sensations kinesthésiques) et sur les sensations cutanées.** Cette modalité permet de transmettre des informations brutes : longueur, direction, forme, texture, angle, durée, effort, position, résistance. Les interactions haptiques sont plutôt écologiques et améliorent l'immersion dans les environnements virtuels. Toutefois, la modalité haptique nécessite un contact permanent entre l'utilisateur et le dispositif (ou avec le sol dans le cas de l'atteinte à l'aveugle). Ce contact continu permet donc la perception progressive des mouvements, et par conséquent, est une solution pouvant potentiellement favoriser une intégration progressive des informations. Néanmoins, dans le cas d'une estimation de valeur, il est nécessaire d'attendre la fin de la transmission pour percevoir entièrement la valeur cible (ex. le bon angle avec le coude, tout le motif d'une texture à motif, etc.).

Dans la classification de (Roudaut and Lecolinet 2007), les modalités d'interaction « relief » et « vibrotactile » sont regroupées dans la modalité « retour tactile », nous avons préféré les dissocier. D'une part, se sont deux modalités haptiques différentes et d'autre part, nous avons souvent utilisé la modalité haptique vibrotactile pour concevoir nos interactions. C'est pourquoi nous traitons cette modalité dans la section suivante et à part entière.

### 2.1.4.2 Modalité d'interaction, haptique, vibrotactile

La modalité vibrotactile décrite ici, correspond à l'utilisation de moteurs vibrants (ex. VPM2<sup>8</sup>) afin de transmettre de l'information. Les moteurs vibrants peuvent être disposés en une seule ligne ou sur une surface (Tan and Pentland 1997).

Les principaux atouts des dispositifs vibrotactiles sont : la mobilité, la simplicité et leurs qualités non-intrusives. Les moteurs vibrants ont une petite taille (ex. Ø12 mm x 3,4 mm ht.) et un faible poids (ex. 1,45 g). Ceci leur permet d'être intégrés dans des tissus ou des accessoires vestimentaires : gants, gilet, ceinture, bracelet, dossier de fauteuil, etc. Ils doivent néanmoins être reliés à un contrôleur (souvent une carte avec sortie analogique, tel que Arduino<sup>9</sup> ou Phidget<sup>10</sup>). Toutefois, l'utilisation des contrôleurs requiert quelques connaissances en programmation. La conception des dispositifs vibrotactiles demande, en outre, des notions de bricolage (ex. soudure à l'étain, fixation par vis, etc.), de couture (ex. fil à coudre conducteur et fil en soie), d'impression 3D (ex. pour faire un coffret de cartes Arduino) ou de réseau informatique pour les dispositifs connectés, mais ne requière pas des notions d'ingénierie avancées à l'instar de beaucoup de dispositifs haptiques (voir ex. du PHANTOM™ ou du CyberGrasp™). Cependant, la transmission d'informations, avec la modalité vibrotactile, s'effectue uniquement dans le sens « machine vers Homme ». Nous présentons ci-après quelques exemples choisis, dans la littérature, d'interaction basée en premier lieu sur la modalité vibrotactile et mise en œuvre dans différentes configurations.

#### **DISPOSITION EN LIGNE**

Les dispositifs en ligne sont souvent des bracelets ou des ceintures, et plus rarement des serre-têtes, des brassards et des serre-cuisses. Le bracelet de (Brunet *et al.* 2013), composé de huit moteurs vibrants (PrecisionMicrodrives 310-113<sup>11</sup>) équirépartis autour du poignet, pouvait transmettre efficacement vingt-un messages différents. Les tactons (motifs vibrants) utilisaient le rythme, la fréquence et la position des vibreurs. La ceinture de (Tsukada and Yasumura 2004), composée de huit vibreurs équirépartis autour de la taille, pouvait transmettre la direction à suivre en faisant vibrer simplement le moteur associé à la direction. L'encodage est donc plus direct. (Heuten *et al.* 2008), ont repris des travaux

<sup>8</sup> <http://www.gotronic.fr/art-vibreux-vpm2-12006.htm>

<sup>9</sup> <https://www.arduino.cc/>

<sup>10</sup> <http://www.phidgets.com/>

<sup>11</sup> <https://catalog.precisionmicrodrives.com/order-parts/product/310-113-10mm-vibration-motor-3mm-type>

existants et notamment ceux de (Tsukada and Yasumura 2004), afin d'améliorer la précision des indications des directions. Ainsi, leur ceinture était constituée de six moteurs vibrants (Samsung A400 vibration motors<sup>12</sup>). La direction était transmise avec deux vibreurs dont l'intensité dépendait de leur éloignement à la direction cible : si une direction cible se trouve entre deux vibreurs (A et B) et, est plus près du vibreur A que du vibreur B, alors le vibreur A vibrera plus intensément que le vibreur B. Ceci octroie à cette ceinture, un angle de rotation potentiel de 360° et requiert une interprétation assez directe.

#### **DISPOSITION SUR UNE SURFACE**

Ainsi, (Toney *et al.* 2003) ont évalué un dispositif vibrotactile placé sur les épaules. Il en ressort des « guidelines » de conception concernant l'efficacité en fonction du nombre de vibreurs, la discrimination de ceux-ci et des résultats qualitatifs. (Jones, Nakamura, and Lockyer 2004) ont fait de même pour leur dispositif dorsal, avec la tâche supplémentaire de transmettre une direction. Dans le contexte des technologies d'assistance, (C. M. Jones *et al.* 2008) ont conçu une interaction capable d'améliorer la perception et la localisation d'un changement dans une scène visuelle. Les sujets étaient adossés à un dossier « vibrotactile » de quatre moteurs vibrants (VBW32<sup>13</sup>), de sorte que les sujets percevaient une vibration indiquant un quadrant de la scène (un quadrant par coin). Ce dispositif peut être intégré au siège conducteur d'une voiture par exemple, aidant ainsi le conducteur à faire attention à un piéton en bord de chaussée. En moto, il n'y a pas de siège (sauf exception), (Prasad *et al.* 2014) ont alors conçu le dispositif « Hapti-moto » afin d'aider le motard dans sa conduite. Leur dispositif est une veste « dorsale » composée de trois vibreurs (un sur chaque omoplate et un au milieu du dos (LilyPad Vibe Boards<sup>14</sup>). En se basant notamment sur les travaux de (Pielot, Poppinga, and Boll 2010), leur dispositif pouvait transmettre efficacement jusqu'à douze messages différents, pour encoder la distance et la direction du prochain tournant. En restant sur les dispositifs d'assistance, nous pouvons parler du prototype de (Lehtinen *et al.* 2012), qui à l'aide de deux vibreurs dans la paume de la main (un en haut, l'autre en bas), guident celle-ci vers une cible se trouvant toujours quelque part en haut à gauche de la position de départ. Cette technique de guidage permet, notamment aux déficients visuels, de localiser une petite cible sur une grande surface. Une autre application, pour les déficients visuels (Nicolau *et al.* 2013), a permis d'évaluer un dispositif

<sup>12</sup> [http://www.activemobiles.com/samsung\\_vibra\\_motors.htm](http://www.activemobiles.com/samsung_vibra_motors.htm)

<sup>13</sup> Voir, Audiological Engineering Corp., Somerville, MA, USA

<sup>14</sup> Lilypad product page. <http://tinyurl.com/k4oql7z>, 2012

vibrotactile (Ubibraille) composé de six vibreurs Lilypad<sup>14</sup> répartis sur les index, majeur et annulaire des deux mains afin de transmettre les lettres de l'alphabet avec un code Braille standard (matrice de 2x3).

#### **CONCLUSION DE LA MODALITE VIBROTACTILE**

La transmission d'informations avec des dispositifs « vibrotactiles » ne nécessite pas un mouvement d'une partie du corps. Ainsi, recevoir des informations « vibrotactiles » n'entrave pas les mouvements, et surtout, laisse les mains libres. Dans un contexte d'assistance, la modalité haptique vibrotactile permet de transmettre des informations contextuelles en avance (anticipation), ce qui n'est pas le cas des autres modalités haptiques telles qu'elles sont utilisées dans la littérature. Notons enfin que l'information véhiculée peut être encodée de différentes manières (répartition des vibreurs, tactons différents, etc.), induisant ainsi une perception plus ou moins directe de l'information.

#### **2.1.4.3 Modalité d'interaction tangible**

---

##### **DEFINITION REDUITE A DES UTILISATIONS NON VISUELLES**

Les objets tangibles sont une forme de Systèmes Interactifs Mixtes (SIM). Ils servent d'intermédiaires entre l'humain et le système. Les objets tangibles sont équipés de capteurs et/ou d'effecteurs afin de transmettre des informations tactiles (moteurs vibrants, centrales inertielles, etc.), auditives (micros, hautparleurs, etc.) ou visuelles (lampes, capteur de lumière, etc.). La transmission d'informations peut être vers l'utilisateur ou vers le système (Shaer and Hornecker 2010). Ainsi, la modalité tangible requiert un ou des objets intermédiaires, par le ou lesquels transiteront les informations. A la différence de la modalité d'interaction haptique, où nous devons avant tout nous focaliser sur nos mouvements (**percevoir pendant l'interaction**), la modalité dite tangible, nécessite que nous nous concentrons sur le ou les objets tangibles impliqués et leur manipulation (**manipuler pour interagir**).

Dans cette section, nous nous concentrerons sur les interactions tangibles impliquant des modalités tactiles (pas d'auditif). Nous verrons, également, qu'ajouter un composant supplémentaire, dans la chaîne de transmission d'information, doit se justifier par le fait de simplifier celle-ci ou d'être la technique d'interaction la plus adaptée à la tâche.

De nombreux dispositifs tangibles existent, mais la très grande majorité est dédiée aux mains. Aussi les tailles et les formes de ces dispositifs sont contraintes par les mains pour faciliter leur préhension et leur manipulation. Ils se doivent aussi d'être légers pour éviter

une fatigue, voir une tendinite, aux bras des utilisateurs. Nous illustrerons, avec des exemples choisis, deux catégories d'objets tangibles : les Smartphones et les cannes. Les premiers, car ils sont présent dans la poche de chacun ; les seconds, car le tangible pour les déficients visuels est une voix permettant d'envisager une transmission d'informations plus riche que le simple « text to speech » (ou synthèse vocale ou modalité auditive de la parole).

### **LES SMARTPHONES**

Les Smartphones, de part leurs nombreux capteurs et effecteurs<sup>15</sup>, transmettent des informations avec les modalités tactiles (pas d'auditif, nous limitons ici notre état de l'art). De plus, ils ont une batterie, sont programmables et possèdent entre autres<sup>15</sup> un GPS, des centrales inertielles et un moteur vibrant. Dans un contexte de transmission d'informations, le Smartphone est un dispositif de choix pour beaucoup d'expériences sur la mobilité. Par exemple, les travaux de (Pielot *et al.* 2011, 2012, 2010) aboutirent à une utilisation innovante du Smartphone, en indiquant une direction égocentrée parmi huit, avec des tactons inspirés du code morse. Cette technique est appelée : « PocketNavigator ». Les huit directions azimutales étaient : devant, derrière, gauche, droite, et les quatre directions intermédiaires ; les tactons (motifs vibrants) étaient composés d'une succession d'une paire de vibrations petites ou longues. Par exemple, le « devant » est codé par une succession de deux courtes vibrations et le devant-gauche (NO) est codé par une longue vibration suivie d'une courte. (Fujimoto and Turk 2014) ont d'ailleurs repris cette interaction pour une expérience visant à transmettre un couple de valeurs à l'utilisateur. Ce couple de valeurs était composé d'une direction et d'une distance, codant ainsi la position d'un danger imminent pour l'utilisateur piéton. La direction était transmise par « PocketNavigator » et la distance par un « compteur Geiger ». Le principe du compteur Geiger est de transmettre une valeur à l'aide d'une **succession continue** de pulsations, où le **tempo** de la succession (allure, rapidité relative, vitesse) code la valeur (ex. le taux de radioactivité). Ainsi, la combinaison des deux encodages décrite par (Fujimoto and Turk 2014) repose sur la variation de la durée des pulsations (« PocketNavigator ») et sur la variation de l'écart de temps séparant deux pulsations (« compteur Geiger »). Par exemple, « devant à 45 mètres » peut être encodé par « deux courtes pulsations de 200 ms séparées par un silence de 450 ms ». Dans cette expérience, le Smartphone était dans la poche des participants, ce qui peut s'apparenter à du « tangible hand-free » et donc en l'occurrence, à du vibrotactile. Toutefois, la possibilité de « scanner » l'environnement avec le Smartphone pour rechercher

<sup>15</sup> <https://fr.wikipedia.org/wiki/Smartphone>

(ou localiser) une cible, classe également le Smartphone comme un dispositif tangible. Voici un autre exemple d'utilisation de Smartphone avec « NaviRadar » de (Rümelin, Rukzio, and Hardy 2011). Ce dispositif est un Smartphone, auquel les auteurs ont ajouté un C2 tactor<sup>16</sup> (moteur vibrotactile haut de gamme). Ce dispositif peut transmettre un couple de valeurs (direction-distance) via des tactons et selon six combinaisons de deux facteurs parmi trois : intensité, rythme, durée (direction) et bruitage du signal (distance). Dans l'expérience, l'utilisateur tenait en main le Smartphone et avait un doigt sur le C2 tactor. Cependant, d'après les spécifications dudit moteur vibrant<sup>16</sup>, celui-ci a été conçu pour que ses vibrations puissent être ressenties à travers des vêtements, comme une poche par exemple. Là encore, le Smartphone peut donc être classé comme un dispositif vibrotactile.

### **LES CANNES**

Une canne pour aveugle se tient à la main et nous ne nous intéressons qu'à une de ses utilisations : détecter les obstacles devant soi. Ainsi, (Farcy *et al.* 2006) ont conçu et évalué plusieurs prototypes dont deux étaient destinés à détecter les obstacles. Le prototype « Tom Pouce » alertait l'utilisateur par une vibration, lorsqu'un obstacle était détecté dans un volume d'un cône, dont la profondeur était de trois mètres au maximum (réglable : 0,5m, 1,5m ou 3m), selon un angle de détection horizontal de 20° et de 50° en vertical (Cône à base elliptique 20°x50° et de profondeur maximale de 3m). Le second prototype, « Télétact II », outre l'interaction audio dont nous ne parlerons pas ici, permettait de transmettre la distance d'un obstacle, devant soi, avant collision. (Farcy *et al.* 2006) conçurent ce second prototype ainsi : ils utilisèrent deux « appareils » vibrants situés sur deux doigts différents (voir Figure 2-6). Le premier doigt était affecté aux détections entre 1,5 et 6 mètres : le signal était un léger pulse (courte vibration) pour des détections entre 3 et 6 mètres et consistait en une vibration plus forte pour une détection entre 1,5 et 3 mètres. Le deuxième doigt était concerné par les alertes, i.e. des détections inférieures à 1,5 mètres. Ce dernier prototype présente donc deux interactions : l'une tangible, pour l'estimation de distance et l'exploration de l'environnement devant soi ; et l'alerte vibrotactile pour avertir l'utilisateur d'un danger proche. (Buchs, Maidenbaum, and Amedi 2015) ont proposé un dispositif se tenant en main, de la taille d'une télécommande, afin de signaler les obstacles jusqu'à cinq mètres. Ce dispositif, « EyeCane », s'utilise comme une canne pour scanner l'environnement, et envoie des stimuli tactiles à l'utilisateur en cas de détection.

<sup>16</sup> [http://bdml.stanford.edu/twiki/pub/Haptics/VibrationImplementation/C2\\_tactor.pdf](http://bdml.stanford.edu/twiki/pub/Haptics/VibrationImplementation/C2_tactor.pdf)



Figure 2-6 : Illustration du TéléTact II (<https://perso.limsi.fr/bellik/TeleTact/TeleTact.htm>)

Dans ces premiers exemples de techniques d'interaction tangible, la modalité d'interaction tangible se base sur la manipulation d'objets « améliorés » par l'ajout d'effecteurs ou de capteurs. Dans un contexte de transmission d'information(s), telles que des directions, des distances ou plus génériquement, des grandeurs physiques, la manipulation de ces objets tangibles implique une recherche active de l'information, tout en percevant continuellement un feedback (tactile ou auditif).

#### **UN AUTRE DISPOSITIF TANGIBLE : LE VOLANT**

Dans un contexte de technologie d'assistance, (Sucu and Folmer 2013) ont utilisé un volant (Logitech G27 Racing Wheel), sur lequel a été ajouté deux autres objets tangibles (Sony Playstation Move), dont les vibrations indiquaient à l'utilisateur la manière de corriger sa trajectoire. (Sucu and Folmer 2014) adaptèrent leur système afin que même un aveugle, dans une voiture autonome<sup>17</sup>, puisse corriger la trajectoire du véhicule en cas de panne du pilote automatique. Ils nommèrent leur expérience « *The Blind Driver Challenge* ». Nous précisons que l'expérience a été réalisée sans risque, avec un simulateur. Cet exemple illustre donc l'utilisation du tangible tactile dans un contexte très concret, où l'information est requise en temps réel et dépend de la position d'un objet tiers (la voiture).

#### **CONCLUSION DE LA MODALITE TANGIBLE**

La transmission d'informations via des dispositifs tangibles, pour des utilisations non visuelles, est souvent synchronisée avec le mouvement du corps : orientation, position, localisation relative au corps ou à la partie du corps. Mais l'utilisateur se concentre, en premier, sur la manipulation de l'objet tangible. De plus, l'interaction tangible est souvent utilisée dans des situations, où la transmission des informations doit s'effectuer en temps réel, comme par exemple le fait de scanner son environnement à la recherche d'obstacle ou

<sup>17</sup> Une voiture autonome est une voiture capable de rouler automatiquement et en toute autonomie dans le trafic réel et sur une infrastructure non spécifique sans l'intervention d'un être humain.

d'appuyer sur un bouton au bon moment (situations que l'on retrouve beaucoup dans les jeux vidéo). Enfin, la manipulation d'un objet tangible améliore potentiellement l'immersion. C'est-à-dire que cela facilite la représentation mentale de l'utilisateur dans son interaction entre lui et la machine, voir même, cela permet de créer un sentiment de présence dans un environnement virtuel (ex. des jeux vidéo). C'est aussi l'un des avantages de la manipulation directe d'une interface (Hutchins, Hollan, and Norman 1985), c'est-à-dire de manipuler un objet tangible qui a la même représentation dans le système : par exemple, je bouge un slider de 3 mm, et dans le système, la représentation virtuelle du slider bouge de 3 mm. Je tourne mon volant tangible de 45° et la voiture virtuelle, que je conduis, incline ses roues à 45°.

### 2.1.5 Apport du tactile aux situations d'interaction non visuelles.

Avec des dispositifs tactiles, nous avons vu qu'il était possible de transmettre de l'information selon trois composantes principales : haptique, vibrotactile et tangible (voir Figure 2-7).

Les informations transmises peuvent avoir plusieurs degrés de précision et un « mapping<sup>18</sup> » soit direct, soit indirect. Les mappings directs regroupent les interactions utilisant nos sens de manière écologique (« au plus proche du réel »). Voici quelques exemples de mapping direct : une vibration à gauche pour transmettre l'information de tourner à gauche ; un bras articulé bloque notre mouvement pour nous indiquer un mur. Les mappings indirects regroupent les interactions nécessitant un traitement plus complexe de l'information : la perception du stimulus est codée. Voici quelques exemples de mappings indirects : un coude en flexion d'un angle de 30° est une commande de zoom à 30% ; le message vibrant (tacton) de trois courtes pulsations, sur trois vibreurs disposés en triangle autour du poignet, signale un danger. De plus, la précision diffère selon la tâche et le dispositif : OUI/NON (détection d'obstacle, alarme), quelques valeurs (8 orientations, 9 angles du coude), une grande variété de messages (12-21 notifications) ou enfin une précision infinie de l'information (localisation, orientation 360°, guidage, distance, texture, force).

Concernant la modalité haptique, nous avons vu que la manipulation active d'objets améliore la reconnaissance de ceux-ci (**percevoir pendant l'interaction**). Nous avons vu

<sup>18</sup> Mapping : (Informatique) Plus génériquement, mise en correspondance : toute opération établissant une correspondance entre les éléments d'un ensemble de données et ceux d'un second (changement de système de coordonnées, assignation de fonctionnalités à des modules, association de champs textuels avec une ontologie, etc.). <https://fr.wiktionary.org/wiki/mapping>

également que l'atteinte à l'aveugle était une alternative efficace à la réponse orale. La modalité vibrotactile a, quant à elle, l'avantage d'être « hand-free » (main libre) et non-intrusive. De plus, les dispositifs vibrotactiles peuvent être facilement conçus et programmés. La modalité tangible, enfin, repose sur des objets de la vie « quotidienne », qui, après transformation, peuvent être utilisés comme des dispositifs d'assistance, et dont leur manipulation permet l'interaction.

Nous pouvons conclure que l'aspect dynamique des interactions tactiles améliore l'efficacité des techniques, car les utilisateurs peuvent facilement « mapper » une perception (longueur, direction, forme, texture, angle, durée, effort, position, résistance, tacton) à une valeur. Toutefois, dans le cas où l'interaction ne peut être dynamique, une phase d'apprentissage est nécessaire, afin de créer un « mapping » ou plutôt un référentiel sur lequel les utilisateurs se baseront pour estimer, au mieux, la valeur correspondante au stimulus. Nous avons d'ailleurs, étudié cet effet dans notre première étude (voir Première étude : techniques d'interaction non visuelle pour estimer une valeur).

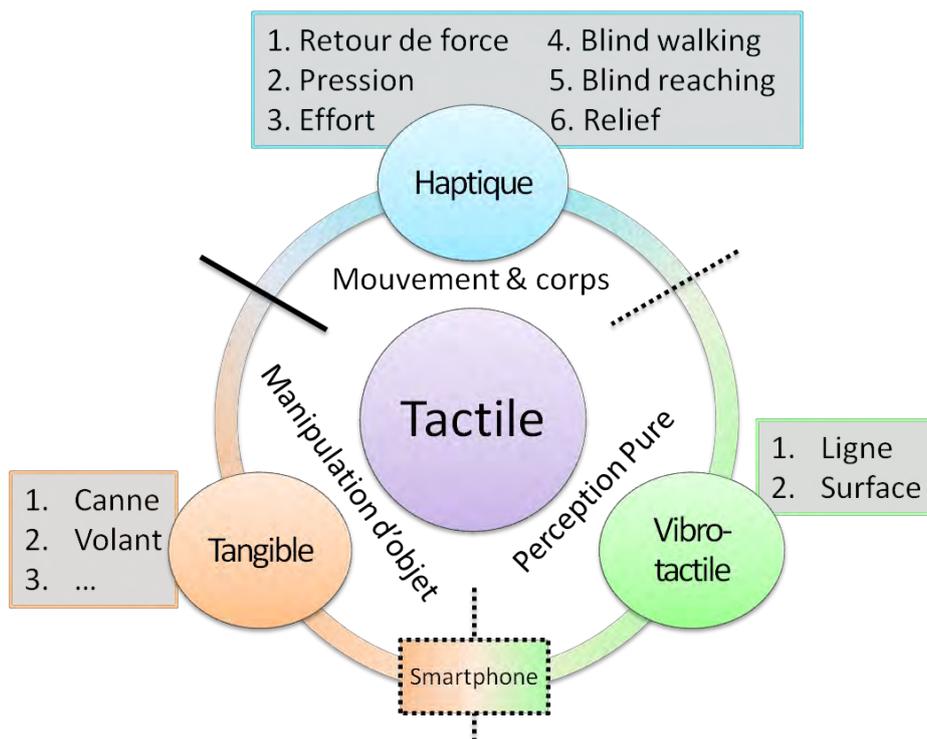


Figure 2-7 : Schéma récapitulatif des trois modalités tactile : haptique, vibrotactile et tangible ; Au centre, leur caractéristique principale les distinguant entre elles. La barre noire indique une séparation **importante** des modalités et les barres en pointillées indiquent que la séparation est moins nette. Pour rappel, « blind walking » signifie « marche à l'aveugle » et « blind reaching » signifie « atteindre à l'aveugle ».

Nous allons maintenant explorer les bases et apports de la perception auditive pour la transmission d'informations.

## 2.2 Modalités auditives

Ce que nous appelons modalités auditives au sens large regroupent : les stimuli auditifs, les modes de communication nécessitant le sens auditif et les interactions auditives. Nous présentons la grande famille des modalités auditives dans les sections suivantes et dans cet ordre.

### 2.2.1 Stimulus auditif

Dans notre contexte, un stimulus auditif est un son, c'est-à-dire une vibration mécanique d'un fluide (ici, l'air), qui se propage sous forme d'ondes longitudinales, grâce à la déformation élastique de ce fluide. A l'instar des vibrations « tactiles », les vibrations « auditives » ont une source, une certaine durée, fréquence et amplitude (ici, intensité). Un son composé peut avoir plusieurs ondes sonores qui « s'entremêlent » ; ce qui donne le timbre d'une note d'un instrument par exemple. La Figure 2-8 expose notre représentation des dimensions physiques et des propriétés d'un son dans le cadre de cette thèse.

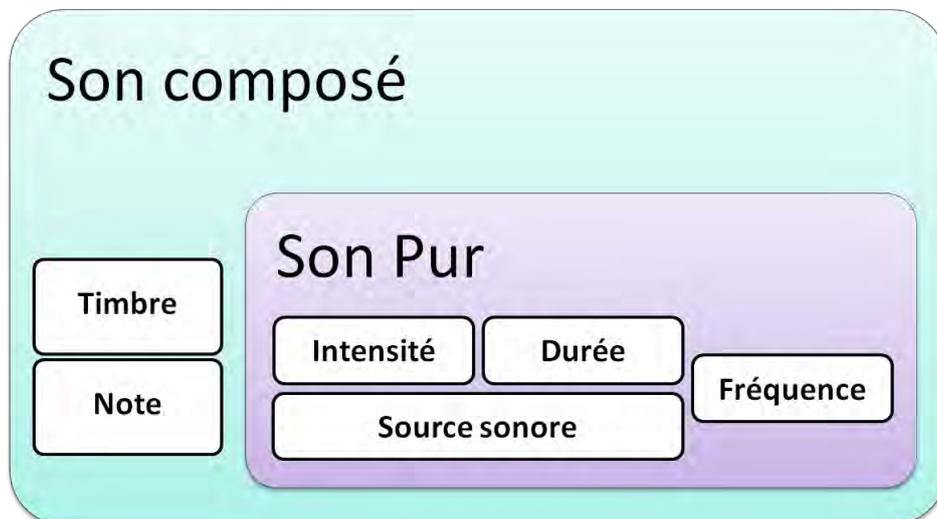


Figure 2-8 : Dimensions physiques et propriétés d'un son composé et d'un son pur.

### 2.2.2 Perception chez l'Homme : modalité sensorielle auditive

Le sens de l'ouïe permet la perception des propriétés des sons composés et purs, à savoir, l'intensité, la fréquence, la durée, la source sonore, le timbre et la note. De plus, un Homme est capable, grâce à ces deux oreilles, de discriminer les différences interaurales d'intensité (IID) et de temps (ITD). Viennent ensuite les propriétés « musicales » telles que la mélodie, le rythme et le tempo. Nous étudierons brièvement la parole, avec le champ auditif humain

et quelques causes de son altération. Une section relative aux « mappings sonores », c'est-à-dire des associations entre un son et une information, illustrera différentes formes de sonification de l'information.

**L'APPAREIL AUDITIF**

Tout d'abord, voici un rappel de la réception d'un son par notre appareil auditif. Il y a trois grandes parties de notre appareil auditif permettant de percevoir un son :

- l'oreille externe qui permet de capter et de concentrer les ondes sonores ;
- l'oreille moyenne qui permet de propager les vibrations jusqu'à l'oreille interne. Elle permet également de réguler la différence de pression entre l'extérieur et le tympan, notamment lorsque la pression sonore (grande intensité) ou environnementale (ex. à trois mètres sous l'eau) est trop forte ;
- l'oreille interne qui permet de convertir une onde sonore en stimulus.

Dans les sections suivantes, nous résumerons l'essentiel de la perception humaine pour chacune des propriétés du son.

**2.2.2.1 Les perceptions d'un son pur**

**L'INTENSITE**

Un son est perceptible à partir d'une intensité minimale de 20 micropascals (soit 0 dB, pression de référence dans l'air) et n'a pas d'intensité maximale. Cependant, l'intensité maximale audible, avant de causer des douleurs et potentiellement des lésions, se situe vers 2 pascals (soit >100dB<sup>19</sup>, voir Figure 2-9).

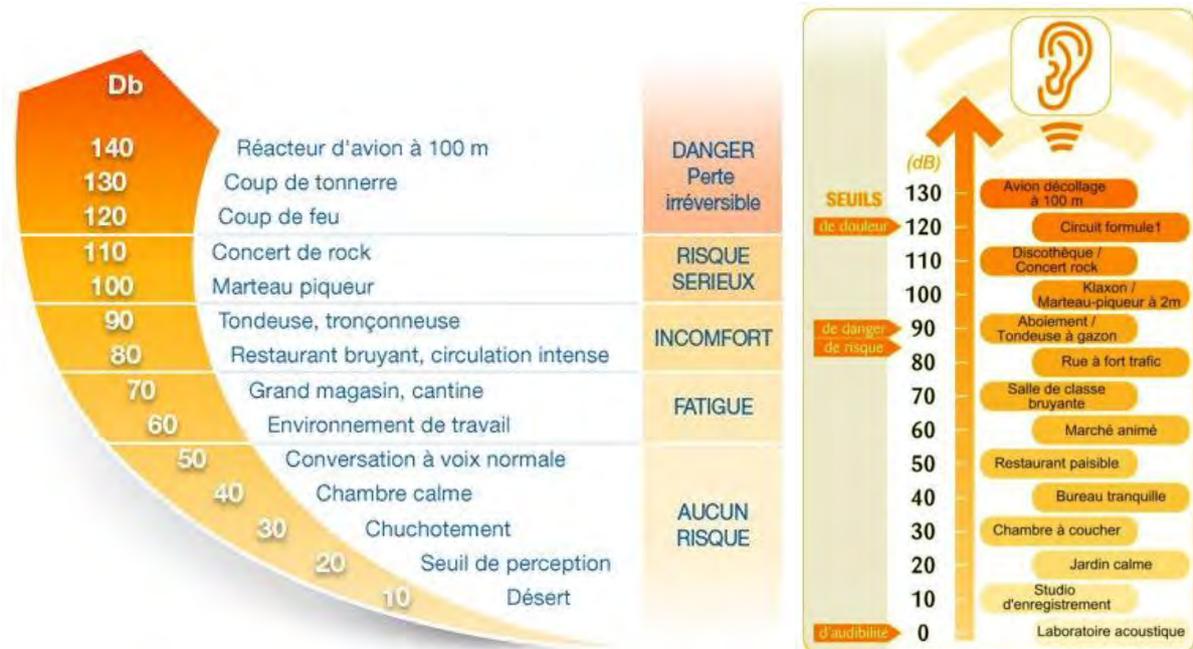


Figure 2-9 : niveaux sonores (en dB) illustrés par des environnements sonores et les seuils de dangerosité, issue de [audioson.wordpress.com](http://audioson.wordpress.com).

<sup>19</sup> Voir Niveau de pression acoustique : dB SPL, la norme ANSI S1.1-1994 et [https://fr.wikipedia.org/wiki/Pression\\_acoustique#Niveau\\_de\\_pression\\_acoustique:\\_dB\\_SPL](https://fr.wikipedia.org/wiki/Pression_acoustique#Niveau_de_pression_acoustique:_dB_SPL).

**LA FREQUENCE**

Un son est perceptible pour l'oreille humaine dans des fréquences comprises entre 16 Hertz (son grave) et 20 000 Hertz (son aigu). En deçà, ce sont les infrasons et au-delà ce sont les ultrasons. D'après les travaux de (Zwicker 1975), une différence entre deux fréquences (sons purs) est perçue, donc discriminable, dès que l'écart est de 3,5 Hz. Cependant, cette JND (*just noticeable difference*) est constante à 3,5 Hz tant que les fréquences sont inférieures à 500 Hz. Au-delà, la JND vaut 7% de la fréquence du son. Par exemple, la JND d'un son pur de 600 Hz est de 42 Hz ( $600 \times 0,07$ ). Il faut donc émettre un son de 558 Hz ou de 642 Hz pour qu'il soit perçu comme différent du son de 600 Hz.

**LA DUREE**

D'après les travaux de (Fraisse 1984), l'Homme est capable de percevoir un son (et toutes ses propriétés) si celui-ci a une durée d'au moins 0,1 seconde. Un son d'une durée inférieure ( $<0,1$  s) est perçu comme un bruit (*son confus*). Enfin, la perception d'un silence entre deux sons nécessite une durée minimale de 0,01 seconde. Quant à l'estimation de la durée d'un son, (Fraisse 1984) nous apprend que l'Homme estime avec précision les durées de sons comprises entre deux et trois secondes avec une limite maximale de cinq secondes. Ainsi, pour transmettre des informations avec un son, il suffit que celui-ci ait une durée entre 0.1s et 5s, en deçà de 0,1s le son est confus et une durée au-delà de 5s n'est pas nécessaire, car l'utilisateur a déjà perçu toutes les propriétés du son ; de plus, pour bien marquer la séparation entre deux sons, différents ou non, il faut un silence d'une durée minimale de 0,01s (voir Figure 2-10).

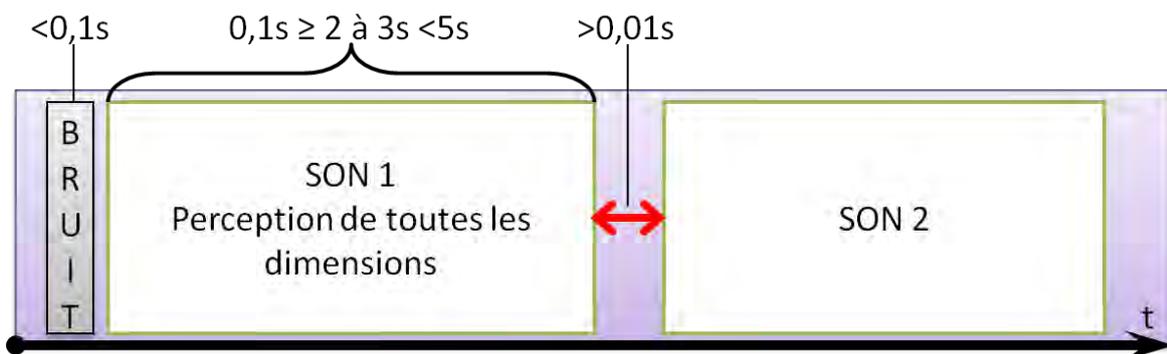


Figure 2-10 : perception et durée d'un son ; la distinction entre deux sons.

**LA DIFFERENCE INTERAURALE D'INTENSITE (IID)**

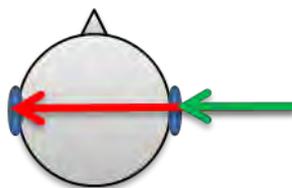
La perception d'une différence d'intensité d'une onde sonore, entre les deux oreilles, dépend de nombreux facteurs : taille de la tête, intensité et fréquence initiales de la source sonore, l'environnement (volume et espace, réverbération et surface, pression de l'air), etc.

Toutefois, cette dimension nous est très utile pour localiser dans l'espace, une source sonore, en dépit de résultats extrêmement variables. Par exemple à 4 000 Hz, les différences d'intensité peuvent atteindre, selon les individus, 15 à 25 dB. Néanmoins, cet indice peut être utilisé comme première approximation d'une localisation, à partir de fréquence de 1 500 Hz (Letowski and Letowski 2012).

#### **LA DIFFERENCE INTERAURALE DE TEMPS (ITD)**

A l'instar de l'IID, la perception d'une différence de temps d'arrivée d'une onde sonore entre les deux oreilles dépend de nombreux facteurs, et nous est également utile pour localiser dans l'espace, une source sonore. Toutefois, l'ITD est plus précise que l'IID dans des fréquences inférieures à 1 500 Hertz, alors que l'IID est plus précise que l'ITD à partir de 1 500 Hz.

Ainsi, lorsque la source sonore est aux azimuts 0° ou 180° (devant, derrière), l'ITD vaudra 0 ms. A l'opposé, l'ITD sera maximale lorsque la source sera aux azimuts 270° et 90° (à gauche, à droite) avec une ITD d'environ 0,7 ms (Voinier and Briolloe 1992). Cela dépend bien sûr de la taille de la tête des sujets : pour rappel, la vitesse du son dans l'air est de 340m/s et une tête humaine moyenne fait 23,8 cm de large ; après un rapide calcul nous obtenons  $0,238/340 = 0,0007s$ , soit 0,7 ms (voir Figure 2-11).



**Figure 2-11 : illustration d'une différence interaurale de temps (ITD) avec une source sonore à 90° sur l'azimut (i.e. à droite, flèche verte) ; la différence est symbolisée rapidement par la flèche rouge (qui traverse la tête) ; l'ITD est de 0,7 ms.**

#### **2.2.2.2 La perception d'une source sonore**

##### **L'AZIMUT**

L'azimut est une direction égocentrée sur un plan horizontal dont l'origine est en face de nous (0°). (Mills 1958) mena une expérimentation « On the Minimum Audible Angle » afin de déterminer chez l'Homme, quel est le seuil de perception d'un changement d'azimut d'une source sonore en fonction de l'azimut du premier stimulus. La mesure observée était la réponse orale du sujet confirmant ou infirmant la détection d'une différence d'angle entre deux stimuli auditifs pour une même fréquence sur le plan horizontale (azimut). Les facteurs sont la position de la première source sonore (0°-90°), la fréquence des stimuli (250 Hz-10 000 Hz) et l'angle (l'arc) entre les deux stimuli (1°-30°). Les résultats principaux sont :

la direction de la source sonore est déterminée de manière optimale, à une fréquence inférieure à 1 400 Hertz, par la différence interaurale de temps (ITD) et pour des fréquences supérieures à 1 400 Hertz par la différence interaurale d'intensité (IID). De plus, (Mills 1958) nous assure que dans des conditions optimales (chambre anéchoïque<sup>20</sup>), le seuil minimal de discrimination pour l'ITD serait de 0,01 ms et de 0,5 dB pour l'IID. Notons également que (Mills 1958) rencontra des problèmes de discrimination pour un azimut proche de 270° ou 90° (gauche, droite), problèmes déjà connus dans la littérature (Bourdon 1914). Toutefois, il est également connu que l'être humain effectue, instinctivement, de petits mouvements de tête, qui lui permettent de créer de nouveaux indices binauraux et monauraux, et ainsi de préciser la position de la source sonore. Cela permet notamment de lever certaines ambiguïtés de localisation, notamment, dans la zone de confusion (azimut à 270° et 90°).

Cependant, (Hartmann and Raked 1989), trente ans après Mills, publièrent une « plainte » contre la méthode expérimentale de (Mills 1958), notamment sur l'interprétation statistique des résultats et la non-prise en compte de biais expérimentaux liés à des stratégies prédictives des sujets. Ainsi, (Hartmann and Raked 1989) conclurent que l'expérience de (Mills 1958) serait plus optimale si l'on considère la tâche comme une identification absolue d'un azimut plutôt qu'une tâche de discrimination d'angle.

Ainsi, la perception de l'azimut serait optimale pour des stimuli en face de soi puis de plus en plus erratique lorsque que la source sonore est proche des azimuts 90° et 270° et de moins en moins erratique lorsque que la source en est éloignée. Enfin, la tendance naturelle de l'Homme à confondre l'arrière avec l'avant (voir (Bourdon 1914)), rend moins efficace la perception d'une source sonore située derrière nous. Notons enfin que notre perception de l'azimut est efficace, pour des basses fréquences avec la différence interaurale de temps, et pour des hautes fréquences avec la différence interaurale d'intensité.

### **L'ELEVATION**

L'élévation est une direction égocentrée sur un plan vertical dont l'origine (0°) est le niveau des oreilles. Dans le plan vertical, les principaux indices utilisés sont des indices monauraux, ils proviennent de la modification d'un son par le torse, la tête et l'oreille externe d'un auditeur. Néanmoins, l'appréciation de la source dans un plan vertical reste cependant plus imprécise que dans le plan horizontal.

---

<sup>20</sup> Une salle d'expérimentation dont les parois absorbent les ondes sonores (ou chambre sourde).

**LE ROLE DE LA TÊTE ET DU PAVILLON DE L'OREILLE**

Les ITDs, les IIDs et les mouvements de tête ne suffisent pas à expliquer complètement nos capacités de localisation et notamment, en élévation. L'ensemble de la tête et du pavillon réalise un filtrage complexe, dont l'effet est notable de 500 à 16 000 Hz. Ce filtrage varie avec la direction d'incidence du son. On en mesure donc les fonctions de transfert, ou HRTF (« Head-Related Transfer Functions ») en fonction de la direction d'incidence du son (azimut et élévation)<sup>21</sup>.

Cette section explique pourquoi il est facile de percevoir une source sonore avec un casque stéréo plutôt qu'avec des enceintes. En effet, la perception d'une source sonore dépend de l'environnement acoustique de l'utilisateur et de sa position par rapport à la source. Dans le cas d'une simulation, les paramètres, liés à l'environnement acoustique et à la position de l'utilisateur sont plus facilement calculables avec un casque stéréo sur les oreilles de l'utilisateur qu'avec des enceintes (voir Artéfacts technologiques pour produire des stimuli auditifs). C'est pourquoi, nous avons utilisé un casque stéréo, dans notre première et deuxième étude, comme dispositif pour transmettre des informations avec la modalité auditive et notamment dans le cas de sons spatialisés.

**2.2.2.3 Perceptions d'un son composé****TIMBRE**

Le timbre est le caractère qui nous permet de différencier deux sons, provoquant la même sensation d'intensité et de même hauteur tonale (fréquence). Le timbre d'un son est essentiellement lié à sa composition spectrale (fréquences), mais également à son évolution dans le temps (forme de l'enveloppe spectrale, l'attaque des sons...). La voix de chaque personne a son propre timbre ; il en est de même pour chaque instrument de musique. Nous verrons quelques exemples d'IHM dans la section 2.2.3 Les mappings sonores.

**NOTE**

La note est une combinaison d'une hauteur tonale (fréquence) et d'une durée. Cette note a également un timbre et une intensité.

Nous utilisons le timbre et la note dans une de nos techniques d'interaction (« Mélodie ») dans notre deuxième étude.

<sup>21</sup> [Cours de psychoacoustique](#) (Pr. Alain Goyé, 2002)

#### 2.2.2.4 Perceptions « musicales »

---

L'homme perçoit également d'autres dimensions souvent regroupées dans le domaine de la musique<sup>22</sup>.

##### **RYTHME**

C'est l'arrangement de la durée des notes et des silences dans le temps. La perception d'un même morceau est différente selon le rythme utilisé. Par exemple, un rythme saccadé peut marquer la précipitation, une situation haletante. Au contraire, un rythme régulier peut donner une impression de calme.

##### **TEMPO**

C'est la vitesse à laquelle est exécutée une œuvre musicale. Une réalisation lente peut suggérer un sentiment de douceur, de lourdeur, de gravité, un thème poétique... A l'inverse, la rapidité peut traduire la peur, le dynamisme, la gaité...

##### **MELODIE**

Une mélodie est une succession de notes dont résulte un air, harmonieux, agréable à entendre ou triste (mélodie mélancolique).

Ces perceptions sont parfois utilisées dans des techniques d'interaction pour transmettre une information (voir par exemple notre section : Les mappings sonores). De plus, dans notre deuxième étude, nous utilisons la « Mélodie » pour transmettre une grande quantité d'informations (voir Deuxième étude : interactions non visuelles pour faciliter le rendu d'une séquence de couples direction-distance).

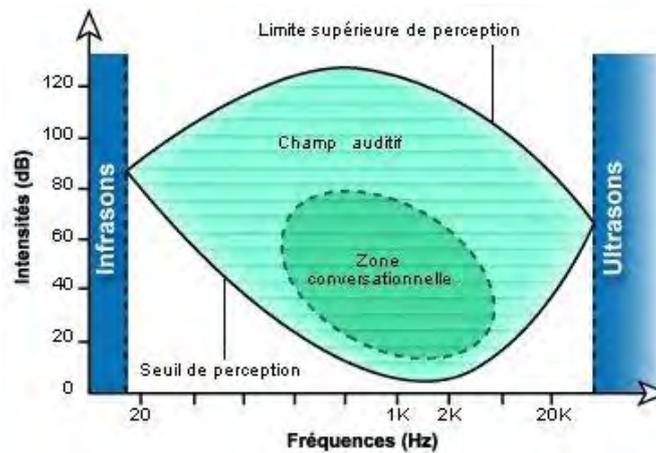
#### 2.2.2.5 Parole et champ auditif humain

---

De nombreux travaux sur l'audition humaine ont permis d'aboutir à une représentation graphique de la capacité humaine à percevoir un son, en fonction de sa fréquence et de son intensité (voir l'étude bibliographique de (Chatillon, 2006)). Ainsi, la voix humaine est bien perçue lorsqu'elle se situe dans la zone conversationnelle (voir Figure 2-12).

---

<sup>22</sup> <http://www.cochlea.org/entendre/perception-de-la-musique> ; Rémy Pujol (PR émérite, CNRS), Stéphan Blatrix, Stéphane Le Merre, Benjamin Chaix. (voir aussi : [www.cochlea.eu](http://www.cochlea.eu))



**Figure 2-12 Champ audible chez l'Homme.** Ce graphique montre l'intensité nécessaire à la perception (en dB) en fonction des fréquences (en Hz). Aux fréquences audibles, l'aire en vert clair délimitée par le seuil d'audition et le seuil d'inconfort (lignes noire sup. et inf.) définit le champ auditif, domaine de sensation auditive ou aire d'audition. L'aire en vert foncé est la zone conversationnelle et définit les sons utilisés pour la communication par la voix humaine (Unité INSERM en 1999 : « Promenade autour de la cochlée »).

#### **DEGRADATION AUDITIVE LIEE A L'AGE : LA PRESBYACOUSIE**

La presbyacousie est un phénomène plus ou moins marqué selon les individus et qui peut résulter du vieillissement. Elle est définie comme une perte progressive de l'audition, liée à l'âge, bilatérale et symétrique, surtout dans les fréquences élevées. La presbyacousie se traduit par une perte de la perception des sons aigus, lesquels permettent une bonne perception des mots. La norme (ISO 7029:2000 2000) indique la modification (diminution) du seuil de perception due uniquement à l'âge, voici les plus importants :

- une des premières conséquences est la perte de la capacité pour les jeunes adultes d'entendre les fréquences au-dessus de 15-16 kHz ;
- la détection de sons très aigus (4 kHz à 8 kHz) devient plus difficile, et la perception de la parole est affectée, surtout pour les consonnes sifflantes (S, Z, Ch) et les fricatives (F, V). Les deux oreilles ont tendance à être affectées.

L'audition peut être également dégradée à causes de malformations congénitales.

#### **2.2.2.6 Conclusion de la perception auditive**

Nous avons vu les limites de perception pour les propriétés d'un son, à savoir : l'intensité, la fréquence, la durée, la position de la source sonore, son timbre etc. Ces propriétés peuvent être utilisées pour estimer une tonie<sup>23</sup>, sonie<sup>24</sup>, chronie<sup>25</sup> etc. Ou alors, pour encoder des messages auditifs abstraits. Dans la section suivante, nous allons aborder les « mapping sonores » qui sont construits sur des messages abstraits auditifs.

<sup>23</sup> Tonie : grandeur « linéaire » subjective associée à la fréquence dans le domaine psychoacoustique.

<sup>24</sup> Sonie : grandeur « linéaire » subjective associée à l'intensité dans le domaine psychoacoustique.

<sup>25</sup> Chronie : grandeur « linéaire » subjective associée à la durée dans le domaine psychoacoustique.

### 2.2.3 Les mappings sonores

L'expression « Mapping sonore » définit le lien entre une représentation sonore et une information. Les mappings sonores peuvent être classés selon plusieurs catégories. (Csapó and Wersényi 2013) ont décrit plusieurs de ces catégories dans le contexte des interfaces Homme-Machine. Nous décrivons brièvement les catégories de (Csapó and Wersényi 2013) et nous nous attarderons sur leurs différences.

#### **ICONES SONORES**

Les icônes sonores sont des métaphores non verbales (sonores) liant un son à un événement (ex. un feedback (retour d'information), une alarme, etc.). Une icône sonore est un son court, facile à interpréter et à apprendre (ex. son d'un système d'exploitation, bip de chantier, etc.). D'après (Gaver 1986), une icône sonore est une représentation directe ou métaphorique d'un son naturel produit par un animal ou un objet (ex. « meuh » pour une vache ou « froissement de papier » pour l'action de supprimer un document numérique dans la corbeille virtuelle) (Blattner, Sumikawa, and Greenberg 1989; Gaver 1986, 1988; Gygi and Shafiro 2009).

#### **EARCONS**

Les earcons sont des représentations abstraites non verbales liant un son à un événement (ex. une alerte, une notification, etc.). Toutefois, à la différence d'une icône sonore dont la métaphore est très connue (ex. bip d'erreur de pc) et facilement interprétable, un earcon est une relation plus difficile à interpréter, et nécessite une phase d'apprentissage conséquente, avant toute utilisation (Blattner *et al.* 1989). Ainsi, un earcon est une représentation sonore abstraite liant un son à un événement, nécessitant un apprentissage, et n'ayant que peu de notoriété dans les populations non-initiées. Par exemple, (Leplâtre and Brewster 1998) ont utilisé des earcons afin de naviguer « plus facilement » dans des menus hiérarchiques de téléphone (fin des années 90).

#### **SPEARCONS**

Les spearcons sont basées sur de la parole accélérée à tel point, qu'il est impossible d'en percevoir les mots (Walker, Nance, and Lindsay 2006). Le principe est de conserver le *pitch* originel (hauteur tonal, fréquence) et de laisser l'utilisateur interpréter le sens du spearcons. Quelques exemples peuvent être entendus en ligne<sup>26</sup>.

#### **ÉMOTICONES AUDITIVES**

<sup>26</sup> <http://sonify.psych.gatech.edu/research/auditorymenus/index.html>.

Une émoticône auditive est utilisée comme une émoticône graphique (☺), c'est un son émotionnel (Froehlich and Hammer 2004). Quelques exemples peuvent être écoutés en ligne également<sup>27</sup>.

#### **SONIFICATION : ALERTES ET AIDE A LA NAVIGATION**

La sonification est définie par (Barrass and Vickers 2011; Kramer 1993) comme : « une utilisation audio non verbale pour transmettre de l'information ». Voici quelques exemples de sonification : les alertes dans les cockpits d'avions civils (Patterson and Civil Aviation Authority 1982) ou comme technologie d'assistance pour les déficients-visuels pour faire de la détection d'obstacle et du guidage (Csapó and Wersényi 2013).

#### **CONCLUSION SUR LES MAPPINGS SONORES**

Précédemment, nous avons vu les limites des propriétés du son : intensité, fréquence, durée, différences interaurales de temps et d'intensité ; et les limites de la parole : champ auditif, dégénérescences liées à l'âge ou innées. Nous avons également vu les composantes et les dimensions composées d'un son : rythme, tempo, timbre, note, mélodie, azimut et élévation. Avec ces propriétés, il est possible de construire des « mappings sonores » afin de lier une information à un son. Toutefois, ces mappings nécessitent d'être appris par les utilisateurs avant toute utilisation. Enfin, ces mappings offrent de multiples façons pour encoder une information. En les combinant, notamment grâce à la sonification, les solutions sont encore multipliées.

Nous illustrons dans la suite, comment des interactions auditives permettent une transmission d'informations, et donc, sont des alternatives aux solutions déjà présentées.

### **2.2.4 Artéfacts technologiques pour produire des stimuli auditifs**

Il existe des technologies capables de produire des stimuli nous faisant percevoir des sons :

- L'écouteur ou oreillette : dispositif contenant, dans la majorité des cas, un transducteur électrodynamique (en langage courant, une sorte de hautparleur miniature) que l'on place dans l'oreille ou que l'on accroche à celle-ci pour l'écoute de sons (voir Figure 2-13).

<sup>27</sup> <http://www.universal-soundbank.com/pleurs.htm> ; [/expressions.htm](#) ; [/onomatopees.htm](#).



Figure 2-13 : oreillette monophonique Bluetooth (sans fil)(source : <https://fr.wikipedia.org/wiki/écouteur>)

- Le casque audio : composé de deux écouteurs, un pour chaque oreille. Les casques stéréo permettent de spatialiser les sons (voir Figure 2-14).



Figure 2-14 : Un casque audio des années 1970(source : [https://fr.wikipedia.org/wiki/Casque\\_audio](https://fr.wikipedia.org/wiki/Casque_audio))

- Le hautparleur : transducteur électroacoustique destiné à produire des sons à partir d'un signal électrique. Il est en cela l'inverse du microphone (premier brevet en 1877 par Werner von Siemens). A la différence d'un casque stéréo, il est très difficile de spatialiser un son avec des hautparleurs, du fait qu'ils sont souvent loin de l'oreille, et aussi, que la position des utilisateurs, dans la pièce, change leur perception du son. Pour ce faire, il faut un système qui calcule la spatialisation du son, en fonction de la réflexivité de l'environnement, et de la position en temps réel de l'utilisateur. Alors qu'avec un casque audio ces deux paramètres sont fixes (voir Figure 2-15).

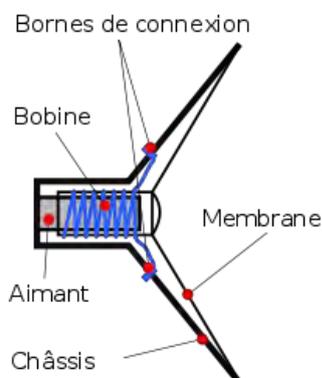


Figure 2-15 à gauche : schéma de coupe d'un hautparleur ; à droite : un hautparleur. (source : <https://fr.wikipedia.org/wiki/Haut-parleur>)

## 2.2.5 Mise en œuvre de la modalité auditive dans des situations interactives

Dans cette section, nous synthétisons des mises en œuvre concrètes de l'auditif dans les situations interactives. Nous avons choisi de considérer trois catégories de situations interactives :

- Les IHM auditives cantonnées à une tâche d'estimation de valeur(s) de type stimulus-réponse.
- Les IHM d'assistance pour du guidage ou de la prévention.
- Les IHM auditives utilisant un mapping sonore pour des tâches de navigation, de sélection, de manipulation ou notification.

### 2.2.5.1 Interfaces auditives pour l'estimation de valeur

Les interfaces auditives, pour l'estimation, transmettent des informations selon les propriétés du son. L'efficacité des propriétés du son à transmettre des informations varie selon leur type et les caractéristiques physiques de l'environnement.

Dans le contexte spécifique d'estimation de distance, (Coleman 1963) a identifié cinq facteurs essentiels dans la perception d'une distance : l'intensité, la fréquence, la réverbération, les différences binaurales de temps et d'intensité. Puis, il s'ensuit une longue discussion sur leur efficacité et l'environnement de contrôle. Par exemple, (Mershon and King 1975) étudièrent, en deux expériences, l'impact d'un environnement anéchoïque<sup>28</sup> (ou non) sur l'estimation d'une distance, par les propriétés d'intensité et de réverbération. Il en est ressorti que l'intensité est une bonne propriété, pour l'estimation de distance. Cependant, dans le cas d'une estimation absolue, l'intensité initiale de la source sonore doit être connue par l'utilisateur, car sans ce repère, point de différence d'intensité et donc point d'estimation possible. De plus, ces études recommandent une différence de vingt décibels pour une perception optimale de la distance. (Mershon and King 1975) observèrent également que plus l'effet de réverbération est important, plus la perception de la distance sera grande. Cependant, l'utilisation de cette propriété nécessite un contrôle de l'acoustique du milieu, ce qui est rarement le cas en situation réelle (et ordinaire). Dans ce même contexte, (Talbot and Cowan 2009) ont réalisé une étude basée sur la fréquence, l'intensité et le rythme. Leurs conclusions sont les suivantes : la fréquence est la propriété la moins effective des trois ; la propriété « rythme » est plus effective que la propriété « fréquence » ; leur interaction « écologique » est la meilleure. Certes, mais quels sont ses paramètres ? (Talbot and Cowan 2009) ont décrit leur interaction écologique ainsi : plus l'intensité décroît plus la distance augmente, elle est bornée entre 60 et 80 dB et est

<sup>28</sup> Une salle d'expérimentation dont les parois absorbent les ondes sonores (ou chambre sourde)

composée de sept niveaux ; de plus, la fréquence est modifiée selon la distance par un filtre passe bas, afin de simuler l'absorption du son par l'air. Dans une deuxième expérience, ils combinèrent leurs deux meilleures techniques (rythme et écologique) et conclurent que cette combinaison était redondante et à peine meilleure que l'interaction écologique.

Dans un autre ordre d'idée, les travaux de (Meijer 1992), décrivent la conception et l'évaluation d'un algorithme de rendu d'« image sonore ». Ainsi, cet algorithme, baptisé « The vOICe » par (Auvray, Hanneton, and O'Regan 2007), est basé sur trois propriétés du son, liées à trois propriétés d'une image : fréquence, durée, intensité (resp. ligne de pixels, colonne de pixel, nuance de gris). Dans leur expérimentation, les utilisateurs devaient « estimer » une image avec du son. Dans les travaux de (Capelle *et al.* 1998), les auteurs proposent un prototype de rendu d'« image sonore » basé notamment sur la fréquence de son, y figure également une comparaison entre leur prototype (bonne résolution du centre de l'image, i.e. un zoom) et celui de (Meijer 1992) (résolution totale satisfaisante de 64x64).

Dans un contexte de transfert d'information, (Oh, Kane, and Findlater 2013) ont utilisé les modalités sonification et parole, afin de concevoir des techniques d'interaction permettant de dessiner à l'aveugle une figure géométrique (ovale ou rectangle). La forme était transmise par le son, plus précisément, par quatre propriétés d'un son : la fréquence avec une gamme de dix fréquences entre 246 Hz et 587 Hz ; l'intensité composée de dix niveaux ; le timbre, ou plutôt la combinaison d'un signal pur en sinusoïde et d'un second en triangle dont le « mélange » variait selon si la cible était à gauche ou à droite ; le pan, qui est la balance des hautparleurs gauche et droit. Ainsi dans une première expérience, la fréquence était utilisée pour guider sur l'axe des ordonnées, tandis que l'axe des abscisses était assigné soit à l'intensité, soit au timbre, soit au pan. Les résultats de cette étude comparative montrèrent que la combinaison « fréquence + pan » était la meilleure. Cette combinaison fut comparée, dans une seconde expérience, à une interaction basée sur la parole. Au vu des résultats de cette étude complémentaire, les auteurs conclurent que les deux modalités (sonification et parole) étaient aussi bien l'une que l'autre, et offraient des avantages complémentaires, avec toutefois une meilleure appréciation de la part des sujets pour la sonification (« fréquence + pan »). D'autres travaux conduits par (Larsen *et al.* 2013), ont comparé, dans une tâche de localisation, trois interactions basées sur le pan (ou **panning**, balance des hautparleurs gauche et droit) et une interaction basée sur la fonction de transfert relative à la tête (HRTF, i.e. caractérise les transformations apportées aux ondes

sonores par le corps d'un auditeur, qui permettent à l'être humain de localiser une source sonore, voir aussi <https://fr.wikipedia.org/wiki/Hrtf>). Les auteurs ont donc évalué des interactions reposant sur les différences interaurales de temps<sup>29</sup> ou d'intensité (panning), et sur l'effet HRTF. Leur conclusion, basée sur des résultats significatifs, recommande l'utilisation de l'effet HRTF dans des techniques d'interaction destinées à des tâches de localisation de sources sonores. De plus, leur technique avec l'HRTF était plus précise et plus rapide pour une tâche de guidage en navigation.

En résumé, il est possible de transmettre des informations avec les différentes propriétés d'un son et même plusieurs informations en simultané avec un son. Ce dernier point est important, car nous utiliserons, nous aussi, les propriétés d'un son pour transmettre des séquences de couples d'information, où chaque couple est transmis en simultané (voir Deuxième étude : interactions non visuelles pour faciliter le rendu d'une séquence de couples direction-distance).

### 2.2.5.2 Technologies d'assistance auditives

Cette catégorie regroupe les techniques d'interaction basées sur des modalités auditives, et ayant au moins une fonctionnalité parmi les trois suivantes : alarme (signal avertissant d'un danger) ; alerte (signal invitant ceux qui le perçoivent à se tenir sur leurs gardes, à faire attention à quelque chose) ; assistance dans une tâche nécessitant la vue (recherche d'objet, navigation, etc.).

#### **ALERTE ET ALARME**

Les travaux de (Edworthy, Loxley, and Dennis 1991) portèrent sur la conception de signaux d'alertes. Ils menèrent une série d'expériences durant lesquelles, la fréquence, la mélodie, l'intensité, le tempo, et le rythme ont été utilisés pour concevoir et évaluer des alertes. Leur expérience finale se concentra sur l'évaluation de treize interactions conçues à partir des résultats issus d'expériences précédentes. Leur conclusion fut que l'ensemble des propriétés du son (citées plus haut) étaient efficaces pour concevoir des signaux d'alertes. Une extension de ces travaux a été menée par (Hellier, Edworthy, and Dennis 1993). Ils ont exploré l'impact des propriétés du son dans la perception d'une urgence, i.e. une alarme. Leurs résultats montrèrent que parmi les dimensions testées, la fréquence permettait de percevoir au mieux trois niveaux d'alarme (faible-75Hz, modéré-168Hz, haut-378Hz). Le

<sup>29</sup> Différence de temps d'arrivée d'un son entre les deux oreilles, voir état de l'art La différence interaurale de temps (ITD) ou (Voinier and Brioloe 1992)

rythme (« speed ») et le tempo (« repetition unit ») ont obtenu des résultats corrects et similaires, mais la discrimination des trois niveaux d'alarme reste plus faible qu'avec la propriété fréquence. Toutefois, (Brock, Ballas, and McFarlane 2005) démontrèrent l'efficacité des propriétés rythme et tempo dans l'encodage d'alarmes. Dans un contexte d'alarme, (Fujimoto and Turk 2014) ont comparé plusieurs techniques d'interaction dans une tâche de localisation d'un danger. Leurs techniques d'interaction d'alarme reposent notamment sur celles de (Etter and Specht 2005) pour l'utilisation de mélodies spatialisées, celles de (Liljedahl *et al.* 2012; Rümelin *et al.* 2011) pour l'utilisation de la métaphore du balayage (« sweep » ou « turning gesture ») et celles de (Jones *et al.* 2008) pour le guidage par variation d'intensité et panning (balance gauche/droite des hautparleurs). Ainsi, lorsqu'un danger approchait, les utilisateurs, marchant d'un pas nonchalant tout en écoutant de la musique, percevaient un changement dans leur musique. Ce changement alertait l'utilisateur du danger imminent tout en lui transmettant la position du danger (direction et distance).

#### **ASSISTANCE DANS UNE TACHE NECESSITANT LA VUE**

La recherche d'information, par des modalités non visuelles dans un contexte de navigation, que ce soit au niveau d'une carte ou en situation écologique (proche du réel), peut s'avérer critique (ex. plan d'évacuation), importante (ex. lieu de rendez-vous) ou intellectuelle (ex. carte géopolitique). Dans ces contextes, la recherche visuelle est souvent sollicitée. Cependant, dans notre contexte de transmission d'informations, via des modalités non visuelles, nous détaillerons plusieurs techniques d'interaction capables d'assister une tâche, nécessitant d'ordinaire, la vue. Ces techniques sont bien souvent adressées, en premier lieu, aux déficients visuels. Par exemple, (Zhao *et al.* 2008) proposent une interaction utilisant les propriétés d'un son (mélodie, gamme de note) et de la parole (synthèse vocale), afin de rendre perceptible la densité de population américaine pour des non-voyants. (Kane *et al.* 2011) ont utilisé la synthèse vocale pour guider la main d'un non-voyant jusqu'à une cible sur une carte. Cette technique utilise, et répète jusqu'à la fin de la tâche, quatre mots pour le guidage : haut, bas, droite et gauche (haut-haut-haut-droite-droite-...). (Panëels *et al.* 2013) utilisèrent la parole, afin de décrire les points d'intérêt (PI) d'une rue ; l'utilisateur faisait glisser son doigt sur un Smartphone, et contrôlait ainsi l'énumération de la liste des PI. Dans un autre ordre d'idée, (Gonot, Chateau, and Emerit 2006) utilisèrent un son spatialisé, pour naviguer dans un environnement virtuel complexe (petite ville). Leurs travaux comparèrent plusieurs dimensions de la sonification (panning, différence binaurale,

HRTF, tempo et timbre<sup>30</sup>) dans deux systèmes de coordonnées différents. L'un, appelé « représentation décontextualisée », donnait la position en coordonnées polaires (direction-distance) et l'autre, appelé « représentation contextualisée », donnait la direction à suivre jusqu'au prochain croisement, ainsi que la distance (totale) qu'il restait à parcourir. De plus, la tâche de navigation était de trouver neuf cibles ; ces cibles étaient virtuelles comme des sources sonores (feu d'artifice, cloche, sirène, ambiance de récréation, etc.). Les résultats de leur étude, montrèrent que l'utilisation d'une « représentation décontextualisée » était plus chronophage lors d'un choix de direction (croisement) que la « représentation contextualisée ». Les auteurs concluent que la transmission d'information était intégrée, plus facilement, avec la représentation contextualisée.

La technologie auditive, ou plus précisément, la manipulation des propriétés du son ou de la parole, dans des techniques d'assistance, rend efficace les interactions pour des tâches d'alerte, d'alarme et d'assistance de navigation. Dans le cadre de cette thèse, nous avons utilisé la technologie auditive afin de transmettre des informations simples de type <valeur> comme une notification (cf. Première étude : techniques d'interaction non visuelle pour estimer une valeur), puis des informations complexes de type <itinéraire> pour faciliter le rendu d'une séquence de couples direction-distance (cf. Deuxième étude : interactions non visuelles pour faciliter le rendu d'une séquence de couples direction-distance).

### 2.2.5.3 Association de sons aux éléments d'interfaces

Cette section regroupe des interactions basées sur des « mapping sonores », i.e. un lien entre un objet sonore et un concept. Nous allons présenter plusieurs applications de ce type de transmission d'informations.

Il existe plusieurs types de mappings sonores. Les plus importants ont été détaillés dans la section Les mappings sonores de ce manuscrit : les mapping sonores sont des relations entre un objet sonore (son ou parole) et un concept (ex. émotion, niveau hiérarchique, objet physique ou virtuel, etc.) ; ces relations doivent s'apprendre.

Commençons par les travaux de (Garzonis, Bevan, and O'Neill 2008), qui comparèrent trois interactions basées sur des mappings sonores dans une tâche de « stimulus-réponse » sur un Smartphone (« j'entends un son, je sélectionne la bonne application »). Parmi les trois interactions, l'une était basée sur des earcons, la deuxième sur des icônes sonores et la troisième sur de la parole (synthèse vocale). Aucune phase d'apprentissage des *mappings*

<sup>30</sup> Voir section Perception chez l'Homme : modalité sensorielle auditive

de test n'a été effectuée, mais seulement une phase de familiarisation aux interactions avec des sons d'animaux (très différents des sons utilisés lors des tests). Toutefois, l'interface de test fut présentée et expliquée aux sujets, afin que les résultats ne soient pas biaisés, par un problème de recherche visuelle ou de compréhension (i.e. que tous les sujets doivent avoir la même représentation de l'interface). Les résultats montrèrent que la modalité de la parole était la plus efficace en termes de temps de réponse (3,2s) et de précision (95,9% de réussite à sélectionner la bonne application). La modalité earcons était la moins efficace (5,1s et 18%), tandis que la modalité icône sonore (4,8s et 37,4%) était légèrement plus efficace que celle des earcons. Les résultats qualitatifs (subjectifs) montrèrent que la modalité parole était préférée, facile à apprendre et à comprendre. Cependant, pour la question : « voudriez-vous l'utiliser ? (l'avoir ?) », la modalité parole obtenait un score (54%) à peine supérieur aux deux autres modalités (46,5% en moyenne). Cette expérience démontra deux résultats : le premier est que sans entraînement préalable, les mappings sonores sont difficiles à interpréter, et le deuxième est que l'utilisation de la parole est autant désirée que celle de mappings sonores non appris. Les auteurs abondent en ce sens dans leur discussion, et prennent en exemple certaines icônes sonores, qui ont eu des résultats de précision légèrement supérieurs à la parole, notamment pour la « vieille » sonnerie de téléphone associée au menu des appels. En effet, à la différence des earcons, les icônes sonores sont des mappings très connus, faciles à interpréter et ne nécessitant que très peu d'apprentissage !

Dans une tâche d'apprentissage de menus, (Palladino and Walker 2007) ont comparé la modalité earcon (mapping abstrait) et la modalité spearcon (parole compressée, seul le pitch originel est conservé). Les auteurs ont pris soin de vérifier leurs hypothèses avec deux types de liste : l'une, était composée de noms communs (océan, voiture, flute, cheval, rire, etc.) et l'autre, était composée de menu « commun » de téléphone (email, ouvrir, aide, qualité d'image, nouveau message, etc.). Leurs résultats montrèrent que les spearcons étaient plus facile à apprendre que les earcons. La modalité spearcon nécessitait en moyenne une série d'apprentissages pour les deux listes, alors que la modalité earcon nécessitait au moins deux séries pour la liste de noms communs ( $4,6 \pm 2,3$ ) et trois séries pour la liste de menu de téléphone ( $6,6 \pm 3,3$ ). La suite de cette expérience fut menée par (Dingler and Walker 2008). Cette fois, après une phase de présentation de vingt menus dans la modalité testée, les participants commençaient la phase de test : ils devaient sélectionner le bon menu parmi vingt, après avoir entendu le stimulus, tous les menus étaient présents

et les sujets pouvaient « rejouer » le stimulus autant de fois qu'ils le souhaitent, avant de prendre une décision finale. Toutefois, si les participants n'obtenaient pas 20/20 (donc aucune erreur), ils recommençaient pour la même modalité : une phase de présentation et une phase de test. Parmi les modalités testées, quatre nous sont connues (earcons, icônes sonores, spearcons et parole), et deux étaient des modalités «hybrides » combinant plusieurs des quatre solutions précédentes. Les résultats montrèrent, une fois de plus, que la modalité spearcon était la meilleure en termes d'apprentissage (nombre de phases de présentation) et de précision (ratio : nombre total de réponses justes / nombre de réponses justes théoriques). Les résultats, pour la modalité parole, étaient les mêmes que pour la modalité spearcon. Toutefois, rappelons que les spearcons sont des stimuli plus rapides que la parole, car ils sont compressés à 50%. Sans surprise, la modalité earcon était la moins performante (au moins quatre fois moins que spearcon) et la modalité icône sonore était meilleure que earcon (et les modalités hybrides) mais moins performante que spearcon (au moins deux fois).

De ces trois expériences, nous pouvons conclure que l'earcon est une modalité peu recommandable, pour des tâches de navigation dans des menus contextuels, car si c'est un avantage d'avoir des possibilités infinies de conception, c'est aussi un désavantage de devoir apprendre chaque relation abstraite dans chaque interaction. Cependant, les icônes sonores, qui pourraient être considérées comme les meilleures earcons possibles pour une interaction, sont largement utilisées dans les interfaces des systèmes (souvent comme feedback (retour d'information)). Dans le cas particulier de la navigation dans des menus contextuels, les icônes sonores se révèlent être moins performantes que les spearcons ou la parole. Enfin, les spearcons semblent être une modalité adaptée à cette tâche de navigation, car une seule phase d'entraînement permet à l'utilisateur d'intégrer le « langage » de l'interface. Sans surprise, la parole est très efficace comme technique d'assistance, mais, souvent, les utilisateurs préfèrent d'autres modalités car la parole est un stimulus relativement long. D'ailleurs, nous pouvons remarquer que les modalités, basées sur la parole (ex. parole, spearcon, etc.), sont dépendantes des langues maîtrisées par les utilisateurs (mots, prononciation, pitch, etc.), tandis qu'une icône sonore, comme par exemple le « froissement d'un papier » pour indiquer la suppression d'un document numérique, semble indépendante de tout langage naturel.

L'utilisation des mappings sonores améliore l'efficacité d'une interaction, dans un contexte de transfert d'information, pour des tâches de navigation, de manipulation ou de notification.

## 2.2.6 Apport de l'auditif aux situations d'interaction non visuelle

### *QUALITES REMARQUABLES DE L'AUDITIF*

Avec une modalité auditive, nous avons vu qu'il était possible de transmettre de l'information. Une modalité sonore peut s'appuyer sur des propriétés du son : les cinq principales, les intrinsèques et les composées. Les informations transmises peuvent avoir un « mapping » soit direct, soit indirect. Les mappings directs regroupent les interactions utilisant nos sens de manière écologique (au plus proche du réel). Voici quelques exemples de mapping direct : un son spatialisé à gauche (azimut) pour transmettre l'information de tourner à gauche ; une intensité d'une source sonore (connue) transmet une distance égocentrée ; et bien sûr, la parole. Les mappings indirects regroupent les interactions nécessitant un traitement plus complexe de l'information car celle-ci est codée dans un stimulus ne reposant pas sur un langage. Voici quelques exemples de mappings indirects : une earcon de « son de trompette » pour confirmer la réception d'un mail ; la durée d'un son pour coder une distance ; un son aigu d'un piano pour indiquer une position en haut à droite d'une cible.

Dès lors qu'un son a une durée supérieure à 0,1 seconde et inférieure à 5 secondes, toutes ses propriétés (voir Figure 2-16) sont perceptibles. Toutefois, la discriminabilité entre deux sons différents dépend des différences entre leurs propriétés. De nombreuses études, dans le domaine de la psychoacoustique, explorent plus avant ce qui a été dit dans cet état de l'art. Néanmoins, nous avons synthétisé l'essentiel de la modalité auditive, pour concevoir des techniques d'interaction innovantes, basées sur cette modalité, afin de transmettre de l'information.

Ainsi, la modalité auditive peut être décomposée en deux modalités : la sonification et la parole. Nous avons vu que la modalité sonification permet de transmettre des informations perceptibles de manière précise. Par exemple, nous avons vu des interactions capables de transmettre des couples de valeurs direction-distance, uniquement par sonification. Mais la modalité parole est très efficace, car c'est un mapping direct, facilement interprétable par les utilisateurs. Nous avons alors vu que des couples de valeurs direction-distance pouvaient être transmis par une interaction basée sur la combinaison de ces deux modalités, i.e. de la

parole sonifiée. Ainsi, la modalité de la sonification permet la fusion des informations dans un même signal, tout en conservant l'aspect écologique d'une interaction. Nous utiliserons cette modalité dans nos première et deuxième études.

A l'instar de la modalité tactile, nous pouvons conclure que l'aspect dynamique des interactions auditives améliore la perception et l'efficacité des techniques, car les utilisateurs peuvent facilement « mapper » une perception (intensité, durée, fréquence, différence interaurale, timbre, rythme, tempo, note, azimuth) à une valeur ou une position dans l'espace. Toutefois, dans le cas où l'interaction ne peut être dynamique, une phase d'apprentissage est nécessaire, afin de créer un « mapping » ou plutôt un référentiel, sur lequel les utilisateurs se baseront, pour estimer au mieux la valeur correspondante au stimulus. C'est pourquoi, toutes nos expérimentations comportent une phase d'apprentissage.

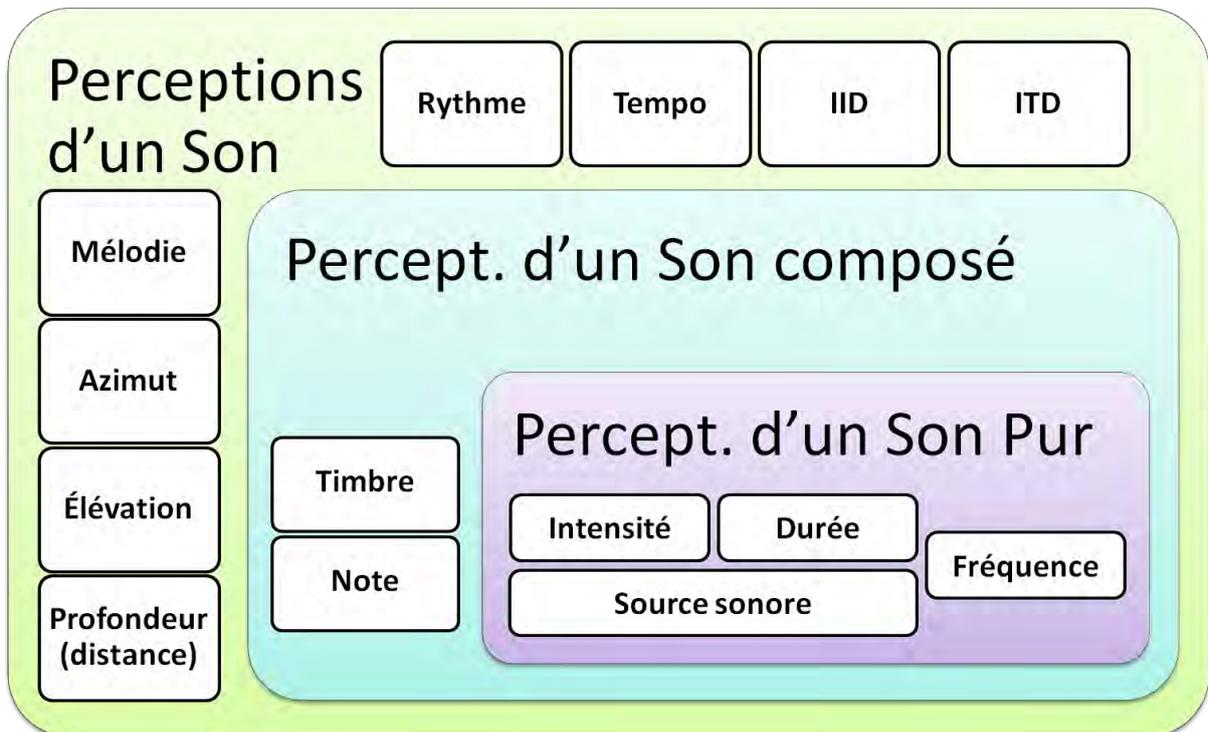


Figure 2-16 Schéma récapitulatif des propriétés perceptibles d'un son.

Dans la section suivante, nous expliquons les termes modalité et multimodalité, et présentons les propriétés CARE et CASE de la multimodalité.

## 2.3 Multimodalité

### 2.3.1 Modalités et multimodalité

#### *LA MODALITE*

Le terme modalité est défini selon deux aspects qui peuvent se confondre. Du point de vue de la perception humaine, ce sont les cinq **modalités sensorielles (audition, vision, toucher, olfaction, gout)**, qui sont également des modes de communication.

Du point de vue de l'Interaction Homme-Machine, les **modalités d'interactions** décrivent la manière d'interagir avec le système. Selon la définition de (Nigay and Coutaz 1996), une modalité d'interaction est un **couple dispositif-langage**. Cependant, selon la définition de (Bellik and Teil 1992a), la modalité est une **forme concrète particulière d'un mode de communication**. C'est de celle-ci dont on traite dans cette section, tout en nous limitant aux modalités auditives et tactiles.

#### *LA MULTIMODALITE*

Ce terme décrit une interaction où plusieurs modalités entrent en jeu : « À partir de la définition d'une modalité d'interaction, nous déduisons qu'un **système interactif est multimodal lorsqu'il possède plusieurs modalités en entrée ou en sortie**. Notons qu'un système peut être multimodal même s'il ne possède qu'un seul dispositif physique » (extrait de (Serrano 2010)). Le premier exemple ou plutôt la première interaction considérée comme étant multimodale est le « Put that there » de (Bolt 1980) dans laquelle l'utilisateur combine la parole et le geste pour exprimer une commande.

Avec ces définitions, nous avons la base pour concevoir et décrire des techniques d'interactions. Lors d'un processus de conception de techniques d'interaction, il est important de définir les modalités mises en œuvre et leurs propriétés.

### 2.3.2 Les propriétés de la multimodalité

#### *LES PROPRIETES CARE*

Les propriétés **CARE** (**C**omplémentarité, **A**ssignation, **R**edondance et **É**quivalence) présentées par (Martin 1994), puis enrichies par (Coutaz and Nigay 1994; Nigay and Coutaz 1997), caractérisent les usages combinés de plusieurs modalités d'interaction. Il existe quatre propriétés : Complémentarité, Assignation, Redondance et Équivalence.

(Nigay and Coutaz 1996) ont défini la propriété de **complémentarité du modèle CARE pour des modalités** en entrée, puis (Chevrin 2004; Cremetz 2007) appliquent toutes les propriétés CARE pour les modalités en entrée comme en sortie.

En nous inspirant des définitions de ces quatre propriétés, nous proposons une mise en œuvre des propriétés CARE adaptée aux modalités utilisées dans le contexte de la transmission d'information : des modalités sont **équivalentes** si chacune permet de transmettre la même information ; une modalité est **assignée** à une information s'il n'existe pas de modalité équivalente pour transmettre la même information. Des modalités sont **complémentaires** si elles sont nécessaires et suffisantes pour transmettre une même information composée de sous-unités, chacune étant exprimée par des modalités complémentaires. Des modalités sont **redondantes** si elles sont équivalentes et utilisées simultanément pour transmettre la même information.

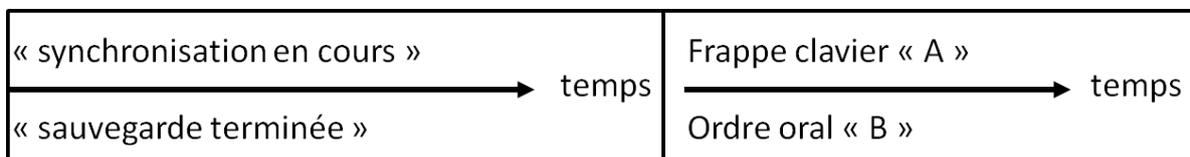
En complément de ces propriétés, centrées sur l'usage de multiples modalités par un utilisateur, nous allons décrire, dans la section suivante les propriétés CASE.

#### **LES PROPRIETES CASE**

Dans un contexte de transmission d'information(s) entre un Homme et une machine, une information peut être générée à partir de plusieurs informations. Par exemple, un message vocal « danger ! » avec un message textuel « Fuite de GAZ » peuvent être adressés à l'utilisateur pour l'alerter des événements **[DANGER]** et **[Fuite de GAZ]**. Dans ce cas précis, les informations sont transmises en parallèle et de manière complémentaire pour transmettre l'information **[Fuite de GAZ dangereuse]**. D'une manière générale, lorsqu'une information est composée de plusieurs sous-unités d'informations complémentaires ou redondantes, alors nous parlons de fusion de sous-unité d'informations ou d'informations dépendantes. Au contraire, dans le cas d'une transmission d'informations indépendantes, c'est-à-dire ni complémentaires, ni redondantes, il n'y a pas de fusion. De plus, les informations, qu'elles soient dépendantes ou indépendantes, peuvent être produites à partir d'une même modalité d'interaction ou à partir de modalités d'interactions différentes. Elles peuvent être transmises en simultané ou séquentiellement. Ainsi, pour définir les différents types de transmissions d'informations, nous nous sommes appuyés sur les propriétés CASE de la multimodalité. Ces propriétés **CASE** (**C**oncurrente, **A**lternée, **S**ynergique, **E**xclusive) ont été définies dans les années 90 par la communauté scientifique du domaine des IHM. Le groupe de travail d'IHM'91 a proposé un espace de classification des systèmes multimodaux, du point de vue du système : les propriétés CASE. Cet espace de travail fut adapté par (Bellik and Teil 1992b) puis approfondi par (Nigay and Coutaz 1993, 1996). Nous avons adapté les propriétés CASE à notre contexte de transmission d'informations. Ainsi, nous pouvons affiner la description d'une transmission

d'informations, en fonction de la nature des données (dépendantes ou indépendantes, basées sur la même modalité ou non) et de l'aspect temporel de la présentation des données (séquentielle ou en parallèle). Ci-après, nous proposons nos définitions des propriétés CASE, adaptées à notre contexte de transmission d'informations, c'est-à-dire que chacune des quatre propriétés CASE représente un des quatre types de transmission d'informations possibles.

La transmission (de type) **concurrente** désigne une transmission, en parallèle, de plusieurs informations indépendantes (ni complémentaires, ni redondantes). Il s'agit par exemple, de deux messages vocaux transmis, en même temps, à l'utilisateur ou le fait que l'utilisateur tape « A » au clavier et ordonne oralement d'écrire « B » en même temps (cf. Figure 2-17).



**Figure 2-17 : Illustration de transmissions concurrentes de plusieurs informations indépendantes : à gauche, deux messages vocaux en sortie, l'utilisateur devra percevoir correctement les deux messages qui s'entremêlent (au point de vue perceptif) ; à droite, deux informations « numériques » en entrée qui sont reçues en même temps par le système, mais affichées en séquence : « AB » ou « BA » ? Le système ne peut savoir quel caractère écrire en premier et encore moins les écrire en simultané.**

La transmission (de type) **synergique** désigne une transmission, en parallèle, de plusieurs d'informations dépendantes (redondantes ou complémentaires), et chacune exprimée par une modalité différente (impossible, sinon, de percevoir plusieurs informations dépendantes car elles seraient indiscernables). Il s'agit ici, d'une fusion de sous-unités d'information. Voici des exemples de transmission d'informations synergiques. D'abord complémentaires : un message vocal dont l'intonation de la voix est une information («DANGER » + voix grave) ou un utilisateur appuie simultanément sur deux touches du clavier pour faire l'action de copier (très utile dans les années 90 pour les combinaisons de touches telles que CTRL+C pour copier du texte) (cf. Figure 2-18). Puis, un exemple de transmission d'informations synergiques et redondantes : une alerte auditive de danger signale « le danger d'une fuite de gaz », et en même temps, des gyrophares signalent visuellement le signal d'alerte « DANGER GAZ ».

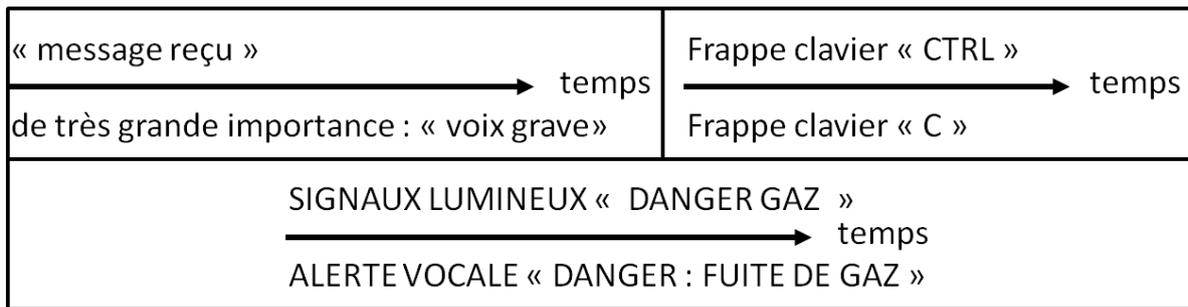


Figure 2-18 : Illustrations de transmission synergique de plusieurs informations dépendantes : à gauche, un message vocal sonifié en sortie. L'utilisateur devra percevoir correctement les deux sous-unités pour comprendre l'ensemble de l'information. A droite, le système devra interpréter deux informations « numériques » en entrée qui sont reçues en même temps comme la commande « copier ». En bas, l'utilisateur pourra comprendre l'information par l'une ou l'autre modalité, car elles sont redondantes.

La transmission (de type) **alternée** désigne une transmission, en séquence, de plusieurs informations dépendantes (redondantes ou complémentaires), et chacune exprimée par une modalité différente. Il s'agit ici, d'une fusion de sous-unités d'information. Par exemple, bouger la souris, puis dire « clic » pour cliquer sur un bouton (cf. Figure 2-19).

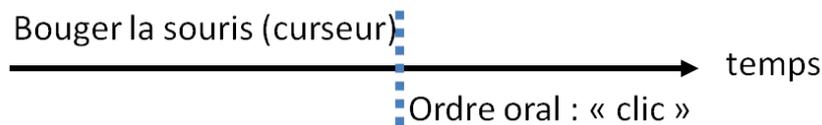


Figure 2-19 : Illustration d'une transmission alternée de plusieurs sous-unités d'information : l'utilisateur pointe un bouton avec sa souris (curseur) et dit la commande vocale « clic » pour ordonner au système de cliquer sur le bouton.

La transmission (de type) **exclusive** désigne une transmission, en séquence, de plusieurs informations indépendantes (ni complémentaires, ni redondantes). Par exemple, un utilisateur programme un « GPS-Météo France » pour se déplacer, le GPS-MF commence par lui donner des indications verbales sur la météo, concernant son itinéraire (prévention), puis, lui affiche l'itinéraire à suivre (guidage). Un autre exemple serait qu'un utilisateur transmette un message par écrit, puis un autre différent, mais toujours par écrit. Notons que, la transmission en séquence d'informations indépendantes peut se faire pour une même tâche ou deux tâches différentes. Par exemple, prévenir par SMS, les habitants d'un quartier, qu'il y a une fuite de gaz, puis prévenir par SMS, les responsables du réseau de distribution qu'il y a une fuite de gaz.

Dans le cas particulier où une séquence d'informations dépendantes serait transmise par la même modalité, cette modalité serait assignée à la tâche et l'interaction serait monomodale exclusive ; la transmission serait de type exclusive. Notons que, nous pouvons appliquer ce raisonnement au cas particulier d'une transmission d'informations dépendantes, en

parallèle, transmises par la même modalité. Cette modalité serait alors assignée à la tâche, l'interaction serait monomodale concurrente et la transmission serait de type concurrente.

### 2.3.3 Mise en œuvre de la multimodalité dans des situations interactives

Cette section a pour but d'illustrer différents types de transmission d'information(s) à partir d'une interaction multimodale. Nous nous intéressons, plus particulièrement, aux situations mettant en œuvre des techniques d'interaction multimodale, basées sur l'utilisation de modalités non visuelles.

(Vieira *et al.* 2015) ont développé le système de rééducation *SleeveAR*, dans le but de rendre autonome un utilisateur pendant ses séances de rééducation (sans que la présence d'un spécialiste soit requise). Le but du système est de faire reproduire un geste (pré-enregistré) à l'utilisateur sans que celui-ci ne se blesse. Ce système multimodal transmet des messages visuels, auditifs et haptiques à l'utilisateur : guidage du mouvement (flèche sur le bras, son, vibration), barre de progression visuelle, alerte tactile, etc. Les informations visuelles sont transmises en **concurrence**, car l'une guide le mouvement, l'autre affiche « d'autres-informations-utiles-et-non-directement-liées-au-mouvement-lui-même » en simultané. Les modalités d'interaction visuelle sont concurrentes. Les informations auditives et haptiques sont transmises en **synergie** et de façon **complémentaire**, pour guider (en simultané) l'utilisateur, lorsque l'information visuelle de guidage ne lui est plus accessible. Les modalités d'interaction auditives et haptiques sont utilisées en synergie et sont complémentaires. D'un certain point de vue, ces deux modalités d'interaction (ensemble) sont équivalentes à la modalité d'interaction visuelle (cf. Figure 2-20), car elles transmettent (ensemble) la même information. Notons également, que la transmission d'informations peut, alternativement, s'effectuer avec l'un ou l'autre mode de communication, car ils sont redondants. Enfin, une seconde modalité d'interaction haptique est **assignée** pour transmettre **exclusivement** une information vitale à l'utilisateur, à propos de son action en cours.



**Figure 2-20 : illustration du système SleeveAR, l'utilisateur effectue un mouvement, il ne peut pas lire les informations visuelles (bras droit) qui sont hors de son champ de vision car le sujet gardait la tête droite pour ce mouvement de rééducation, un guidage audio et haptique prend le relais.**

(Bolelli *et al.* 2004) ont présenté la preuve de concept d'un système collaboratif de gestion de crise facilitant la prise de décisions et leurs diffusions. En s'appuyant sur le principe du système existant GeoMIP (Agrawal *et al.* 2004) et sur l'image centrale de la Figure 2-21, voici une explication du fonctionnement de leur interaction : le coordinateur pointe du doigt une zone vers la carte (l'écran) tout en parlant, le système récupère ces informations qui lui sont transmises en **synergie** (en même temps) ou en **alternance** (en décalé) par deux modalités d'interaction, auditive et haptique, utilisées de manière **complémentaires** par l'utilisateur. Les autres participants reçoivent l'information par une transmission d'informations dépendantes visuelles et auditives, en synergie ou en alternance. Sur les images latérales de la Figure 2-21, nous pouvons voir deux autres illustrations de transmission d'informations **complémentaires** en **synergie** ou en **alternance** : stylos (modalité tangible) + parole (modalité auditive) et souris (modalité tangible) + parole (modalité auditive).



**Figure 2-21 : Illustration de la preuve de concept d'un système collaboratif de gestion de crise inspirée du GeoMIP. Les interactions permettent une transmission d'informations complémentaires en synergie ou en alternance.**

Dans le contexte d'utilisation spécifique aux interfaces « Paper-based », (Liao, Guimbretière, and Loekenhoff 2006) ont conçu le prototype « Pen-top » afin de tester trois catégories de techniques d'interaction (dans trois tâches) : la navigation, la notification et la sélection. Ces

interactions multimodales utilisent des modalités d'interaction visuelles, auditives et tactiles (cf. Figure 2-22). Prenons l'exemple de la tâche de navigation dans leur *pie menu* (menu circulaire), où les modalités sont complémentaires. Plus précisément, lors d'un changement de menu, les modalités sont **complémentaires** et utilisées en **alternance** pour transmettre l'information. Ainsi, chacune des modalités d'interaction est **assignée** à une sous-unité d'information : visuelle pour la direction, tactile pour prévenir que la position du stylo est proche de deux menus, sonore pour un survol de la frontière entre deux menus et enfin la parole pour indiquer le titre du menu survolé. Prenons un autre exemple de l'une de leurs techniques dédiée à une tâche de recherche de mots-clés. L'utilisateur doit rechercher un mot dans un texte, pour ce faire il survole le texte avec un stylo puis lorsque le « stylo » reconnaît le mot-clé dans le texte ou la ligne courante (survolée), des informations visuelle et auditive sont transmises à l'utilisateur. Ainsi, les modalités d'interaction auditives et visuelles sont **assignées** et **exclusives** : la parole indique la présence de l'élément recherché dans la page, puis un indice visuel informe l'utilisateur que la ligne qu'il survole avec le « Pen-top » contient l'élément recherché. Enfin, dans une tâche de sélection, le système transmet l'information « la confirmation finale » avec une modalité d'interaction visuelle (**assignation exclusive**) pour indiquer à l'utilisateur que son mot est enregistré pour une future action (copier-coller par ex.).

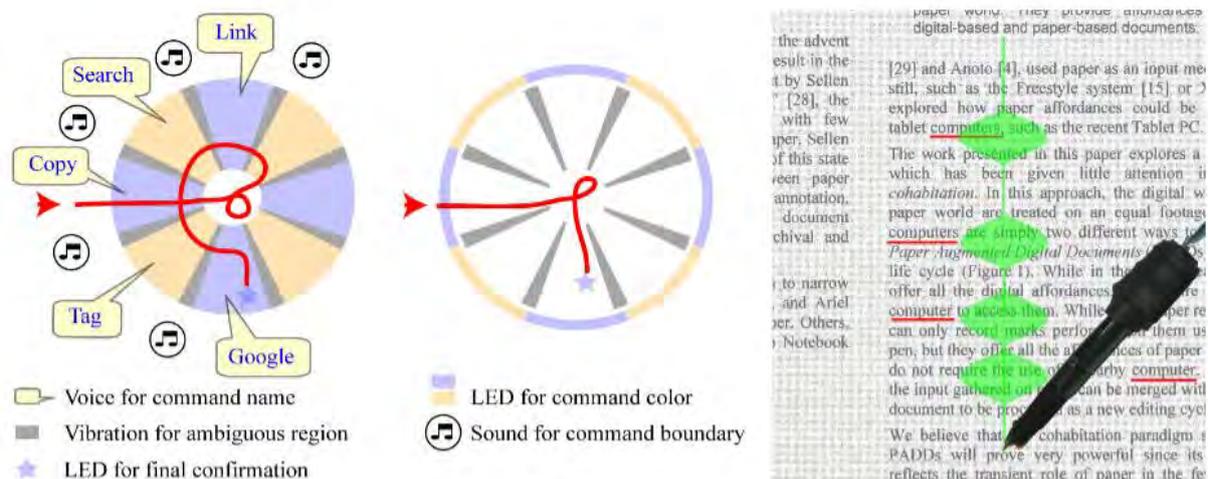


Figure 2-22 : Illustration issue de (Liao et al. 2006) des interactions du « Pen-top » : à gauche, un utilisateur novice recevant des informations relatives à sa navigation ; au milieu, la même navigation mais dans le mode expert (donc moins d'indices) ; à droite, la recherche du mot « computer » dans un texte.

(Cassell et al. 1999) ont présenté un système multimodal de communication basé sur les gestes et la parole. Ce système de communication entre un Homme et une machine (robot, IA, etc.) dépendait d'une machine à états (voir grammaire et théorie des langages). Ces

auteurs ont conçu un robot intelligent, nommé « Rea », capable de traduire certains gestes et certaines paroles de l'utilisateur en commandes (cf. Figure 2-23). En voici quelques exemples : lorsqu'un utilisateur faisait une phrase déclarative suivie d'une pause de plus de 500 ms sans faire aucun geste, Rea comprenait que l'utilisateur lui laissait la « parole ». Par contre si l'utilisateur faisait des gestes, Rea lui redonnait la parole. Lorsque Rea souhaitait la parole, elle regardait l'utilisateur, levait la main et prononçait l'onomatopée « hum ». Ainsi, les modalités d'interaction étaient utilisées soit en **alternance mais de manière complémentaire** (ex. l'utilisateur parlait puis effectuait un geste pour signifier à Rea qu'il souhaitait garder son tour de parole), soit en **exclusif assigné** (ex. la parole seule permettait à Rea de prendre le tour de parole, un geste permettait à Rea de signifier à l'utilisateur qu'elle attendait son tour).



Figure 2-23 : Illustration issue de (Cassell *et al.* 1999) : « Rea le robot », en « pleine conversation » avec un participant.

(Hinckley and Song 2011) ont exploré les interactions multimodales avec un Smartphone en combinant le « toucher » et le « mouvement ». Les auteurs sont partis sur deux axes de conception : « *touch in motion* » et « *motion in touch* ». Dans le premier axe, les gestes incluent, en plus des informations de mouvements, des informations relatives au toucher, comme le nombre de point de contacts sur l'écran et leurs positions. Une des interactions conçues est la « Tilt-to-Zoom » qui facilite l'action de zoomer avec une seule main sur un écran tactile : poser le pouce sur l'écran puis incliner le Smartphone vers l'avant pour zoomer, vers l'arrière pour dé-zoomer (**modalités complémentaires synergiques**). Dans l'axe « *motion in touch* », les gestes incluent, en plus des informations « tactiles », des informations relatives au mouvement, comme la force, la vitesse, l'angle, etc. Ceci ajoute au « touch » une signification plus expressive. Dans ce second axe, les auteurs font la distinction entre le « soft tap » et le « hard tap » avec la règle suivante : deux « soft tap » équivalent à un « hard tap » (**transmission exclusive d'informations dépendantes, dont la**

**modalité d'interaction est le « soft tap » sur l'écran, une séquence de deux « soft tap » donne l'information du « hard tap », c'est un cas d'interaction monomodale exclusive avec informations dépendantes).** Il existe une vidéo de démonstration sur le site de l'ACM<sup>31</sup>.

#### 2.3.4 Conclusion sur la multimodalité

En résumé, nous avons vu les définitions de modalité et multimodalité ; nous avons vu également les propriétés (CARE) s'appliquant aux dispositifs, aux langages et aux modalités. Enfin nous avons rappelé quatre types de multimodalité (CASE) et les avons adaptées à un contexte de transmission d'informations (quatre types de transmissions d'informations).

Dans notre contexte de transmission d'informations, l'équivalence de modalités d'interaction transmettant la même information permettrait de choisir la technique d'interaction la plus adaptée au contexte d'utilisation ou alors l'utilisateur pourrait choisir entre plusieurs techniques (préférence). De plus, la redondance d'une information, par l'utilisation de plusieurs modalités différentes, assurerait une plus grande robustesse de la transmission de cette information.

Enfin toujours dans un contexte de transmission d'information(s), une information peut être composée de plusieurs sous-unités d'information (cas de fusion) ou être divisée en plusieurs sous-unités d'information (cas de fission). Toutefois, lors d'une transmission d'information, celles-ci peuvent être réparties sur un ou plusieurs dispositifs (médias). (Nigay and Coutaz 1993). Ces mécanismes, de fusion et de fission, peuvent être utilisés dans des cas d'interaction multimodale alternée (voir l'exemple du robot Rea de (Cassell *et al.* 1999) ou synergique (voir l'exemple du système *SleeveAR* de (Vieira *et al.* 2015)).

#### 2.4 Conclusion de l'état de l'art

Les modalités d'interaction tactiles et auditives permettent la transmission d'information, dans des situations d'assistance, où la modalité visuelle n'est pas adaptée. La nature des informations transmises du système vers l'utilisateur, peut-être une alerte, une alarme, une notification, un feedback (retour d'information). Dans le sens inverse, humain vers système, la nature des informations est moins bien définie, cela peut-être une commande précise (« je valide ») ou l'envoi d'un stimulus (ex. « je souffle », « je bouge », etc.) qui sera capté et analysé par le système.

<sup>31</sup> [Sensor synaesthesia : touch in motion, and motion in touch, démonstration en vidéo.](#)

Nous avons vu que certaines techniques d'interaction d'assistance utilisant les modalités d'interaction tactiles ou auditives, permettaient de guider très efficacement l'utilisateur. Il ressort aussi que des interactions basées principalement sur la modalité tactile, augmentent l'immersion dans des environnements virtuels et améliorent la représentation mentale d'un utilisateur pour l'objet ou l'espace qui l'entoure (réel ou virtuel), notamment dans le cas d'interactions actives. Enfin, il apparaît qu'une combinaison de modalités peut être envisagée sous différentes formes (propriétés CARE) et donner lieu à des usages de différentes natures (propriétés CASE) susceptibles d'influencer la production et la perception des informations.

La modalité d'interaction de la parole est, à elle seule très efficace dans le transfert d'information. Cependant, nombre d'études comparatives ont révélé que les utilisateurs préféreraient autant les interactions basées sur la sonification que celles basées sur la parole, avec dans certains cas, une nette préférence pour la parole sonifiée.

La sonification est une modalité d'interaction, composée d'au moins douze paramètres, dont l'efficacité dépend énormément de l'environnement sonore (ex. bruit ambiant ou silence) et physique (ex. anéchoïque ou non). Toutefois, cette modalité permet la transmission d'informations simples (tonie<sup>32</sup>, sonie<sup>33</sup>, chronie<sup>34</sup>, etc.), spatialisées (azimut, élévation et distance) ou émotionnelles (voir musique). Cette modalité peut être combinée à la parole, mais aussi aux modalités tactiles. Toutefois dans ce dernier cas, l'interaction passera par deux modalités issues de deux grandes familles de modalités : tactile et auditive. Ce faisant, le flux d'information ne pourra pas être fusionné sur un même dispositif. Remarquons que, là encore, le Smartphone pourrait être le chaînon manquant.

En résumé de cet état de l'art, portant sur l'utilisation d'interactions, mono ou multimodales, basées sur les modalités tactiles et auditives de techniques de transmission d'information(s), nous retenons trois points.

Le premier point est qu'il existe beaucoup de techniques d'interaction non visuelles pour transmettre de l'information. Cette multiplicité des techniques couvre les besoins des utilisateurs et ce, dans des contextes d'usages différents (assistance, guidage, navigation, alarme, alerte, notification, etc.). Toutefois, dans certaines situations, la modalité tactile serait plus adaptée ou préférée à la modalité auditive (ex. perception de la forme d'un

---

<sup>32</sup> Tonie : grandeur « linéaire » subjective associée à la fréquence dans le domaine psychoacoustique.

<sup>33</sup> Sonie : grandeur « linéaire » subjective associée à l'intensité dans le domaine psychoacoustique.

<sup>34</sup> Chronie : grandeur « linéaire » subjective associée à la durée dans le domaine psychoacoustique.

objet), et inversement (ex. perception d'une mélodie). Néanmoins, les modalités tactiles et auditives permettent d'estimer, et donc de percevoir, des dimensions spatiales (ex. distances et directions), des dimensions temporelles (ex. durée et rythme) et des grandeurs comme par exemple une intensité ou une fréquence.

Les propriétés d'un son, d'une vibration, d'un mouvement, d'un langage oral ou gestuel, offrent des possibilités quasi infinies de dérivation des techniques d'interaction. Ajoutons l'utilisation des propriétés CARE pour les modalités et CASE pour les interactions multimodales et les types de transmission d'informations et, nous obtenons un immense champ des possibles en matière de conception de technique. Ce deuxième point est important car il justifie l'utilisation de modèles (descriptions abstraites) pour décrire ces possibilités de dérivation.

Le troisième point est l'importance des études comparatives de techniques d'interaction. Ces études comparent des techniques différentes dans un même cadre expérimental, pour les mêmes tâches et avec les mêmes populations. Car, si nous avons vu que les modalités tactiles et auditives permettaient de transmettre efficacement de l'information, cela ne nous a pas renseigné sur leurs similitudes, ni sur leurs différences. Or comment choisir la technique la plus adaptée à une situation, s'il n'y a qu'une solution proposée ? Ainsi, nous avons pris soin, dans les expériences menées dans le cadre de cette thèse, de privilégier l'approche transversale<sup>35</sup> et donc, les études comparatives, qu'elles soient sous la forme de pré-tests, d'étapes de conception ou d'études exploratoires et expérimentales.

Toutefois, pour organiser ce champ immense des possibles et structurer les multiples résultats expérimentaux, nous consacrons le chapitre suivant à un cadre analytique (conception primitive d'un modèle). Nous y détaillerons des propriétés extraites de la littérature, ainsi que nos techniques d'interactions conçues et évaluées dans le cadre de cette thèse.

Plus globalement, du point de vue des techniques d'interaction non visuelle, nous avons mis en évidence du multimodal basé uniquement sur de l'haptique ou une combinaison de modalités haptiques et auditives. Dans les deux cas, différentes déclinaisons de situations CASE et CARE ont été illustrées par des exemples de la littérature. Devant ce large choix, apporter une aide conceptuelle à la réflexion, pour une conception maîtrisée de techniques multimodales non visuelles, semble incontournable (voir chapitre suivant).

---

<sup>35</sup> L'approche transversale a la particularité de consister à étudier, à un moment donné, en une seule fois, plusieurs techniques d'interaction.

# **Chapitre 3**

## **Utilisation/Définition d'un cadre analytique**



### 3 Utilisation/Définition d'un cadre analytique

#### 3.1 Motivation

Dans l'état de l'art, nous avons vu que la capacité humaine, à percevoir et à générer des informations (stimuli), varie en fonction de la modalité sensorielle utilisée par les techniques d'interaction. Cependant, la définition de la modalité d'interaction de (Nigay and Coutaz 1996), et donc la définition d'une technique d'interaction, permet de décrire une technique, uniquement, d'un point de vue système. Ce point de vue système permet de décrire une technique d'interaction avec un couple dispositif-langage et avec plusieurs couples dans le cas de techniques multimodales. Toutefois, dans le cadre de cette thèse, nous avons besoin de décrire une technique d'interaction en considérant, également, le point de vue humain. Dans ce but, nous proposons un cadre d'analyse, visant à orienter nos choix de conception, afin d'optimiser les interactions pour de la transmission d'information en considérant les modalités sensorielles (cf. Etat de l'art sur la transmission d'information basée sur les modalités tactiles et/ou auditives), les modalités d'interaction (Nigay and Coutaz 1996) et les propriétés de la multimodalité (cf. Multimodalité). C'est pourquoi, en plus de décrire une technique d'interaction par sa modalité d'interaction (couple dispositif-langage), nous décrivons la technique par la zone du corps qu'elle sollicite et par le type de transmission qu'elle utilise (cf. Les propriétés de la multimodalité) pour transmettre les informations.

Néanmoins, dans la littérature présentée dans le chapitre 2, nous avons remarqué que certaines interactions faisaient participer activement les utilisateurs dans des tâches d'estimation (Proffitt *et al.* 2003) ou de reconnaissance de forme (Gibson 1962) et que cela avait un impact sur les estimations. Bien sûr, nous avons, également, vu des techniques efficaces ne faisant pas participer activement les utilisateurs pour percevoir l'information, c'est-à-dire que les utilisateurs étaient « passifs » (par ex. (Mershon and King 1975), (Brunet *et al.* 2013), etc.). Nous avons également observé, que certaines transmissions d'information étaient basées sur des stimuli continus (par ex. (Fujimoto and Turk 2014) et (Frasse 1984)), alors que d'autres étaient basées sur des stimuli discrets, par opposition à continu-pendant-toute-la-durée-de-la-tâche. Un exemple de stimulus discret serait un feedback (retour d'information) auditif pour signaler la fin d'un transfert de fichier et un exemple de stimulus continu serait le feedback visuel de la progression du transfert (la « progress bar » par exemple).

Enfin, nous avons extrait deux propriétés spécifiques à notre contexte de transmission d'information. La première est le type de codage de l'information, nous parlons ici du système représentationnel sur lequel repose le langage d'une modalité d'interaction (modalité = couple dispositif-langage, selon Nigay 96). En premier lieu, nous avons le codage par les langages naturels (parole, Braille, écriture, etc.) ou abstraits (tacton, earcons, gestuel, etc.). Puis d'autres, comme par exemple, la durée d'un signal (ex. note de musique) ou son intensité. La seconde propriété concerne la présence ou l'absence de référence dans une transmission d'informations, et plus précisément, juste avant ou après le signal contenant l'information. Nous avons extrait cette propriété d'après les travaux de (Fraisie 1984), où P. Fraisie suggère que la présence d'un son avec durée fixe, juste avant ou après un son avec une durée variable améliorerait la précision de l'estimation de la durée (variable) du son.

Ainsi, ces propriétés, que nous avons extraites et cartographiées à partir des techniques présentées dans la littérature, ont été regroupées sous la forme de facteurs de conception. Dans ce chapitre, nous posons les bases d'un cadre analytique rassemblant les multiples facteurs identifiés comme support de description structurée et systématique, de techniques d'interaction pour de la transmission d'informations. Nous l'utiliserons ensuite, pour décrire nos techniques et montrer leurs différences, déterminer l'influence de certains facteurs et expliquer ainsi, les différences d'utilisabilité entre les techniques.

Nous avons procédé en trois temps : l'**extraction** des propriétés générales issues de l'état de l'art, leur **organisation** dans un cadre d'analyse et l'**illustration** de son utilisation par la génération de techniques d'interaction pour de la transmission d'informations.

### 3.2 Présentation des facteurs de conception de notre cadre analytique

Nous avons cartographié les propriétés de conception traitées et observées dans la littérature et avons retenu huit facteurs :

1. La modalité est la base pour décrire une technique d'interaction en IHM. Il existe plusieurs modalités non visuelles : dans le cadre de cette thèse nous avons choisi de travailler sur des modalités auditives ou tactiles. Il convient de préciser sur laquelle de ces deux grandes familles de modalités, se basent nos techniques. Usuellement, une modalité d'interaction haptique requière la modalité sensorielle tactile et une modalité d'interaction auditive requière la modalité sensorielle auditive. Rappelons

également que selon la définition de (Bellik and Teil 1992a), la modalité est une forme concrète particulière d'un mode de communication.

2. Le dispositif, car nous nous basons sur la définition de modalité d'interaction de (Nigay and Coutaz 1996) <dispositif-langage> pour décrire nos techniques. Ainsi, le dispositif est un facteur de conception à part entière, ce qui implique la description du dispositif pour chacune de nos interactions.
3. La partie du corps sollicitée par l'interaction. De plus, il est possible de détourner un dispositif existant pour de nouvelles interactions (ex. une pédale à pied utilisée avec une main).
4. Le type de codage, car nous nous appuyons sur la définition de modalité (Nigay and Coutaz 1996) <dispositif-langage> où le langage englobe le système représentationnel dans lequel est codée l'information. Nous avons choisi d'explicitement le langage utilisé. Par exemple, le langage gestuel, qui est en fait un mouvement ou une succession de mouvements d'une partie mobile, peut à lui seul, encoder une information, par la durée d'un mouvement, sa force, ses rotations, ses translations ou alors par le geste lui-même (i.e. un « mot » abstrait du langage gestuel).
5. Le type de signal, car nous avons besoin de connaître le type des signaux utilisés pendant la transmission d'une information. Par exemple, pour indiquer à l'utilisateur qu'une copie de fichier est en cours, le système affiche en continu la progression de la copie et lorsque celle-ci est finie, le système envoie un bip audio. Ce bip audio est de type discret, car il est juste envoyé à la fin de la tâche et plus précisément pour signifier la fin de la tâche. Nous nous sommes appuyés sur l'espace de classification de (Roudaut and Lecolinet 2007) où les dispositifs sont décrits comme étant continus ou discrets, c'est-à-dire qu'ils transmettent soit de l'information en continu, soit de manière ponctuelle (discrète). Nous avons adapté leurs définitions dans notre contexte de transmission d'information non visuelle.
6. Le type d'utilisation de la modalité d'interaction car, comme nous le faisons couramment pour décrire le sens d'utilisation d'une interaction (entrée et/ou sortie), il est important de caractériser le sens d'une transmission d'information : de l'Homme vers la machine, de la machine vers l'Homme ou les deux. Nous expliquons dans la section suivante les différents types d'utilisation, et ce que cela implique dans la conception. Notons que l'interaction peut être en « entrée » et donc en

sortie utilisateur, ou en « sortie » et donc en « entrée utilisateur ». Bien que ce facteur soit redondant avec la zone du corps, nous souhaitons garder ces deux facteurs afin d'explicitier le lien entre le point de vue machine et le point de vue Homme.

7. Les catégories de techniques <actives> ou <passives>, car nous avons vu l'importance d'être actif dans une tâche de reconnaissance de formes (cf. L'« active touch », (Gibson 1962)). Cependant, nous nous posons la question de la transitivité de cette propriété dans notre contexte de transmission d'information non visuelle : être actif dans une tâche d'estimation permet-il à l'utilisateur d'être plus efficace que s'il est passif ? La satisfaction des utilisateurs est-elle meilleure pour des techniques dites actives que pour des techniques dites passives ? Nous avons consacré une partie de notre première expérimentation à ces questions.
8. Le nombre de références, car même si cela semble évident que le fait d'ajouter des références peut faciliter l'estimation des valeurs transmises, et que cela est conseillé dans la littérature (Fraisie 1984), nous nous posons la question de l'impact de ce facteur dans des tâches d'estimation :
  - sur la précision et la dispersion de l'erreur d'estimation ;
  - sur la dérive de l'erreur au cours du temps ;
  - sur la satisfaction des utilisateurs.

Nous avons consacré une partie de notre première expérimentation à cette question.

La section suivante détaille nos huit facteurs de conception.

### **3.2.1 Cartographie des facteurs pertinents pour une transmission d'information**

#### ***DEUX MODALITES NON VISUELLES CONSIDEREES***

Le cadre de nos travaux se limite aux interactions non visuelles, auditives et tactiles. Cela inclut les modalités sensorielles de l'Homme et les modalités d'interactions entre l'Homme et la machine. Dans l'état de l'art, nous avons décrit les Modalités tactiles et les Modalités auditives sur lesquelles nous avons travaillé, dans le cadre de cette thèse. Il convient de préciser sur laquelle de ces deux grandes familles de modalités se basent nos techniques.

Le facteur modalité fait partie de notre cadre d'analyse.

#### ***DISPOSITIFS ET PARTIES DU CORPS***

Dans une interaction Homme-Machine, l'Homme utilise nécessairement une partie de son interface sensori-motrice et manipule éventuellement un objet tangible. La machine utilise des dispositifs d'entrées/sorties (ou média) pour communiquer.

Le Tableau 3-1 présente des dispositifs utilisés dans des techniques d'interaction issues de notre état de l'art, avec la partie du corps associée.

Dispositifs physiques	bracelet vibrotactile (cf. Modalité d'interaction, haptique, vibrotactile)	Cybergrasp™ (cf. section Modalité d'interaction haptique)	Smartphone / canne (cf. Modalité d'interaction tangible)	casque audio (cf. Modalités auditives)
partie du corps	poignet	doigts	main	oreilles

**Tableau 3-1 : exemples de dispositifs utilisés dans des techniques d'interaction de l'état de l'art, avec la partie du corps associée.**

Les dispositifs sont décrits du point de vue machine et peuvent être regroupés selon trois catégories :

1. Les dispositifs de type Capteurs :

Un capteur prend de l'information et la transmet au système ;  
Ex : capteur de lumière, d'humidité, gyroscope, etc.

2. Les dispositifs de type Effecteurs :

Un effecteur transmet de l'information à l'utilisateur ;  
Ex : hautparleur, écran de télévision, moteur vibrant, etc.

3. Les dispositifs de type Objets Tangibles :

Les objets tangibles en IHM servent d'interfaces (tangibles) intermédiaires entre l'Homme et la machine. **Les objets tangibles sont équipés de capteurs et/ou d'effecteurs afin de transmettre des informations** tactiles (moteurs vibrants, centrales inertiels, ...), auditives (micros, hautparleurs, ...) ou visuelles (lampes, capteur de lumière, ...). Ils sont très souvent utilisés dans les systèmes d'interactions mixtes (SIM) (cf. Modalité d'interaction tangible). Ils ne sont pas de simples capteurs ou effecteurs, mais un ensemble de capteurs et/ou d'effecteurs. Les plus connus

portent des noms comme les Smartphones, sinon certains les appellent « artéfacts »<sup>36</sup>.

Ex : un cube avec un code QR sur l'une de ses faces posé sur une table interactive, une maquette d'avion avec des gyroscopes pour « conduire » un avion dans un simulateur, etc.

Le facteur média (dispositif) fait partie de notre cadre d'analyse.

### **Pourquoi ajouter la partie du corps associée à notre cadre d'analyse ?**

Nous avons observé que les interactions étaient souvent décrites comme des couples <dispositif-langage>, mais l'explication de la technique explicite systématiquement la partie du corps sollicitée par l'interaction elle-même. De plus, dans le cadre de cette thèse, nous avons vu que les différentes parties d'un corps humain avaient des propriétés différentes en termes de sensibilité (ex. le visage est sensible à une pression de <20mg, tandis que le pied a un seuil de 400 mg), de seuil de perception (ex. les mains ont une haute résolution pour discriminer deux points de contact), de capacité de perception tactile (ex. vibration optimale à 250 Hz). Enfin, bien souvent, le nom des dispositifs renseigne sur la partie du corps concernée (un brassard pour le bras ou l'avant-bras ; un bracelet pour le poignet ou la cheville ; un gilet pour le dos ou le ventre ; etc), cela devient moins évident pour les objets tangibles, même s'ils sont souvent dédiés à un usage manuel et donc aux mains (ex. Smartphone, Canne, etc.). Et encore moins évident avec certains dispositifs commerciaux car leurs noms ne sont pas explicites (Phantom, Cybergrasp, TéléAct II, etc.). Enfin, même si un dispositif est très connu pour être tenu en main, celui-ci peut exceptionnellement requérir une autre partie du corps : par exemple, un Smartphone peut être tenu en main ou attaché à l'avant-bras ou une pédale à pied utilisée avec une main.

Afin d'assurer la meilleure description possible, nous ajoutons le facteur : « partie(s) du corps » dans notre cadre d'analyse afin de lever toute ambiguïté et de prévenir des potentiels défauts de conception n'ayant pas pris en compte l'aspect cognition spatiale dans la conception d'une technique d'interaction (ici, propriétés différentes en fonction des zones du corps sollicitées).

#### ***DIFFERENTS TYPES DE CODAGES***

Dans notre contexte de transmission d'information, nous nous appuyons sur la définition de modalité (Nigay and Coutaz 1996) <dispositif-langage> pour décrire nos techniques, et plus

<sup>36</sup> <http://www.void-spaceartifact.com/fr>

précisément, sur le système représentationnel dans lequel est codé l'information. Bien que, nous disposons déjà des facteurs modalité, dispositifs et partie du corps pour décrire une interaction, cela ne suffit pas pour décrire le type de codage de l'information. Prenons l'exemple d'une des modalités haptiques d'interaction, le « Blind reaching » avec un mouvement de main. A ce stade de la description, nous savons que l'utilisateur fait un mouvement de la main pour interagir et transmettre une information au système. Mais quel est le code principal utilisé afin de transmettre cette information ? Un mouvement de main peut, par exemple, indiquer un angle de rotation, un vecteur de translation, ou alors une durée de mouvement : « tant que la main est levée, le son monte » ; « si la main fait un cercle vers la droite alors le son monte » ; « si la main se déplace vers le haut alors le son monte ». Le système représentationnel permet d'indiquer le « code » utilisé pour encoder l'information et nous avons déterminé six formes de codes distincts, que nous avons regroupées dans le facteur de conception : Type de codage.

Ainsi, ces six types de codage différents permettent de décrire le codage principal pour transmettre une information :

1. Angle (A) → technique qui implique une rotation ou une inclinaison  
Ex : tourner un volant, un cube, un cylindre, incliner une tablette, sa main, etc.
2. Longueur (L) → technique qui implique une translation/glissement  
Ex : translater une souris, faire glisser un slider/son doigt, sa main, etc.
3. Durée (D) → le temps écoulé entre deux évènements temporels.  
Ex : entre le début et la fin d'un son, entre le début d'un clic et la fin d'un second (ex. le double-clic), entre la fin d'une frappe et le début d'une autre, entre le début d'un mouvement de main et sa fin, etc.
4. Force (F) → la force ou l'intensité d'un stimulus  
Ex : appuyer sur un bouton, variation d'une pression auditive (volume sonore) ou tactile. Elle est souvent associée à des tâches d'estimation de magnitude, i.e. de force ou d'intensité.
5. Pulse (P) → signal qui se compose d'une unique ou d'une succession de pulsations auditives (ex. bips), tactiles (ex. vibrations) ou visuelles (ex. flashes). Ce type de codage donne seulement l'information d'un nombre (de pulsations).  
Ex : flash lumineux, un bip ou une note de musique, une vibration  
Ex2 : succession de flashes, de notes ou de vibrations  
Mais pas d'earcons ou de tacton (langage) !  
Ni de variation des pulsations (voir codage Force/intensité, Angle, Longueur, Durée) !
6. Langage (Lg) → le code utilise le langage naturel ou abstrait

Ex : la parole (langage naturel), les earcons (langage abstrait auditif), les tactons (langage abstrait tactile), les gestes (langage abstrait gestuel), des symboles graphiques (langage abstrait symbolique).

Le facteur type codage fait partie de notre cadre d'analyse.

#### **DIFFERENTS TYPES DE SIGNAL**

Nous avons observé au travers des exemples d'interaction issus de la littérature, la dichotomie entre un signal discret et un signal continu dans une transmission d'information. Il nous semble pertinent de décrire les types des signaux utilisés dans une transmission d'information. Comme point de départ à la spécification des termes « signal discret » et « signal continu », nous nous sommes basés sur l'espace de conception de (Roudaut and Lecolinet 2007) où les notions de dispositifs discrets ou continus sont définis ainsi :

✚ Dispositifs continus : « *Les écrans tactiles présents sur de nombreux terminaux mobiles sont des dispositifs continus. C'est-à-dire qu'ils récupèrent un signal continu représentant la position du pointeur à l'écran (rien ou plusieurs valeurs).* » ;

➔ Ainsi, le signal continu de la position du curseur est transmis en permanence au système.

✚ Dispositifs discrets : « *A la différence des dispositifs continus, l'information récupérée par un dispositif discret est une valeur booléenne (par exemple, une touche de clavier est appuyée ou non).* »

➔ Ainsi, le signal discret de l'état de la touche est transmis au système seulement lors d'un changement d'état.

Dans ces définitions, nous avons les notions de signal continu et de signal discret. Nous avons adapté ces définitions à notre contexte de transmission d'information afin d'apporter une meilleure description de la technique :

Signal continu : signal transféré pendant toute la durée de la tâche en continu ou par de rapides successions ou répétitions du signal.

Ex : signal du « bip audio pour signaler que les phares de la voiture sont allumés » ; voyant sur le tableau de bord pour signaler « qu'une porte de la voiture est restée ouverte » ; signal de changement de direction : « le clignotant d'une voiture » ; etc.

Signal discret : signal transféré une seule fois pendant la tâche (ne couvre pas toute la tâche) ou alors une succession/répétition du signal, très espacée temporellement.

Ex : un tacton pour signaler l'arrivée d'un nouveau message ; les gongs ou carillons d'une horloge pour indiquer les heures.

Le Tableau 3-2 , donne d'autres exemples de signaux continus et discrets :

Signal discret	Signal continu
vibration arrivée message	icone arrivée message
bip d'erreur de commande système	bip pour phare allumés et porte ouverte
appel de phare	clignotant pour changer de direction
horloge parlante toutes les heures	compte à rebours

**Tableau 3-2 : exemples de signaux continus et discrets.**

Le facteur type signal fait partie de notre cadre d'analyse.

#### **DEUX TYPES D'UTILISATION D'UNE INTERACTION**

Dans notre contexte de transmission d'information, nous avons exploré deux types d'utilisation possibles d'une technique d'interactions. Dans le cas général d'une interaction H-M, l'Homme agit sur le système et le système lui donne une information en retour, c'est le cas classique d'une interaction : l'action de l'utilisateur sur le système est toujours suivie par un feedback (retour d'information) dudit système en signe de bon fonctionnement. Car si le système ne renvoie pas d'information, suite à une action de l'utilisateur, cela indique généralement que le système est en panne ou alors qu'il n'y a pas d'interaction. Outre le cas classique, il est possible que le système envoie une information à un utilisateur sans que celui-ci n'en soit à l'initiative. Ainsi, nous considérons deux types d'utilisation possibles pour une technique d'interaction décrite (ou conçue) avec notre cadre d'analyse et voici la signification de ces deux types :

1. IN/FEEDBACK → l'utilisateur fait une action sur le système et le système lui envoie un feedback immédiat. Selon (Bastien and Scapin 1992), un « *feedback immédiat concerne les réponses de l'ordinateur consécutives et immédiates aux actions des utilisateurs quelles qu'elles soient. Ces réponses doivent renseigner sur l'action accomplie et le résultat. Le délai de réponse doit être homogène et approprié selon les types de transactions. Dans tous les cas, une réponse, aussi immédiate que possible, doit être fournie à l'utilisateur, le renseignant sur l'action accomplie et sur son résultat ».*
2. -/OUT → Signal porteur d'information destiné à un utilisateur, sans que celui-ci ne soit à l'origine du signal.  
Ex : une alerte, une notification, etc.

Le facteur type d'utilisation fait partie de notre cadre d'analyse.

#### **CATEGORIE DE TECHNIQUE : ACTIVE OU PASSIVE ?**

(Gibson 1962), dans son expérience « *the great cookie experiment* », introduit les deux concepts clefs : « active touch » et « passive touch ». Sa technique « passive touch » se base

sur la modalité sensorielle cutanée (moule à cookies pressés dans la main) et sa technique « active touch » repose sur la modalité sensorielle haptique (manipulation directe des moules à cookies).

En ce qui concerne les techniques passives, nous pouvons retrouver la notion de « modalité passive » dans l'espace de classification de (Roudaut and Lecolinet 2007). Toutefois, leur définition s'applique au point de vue système, alors que nous, nous sommes du point de vue humain. Ainsi, nous avons préféré la définition de (Turk and Robertson 2000) pour définir les techniques passives : « **Une interaction Homme-Machine est définie comme passive, si elle ne requière pas un contrôle explicite de la part de l'utilisateur pour exécuter une tâche** ». Le terme de contrôle explicite, définie par (Bastien and Scapin 1992), concerne à la fois la prise en compte par le système des actions explicites de l'utilisateur et le contrôle qu'a l'utilisateur sur le traitement de ses actions. C'est-à-dire que le système doit exécuter seulement les opérations demandées par l'utilisateur, aucune autre, et ce, au moment, où il les demande (action explicite). De plus, l'utilisateur doit toujours « avoir la main », c'est-à-dire de pouvoir contrôler le déroulement des traitements informatiques en cours (contrôle utilisateur, ex. : interrompre, reprendre, avancer, reculer, etc.).

En nous appuyant sur les définitions de « feedback immédiat » et de « contrôle explicite » de (Bastien and Scapin 1992), nous proposons une définition de l'interaction active : **Une interaction Homme-Machine est définie comme active, si elle requière un contrôle explicite de la part de l'utilisateur pour exécuter une tâche et qu'elle fournit un feedback immédiat (retour d'information) de l'interaction à l'utilisateur.**

Ainsi, avec cette définition, l'interaction de « commander un café à un distributeur » est une interaction passive : pas de contrôle explicite (*alea jacta est*). Elle serait active si par exemple, l'utilisateur fabriquait lui-même son café (ou son jus d'orange) car il aurait alors le contrôle explicite et un feedback immédiat de son action en cours.

Le facteur type d'interaction fait partie de notre cadre d'analyse.

**PRECISION SUR LES TYPES D'UTILISATION D'UNE INTERACTION PASSIVE ET ACTIVE**

Afin de lever toute ambiguïté sur les termes : active, passive et les types d'utilisation, le Tableau 3-3 combinatoire ( $\cap$ ) présente pour chaque combinaison, un exemple d'interaction.

$\cap$	IN/FEEDBACK	-/OUT
Passive	« cliquer sur un bouton pour commander un café »	« alarme/alerte »
Active	« manipuler le potentiomètre du volume audio d'une chaîne Hi-Fi »	? ?(voir légende)

**Tableau 3-3 : tableau combinatoire ( $\cap$ ) afin de lever toute ambiguïté sur les termes : active, passive et les types d'utilisation. Notons l'absence de proposition dans la case [-/OUT  $\cap$  Active], cette case peut être remplie par des interactions non connues des utilisateurs. Prenons l'exemple de l'interaction suivante : « l'utilisateur entre dans une pièce ; plus l'utilisateur va vers le fond de la pièce, plus l'éclairage sera amplifié ; et inversement » ; [??] Si l'utilisateur n'a pas conscience de cette interaction, alors pour lui et du point de vue conception, cela sera une [-/OUT  $\cap$  Active], par contre, s'il en a conscience cela sera une [IN/FEEDBACK  $\cap$  Active].**

**CATEGORIE DE TECHNIQUE : AVEC RÉFÉRENCES OU SANS RÉFÉRENCE ?**

Dans un contexte de transmission d'information et plus particulièrement dans le cas d'une estimation de valeur, la représentation interne de l'échelle des valeurs est primordiale. Cette représentation interne s'acquiert par l'apprentissage des relations entre les valeurs et la perception des stimuli. Le terme apprentissage est défini ci-après :

*Dans la psychologie inspirée du béhaviorisme<sup>37</sup> (ou du comportementalisme), l'apprentissage est : « la mise en relation entre un stimulus et une réaction adéquate du sujet, qui cause un changement de comportement qui est persistant, mesurable, et spécifique ou permet à l'individu de formuler une nouvelle construction mentale ou réviser une construction mentale préalable<sup>38</sup> ».*

Ainsi, dans notre contexte de transmission d'information, nos protocoles expérimentaux comprennent tous une étape d'apprentissage<sup>39</sup>, afin de permettre aux utilisateurs de se construire une représentation interne (mentale) de l'échelle des valeurs, avant les phases de test. Enfin, certaines études suggèrent qu'il est possible de renforcer cette représentation interne par l'ajout de « références explicites » dans le signal codant l'information. Prenons l'exemple dans la revue de (Fraisse 1984) de l'estimation d'une valeur codée avec la durée d'un son. La technique sans référence correspond simplement à l'envoi d'un stimulus auditif continu d'une durée variable (en fonction de la valeur à transmettre) vers l'utilisateur. Cette revue propose également, une version avec références de cette technique : « *il faut ajouter*

<sup>37</sup> Cf. John B. Watson, 1913 : concept clés : conditionnement classique ou opérant, et le renforcement

<sup>38</sup> <https://fr.wikipedia.org/wiki/Apprentissage>

<sup>39</sup> Cette étape, appelée phase de familiarisation, était systématiquement présente dans tous nos protocoles et avant la phase de test.

*explicitement un stimulus d'une durée de référence, juste avant l'envoi du stimulus avec une durée variable* ». L'une de leurs conclusions est que la présence de références explicites, juste avant ou après le signal variable, améliore la précision de l'estimation de la durée variable du stimulus. Cette conclusion apparaît également dans les travaux de (Mershon and King 1975), où l'estimation d'une distance, en utilisant la variation d'intensité d'une source sonore, comme indice perceptif, dépendait inexorablement de la connaissance préalable de l'intensité initiale de ladite source sonore (variation d'intensité relative à la distance et à l'intensité de la source sonore).

En nous appuyant sur leurs conclusions, nous définissons les termes « **technique avec références** » et « **technique sans référence** », comme **actant de la présence ou de l'absence de point de repères explicites dans la technique**, qu'ils soient physiques (une butée, un relief, une forme, etc.), temporels (signal avec une durée fixe ou périodique) ou abstraits (ex. earcons, tacton).

Le facteur Référence(s) fait partie de notre cadre d'analyse.

### 3.2.2 Conclusion sur l'extraction des facteurs de conception

Nous avons vu huit facteurs de base, tous extraits de propriétés générales sélectionnées de la littérature. Une technique d'interaction peut être décrite à travers ces facteurs. Cependant, il nous reste à organiser ces facteurs, dans un cadre d'analyse, sous la forme de paramètres.

### 3.3 Organisation des facteurs de conception dans notre cadre d'analyse

Nous avons choisi de représenter les facteurs de conception décrits dans la section précédente, dans un tableau de synthèse sous la forme de paramètres (Tableau 3-4), auxquels nous ajouterons quatre spécifications optionnelles.

#### LE TABLEAU DE SYNTHÈSE : HUIT PARAMÈTRES DE BASE

**Paramètre 1.** Nombre de références explicites : il y a cinq valeurs possibles à ce niveau de détail, « 0 » pour indiquer qu'il n'y a aucune référence explicite, « 1, 2, 3 » pour indiquer par exemple le début, le milieu et/ou la fin d'une plage de valeur et « + » pour indiquer qu'il y a plus de 3 références.

**Paramètre 2.** Catégorie de l'interaction : elle est soit passive, soit active ; en accord avec les définitions (cf. Catégorie de technique : Active ou Passive ?).

**Pamamètre 3.** Le type d'utilisation : modalité en entrée (E) ou en sortie (S), sachant que lorsque l'utilisateur exécute une interaction en entrée du système, le système doit lui fournir nécessairement un feedback immédiat (F).

**Pamamètre 4.** Zone(s) du corps : pour simplifier le tableau, nous avons seulement le sens de l'interaction au point de vue de l'utilisateur, i.e. « U<- » l'utilisateur perçoit, « U-> » l'utilisateur agit. Dans la description détaillée, il conviendra de nommer la ou les parties du corps utilisées.

**Pamamètre 5.** Média (ou dispositif) : pour simplifier le tableau, nous avons les trois possibilités : l'effecteur (Ef.), média capable de transmettre de l'information à l'utilisateur ; le capteur (Ca.), média capable de transmettre de l'information au système ; et l'objet tangible (O.T.) (cf. Modalité d'interaction tangible).

**Pamamètre 6.** Modalité : elle est soit Auditive ou soit Tactile, dans la version détaillée nous pourrions préciser quelle « sous-modalité » est utilisée (cf. Etat de l'art sur la transmission d'information basée sur les modalités tactiles et/ou auditives).

**Pamamètre 7.** Le type de signal : continu ou discret (cf. Différents types de signal)

**Pamamètre 8.** Le type de codage : Angle(A) / Longueur(L) / Durée(D) / Force(F) / Pulse(P) / Langage(Lg) (cf. Différents types de codages).

param1	Référence(s)	0, 1, 2, 3, + →	
param2	Type TI	Passive	Active
param3	Type d'utilisation	S	F E
param4	Zone(s) du corps	U<-	U->
param5	Média	Ef.	O, T. Ca.
param6	Modalité	Audio	Tactile
param7	Type Signal	Cont.	Dis.
param8	Type Codage	A L D F P Lg	

**Tableau 3-4 : tableau de synthèse composé de huit paramètres de base**

Ce tableau de synthèse permet de « mapper » une information. Pour l'utiliser, il suffit de tracer un trait, du haut vers le bas, en passant sur les bonnes valeurs (colonnes/cases) pour chacun des ses paramètres (lignes). La Figure 3-1, illustre deux exemples pour « mapper » une information.

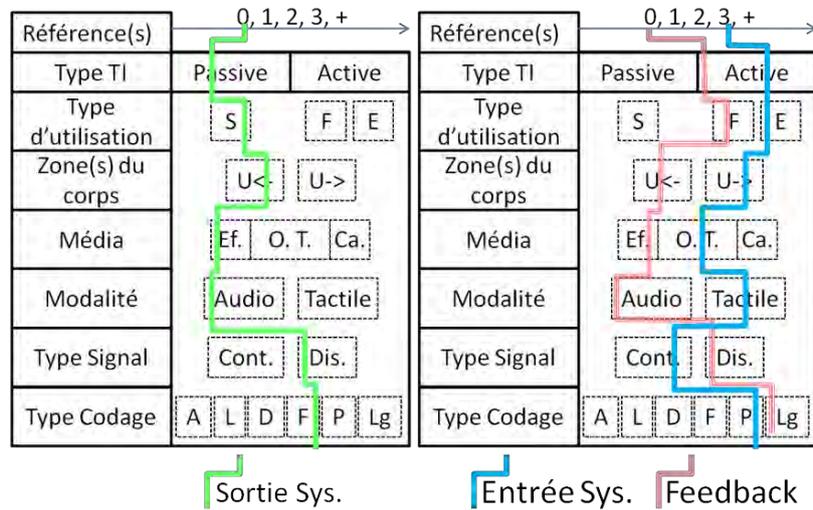


Figure 3-1 A gauche, une **INFORMATION MAPPÉE** par une « technique d'interaction sans référence, passive, signal en sortie du système vers l'utilisateur, avec un effecteur capable de transmettre un signal audio de manière discrète (ponctuelle) et dont la force (intensité) du signal code une information » ; à droite, une **INFORMATION MAPPÉE** par une « technique d'interaction avec 3 références et active, signal en entrée du système de la part de l'utilisateur, avec un objet tangible capable de transmettre un « signal tactile » de manière continue et dont les pulsations du signal codent une information ; l'utilisateur recevra un feedback de son action, avec un effecteur capable de transmettre un signal auditif discret (ponctuel) basé sur un langage oral naturel (parole) ou abstrait (earcons) ».

Ce tableau permet de décrire rapidement une technique transmettant de l'information. Toutefois, il est possible d'ajouter plusieurs spécifications afin de décrire plus précisément la technique.

**QUATRE SPECIFICATIONS : L'ÉCHELLE, LA TÂCHE UTILISATEUR, L'ASPECT TEMPOREL ET LE DISPOSITIF PARTAGE**

Nous pouvons utiliser le tableau de synthèse afin d'avoir une vue *synthétique* du mapping d'une information. Toutefois, cette vue synthétique peut être complétée afin de spécifier plus précisément une technique. Nous proposons d'ajouter quatre spécifications optionnelles, telles que l'échelle, la tâche, l'aspect temporel de la transmission d'information et la notion de « dispositif partagé » :

**L'échelle** est la relation liant un codage et une information, comme par exemple la relation ANGLE == VOLUME AUDIO. Nous pouvons également ajouter des unités à la relation : « 90° d'angle == 100 Niveaux de volume » ou une plage de valeur « ANGLE[0°-90°] == VOLUME[1-100] » (i.e. 9° équivaut à la valeur 10 par exemple). Nous pouvons également coder un message avec un langage (ex. TACTON == MESSAGE) ou un nombre (PULSATION == NOMBRE, où « X pulses » = nombre X).

**La tâche** permet d'en dire plus sur l'utilisation de la technique et contraint également les choix de conception. Par exemple, la tâche « estimer une DISTANCE entre 45 et 500 », contraint l'échelle à respecter la relation : CODAGE X == DISTANCE [45-500], reste à savoir quel CODAGE sera utilisé. Il peut y avoir des tâches plus complexe d'estimation, comme par exemple : « l'estimation d'un itinéraire composé de couples direction-distance ». Dans ce

cas, il faudra prévoir la transmission de chaque élément de l'itinéraire : au niveau des informations composant le couple (une direction et une distance) puis au niveau des composants de l'itinéraire (les couples). Cela nous amène également à traiter l'aspect temporel de la transmission des informations.

**L'aspect temporel** permet de spécifier quand les informations sont présentées à l'utilisateur : en séquence ou en parallèle et si elles sont indépendantes ou non (cf. Les propriétés de la multimodalité). Il y a quatre aspects temporels pour qualifier une transmission d'information via des techniques multimodales (Nigay and Coutaz 1996) :

1. **Concurrent** : transmission parallèle d'informations indépendantes (voir Figure 3-2-4).
2. **Alterné** : transmission séquentielle d'informations dépendantes. De plus, les informations dépendantes peuvent être soit redondantes, soit complémentaires (voir Figure 3-2-1).
3. **Synergique** : transmission parallèle d'informations dépendantes. De plus, les informations dépendantes peuvent être soit redondantes, soit complémentaires (voir Figure 3-2-2).
4. **Exclusif** : transmission séquentielle d'informations indépendantes (voir Figure 3-2-3).

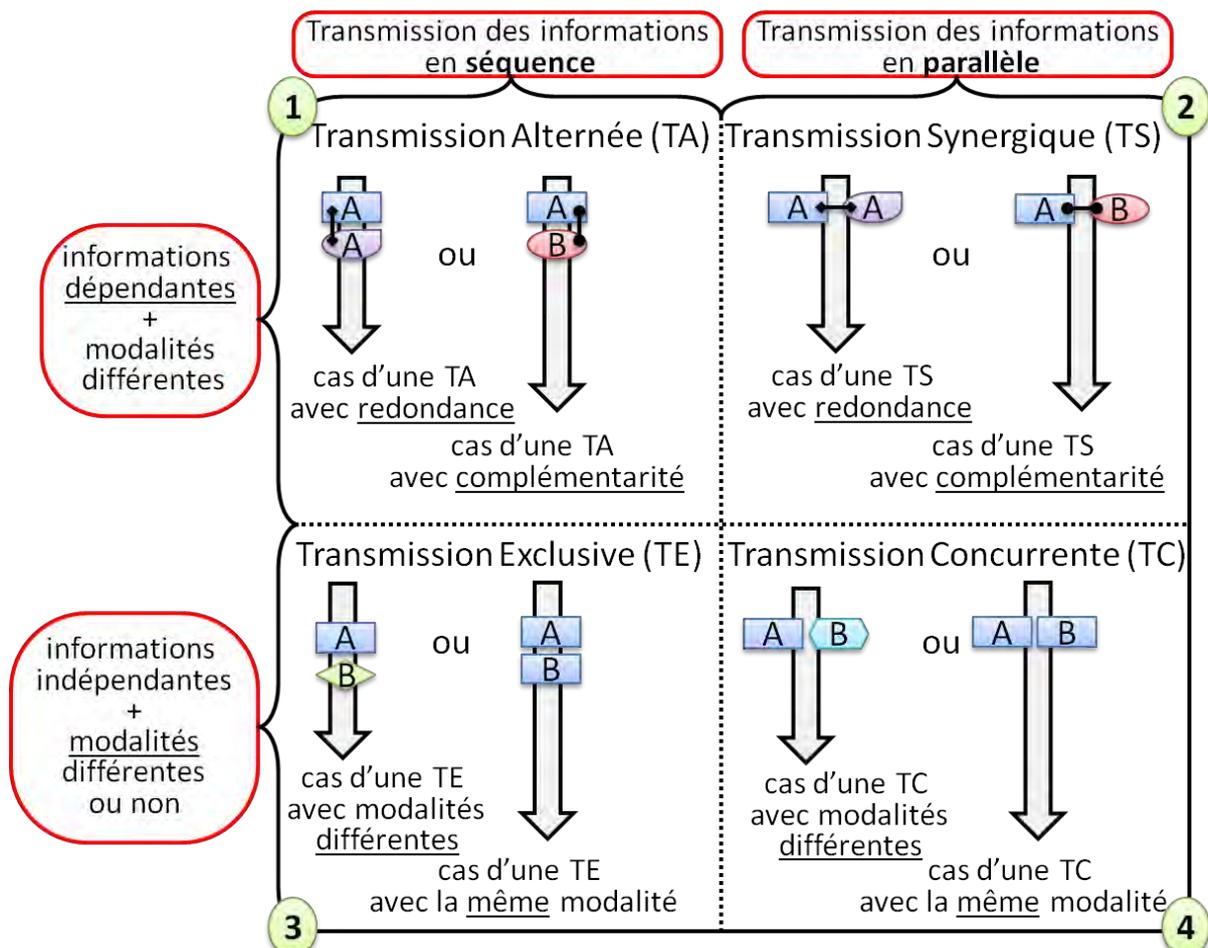


Figure 3-2 : illustration des propriétés CASE adaptées à notre contexte de transmission d'information(s). Les flèches représentent des lignes temporelles, les formes géométriques représentent des modalités (des techniques d'interaction) et les lettres représentent des informations. 1 et 3, sont des transmissions séquentielles ; 2 et 4 sont des transmissions parallèles ; 1 et 2 sont des transmissions d'informations dépendantes avec des modalités différentes ; 3 et 4 sont des transmissions d'informations indépendantes.

**Remarque** : il existe deux cas particuliers, le premier concerne les transmissions séquentielles d'informations dépendantes avec la même modalité, pour lesquelles nous avons choisi de garder le terme « transmission exclusive » afin de les qualifier. Ce qui nous amène au second cas particulier, les transmissions parallèles d'informations dépendantes avec la même modalité pour lesquelles nous avons choisi de garder le terme « transmission concurrente » afin de les qualifier. **Toutefois, les techniques d'interactions concernées par ces deux cas particuliers sont monomodales.**

Évidemment, pour la transmission d'une valeur la question de l'aspect temporel ne se pose pas. Mais l'aspect temporel devient essentiel pour la transmission d'un couple ou d'un ensemble de couples d'information. Nous verrons ces deux derniers cas dans les chapitres 5 et 6 de ce manuscrit.

**Le partage d'un ou des dispositifs** : il est possible qu'un dispositif soit utilisé pour transmettre plusieurs informations. Par exemple, un écran peut afficher plusieurs informations, un casque audio peut transmettre un son spatialisé (stéréo), etc. Lorsque plusieurs informations sont transmises avec le même dispositif, celui-ci est alors partagé.

### 3.3.1 Conclusion de l'organisation des facteurs de conception

Nous avons vu huit facteurs de base décrivant une interaction. Nous avons également ajouté au cadre d'analyse, quatre propriétés (facultatives) affinant la description : l'échelle, la tâche, l'aspect temporel et le partage d'un dispositif. L'utilisation graphique du cadre d'analyse permet, outre une représentation graphique, de concevoir rapidement une interaction et de la comparer à d'autres. La description détaillée de la représentation graphique permet d'ajouter quelques éléments spécifiques à la technique (ex. le dispositif, la zone du corps sollicitée, etc.).

Dans le cadre de nos travaux, sur la conception et l'évaluation de techniques d'interaction non visuelle optimisées pour de la transmission d'information(s), nous avons utilisé ce cadre d'analyse afin de générer toutes nos techniques. Mais en parallèle, nous avons étudié l'impact de certains facteurs de conception sur l'efficacité des techniques et la satisfaction des utilisateurs vis-à-vis desdites techniques. Notons que certains facteurs, comme la participation active d'un utilisateur ou la présence de références, ont été importés des domaines des sciences cognitives. Ils ont été jugés pertinents pour faire partie du processus de conception de techniques optimisées pour de la transmission d'information(s). Dans la

section suivante, nous utilisons notre cadre d'analyse pour générer nos techniques d'interaction par combinaison de facteurs.

### 3.4 Illustration de génération de techniques avec notre cadre d'analyse

Dans un contexte de transmission d'information, et pour une tâche d'estimation d'une valeur sans la vue, nous avons généré huit techniques d'interaction différentes. La génération de ces techniques a trois objectifs :

- ✚ Le premier est l'exploration de plusieurs combinaisons de facteurs pour concevoir des techniques d'estimation d'une seule valeur. Ainsi, nous avons fait varier trois facteurs pour concevoir nos techniques : la présence de référence, la participation de l'utilisateur et le type de codage. Deux facteurs issus des sciences cognitives et un facteur explorant trois types de codages liés aux dimensions de l'espace-temps : angle, longueur et durée. Ainsi, nous nous basons sur des codes « naturels » de notre environnement physique (espace-temps) et sur des principes issus du domaine des sciences cognitives.
- ✚ Le deuxième est l'étude des relations potentielles entre trois facteurs tels que, Référence(s) « Avec Réf ou Sans Réf », le type d'interaction « Passif ou Actif » et le type de codage « ANGLE, LONGUEUR ou DURÉE ». Ce second objectif permettra de tester l'influence de ces facteurs sur l'efficacité des techniques et la satisfaction des utilisateurs vis-à-vis desdites techniques. Puis, de connaître leurs limites et éventuellement de les hiérarchiser.
- ✚ Le troisième est de valider expérimentalement un panel de techniques efficaces, dans une tâche d'estimation d'une seule valeur, afin de concevoir des techniques capables de transmettre efficacement des information(s) complexe(s), comme par exemple, un couple de coordonnées pour localiser précisément une position ou une séquence de couples direction-distance pour faciliter la perception et la mémorisation d'un itinéraire.

Pour toutes ces techniques, nous avons fixé le facteur type de signal sur « signal continu » et l'échelle sur la relation « Codage X == 100 valeurs ». Nous avons également exploré plusieurs types de codage différents, ce qui fait de cette étude une exploration large bande. Nous avons conçu huit techniques de bases. Ces techniques se distinguent les unes des autres par leurs conception différentes : cinq sont actives, trois sont passives ; cinq sont

avec des références, trois sans références ; enfin quatre sont basées sur un codage ANGLE, trois sur un codage DURÉE et une sur un codage LONGUEUR. Le Tableau 3-5 synthétise les choix de conception de nos huit techniques.

Ti (ligne)-facteurs de conception (colonne)	Param. 1	Param. 2	Param. 3	Param. 4	Param. 5	Param. 6	Param. 7 « fixe »	Param. 8
Palmgauge	AR+	P	-/Sortie	Main	Cylindre+aiguille	Pression & contact	Continu	ANGLE
Singlestimulus	AR0	P	-/Sortie	Oreilles	Casque audio	Sonification	Continu	DURÉE
Dualstimuli	AR1	P	-/Sortie	Oreilles	Casque audio	Sonification	Continu	DURÉE
ClicOnMe	AR0	A	Entrée / Feedback	Main	Souris	Pression & contact	Continu	DURÉE
Pipe	AR0	A	Entrée / Feedback	Poignets	Cylindre	Blind reaching	Continu	ANGLE
Volant	AR3	A	Entrée / Feedback	Bras	Volant	Blind reaching	Continu	ANGLE
Pédale	AR2	A	Entrée / Feedback	Poignet	Pédale	Blind reaching	Continu	ANGLE
Slider	AR2	A	Entrée / Feedback	Main	Slider-phidget	Blind reaching	Continu	LONGUEUR

**Tableau 3-5 : synthèse présentant les choix de conception pour les huit techniques d'interaction. AR signifie AvecRéférences, 0,1,2,3 = le nombre de référence et le « + » signifie plus de 3.**

Dans les sections suivantes, nous présentons, de manière synthétique et structurée, les différentes techniques conçues, puis nous discutons le choix des techniques d'interaction retenues.

### 3.4.1 Présentation de techniques d'interaction non visuelle optimisées(?) pour transmettre une valeur

#### 3.4.1.1 Le Palm Gauge

**Principe** : similaire à celui d'un compteur de vitesse. La position de l'aiguille (angle) correspond à la valeur (ex. une distance). L'aiguille va se déplacer, sous la main de l'utilisateur, en suivant un cercle imaginaire de la gauche (0) vers la droite (100) (voir Figure 3-3 gauche). La plage de valeur est comprise entre 1 et 100. Les références sont décidées par l'utilisateur. En posant sa main sur l'objet, il repère physiquement plusieurs valeurs, par exemple, celles sous ses doigts ou entre ses doigts.

**Illustration** :

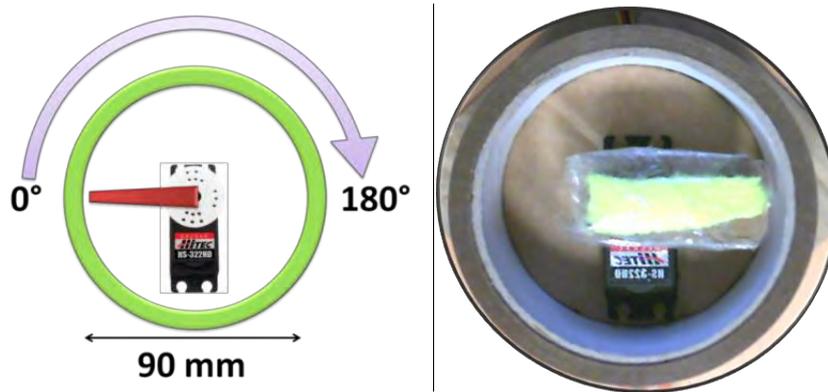


Figure 3-3 : illustration du prototype palm gauge, à gauche le concept, à droite le prototype vu de haut. Exemple d'utilisation : « Je perçois la rotation d'une aiguille sous ma main, elle s'arrête puis revient à sa position d'origine, j'estime la valeur transmise»

**Cadre d'analyse** : technique d'interaction passive, avec plus de trois références (doigts de l'utilisateur). Le signal, en sortie du système vers l'utilisateur, s'effectue avec un objet tangible capable de transmettre un stimulus tactile, de manière continue, et dont l'angle code une valeur entre 1 et 100 ( $180^\circ = 100$  valeurs) (voir Figure 3-4).

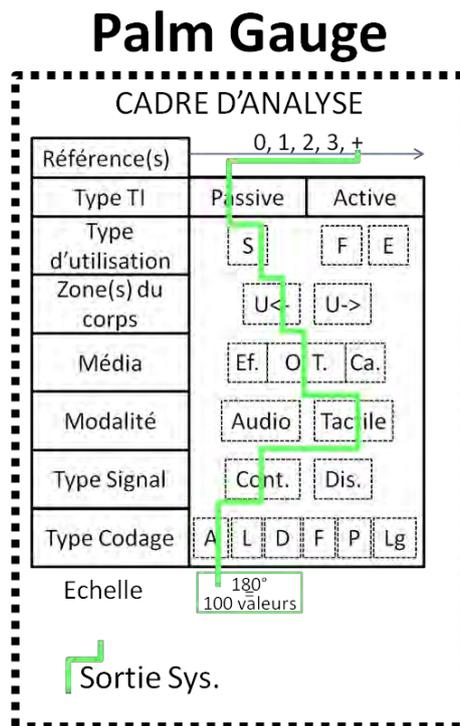


Figure 3-4 : illustration de l'analyse de la technique Palm Gauge.

**Matériel** : le Palm Gauge est constitué d'un servomoteur [HS-322HS Servo] relié à une carte Arduino [1061\_1 PhidgetAdvancedServo 8-motor]. Le support pour la main est un [cylindre en cartons] de 50 mm de hauteur, 90 mm de diamètre et de 5 mm d'épaisseur, dans lequel est fixé le servomoteur, en son centre. L'aiguille est faite à partir d'une [balle de tennis]

attachée au moyeu du servomoteur, avec du fil de fer et du ruban adhésif (voir Figure 3-3 droite).

### 3.4.1.2 Le Single Stimulus

**Principe :** la durée d'un bip -sonore- représente une valeur entre 1 et 100.

**Illustration :**



Figure 3-5 : Illustration de la technique « single bip ». Exemple d'utilisation : « J’entends un son pendant une certaine durée, puis plus rien. J’estime la valeur transmise »

**Cadre d’analyse :** technique d’interaction passive sans référence. Le signal, en sortie du système vers l’utilisateur, s’effectue avec un dispositif (effecteur) capable de transmettre un stimulus audio, de manière continue et dont la durée code une valeur entre 1 et 100 (1500 ms = 100 valeurs) (voir Figure 3-6).

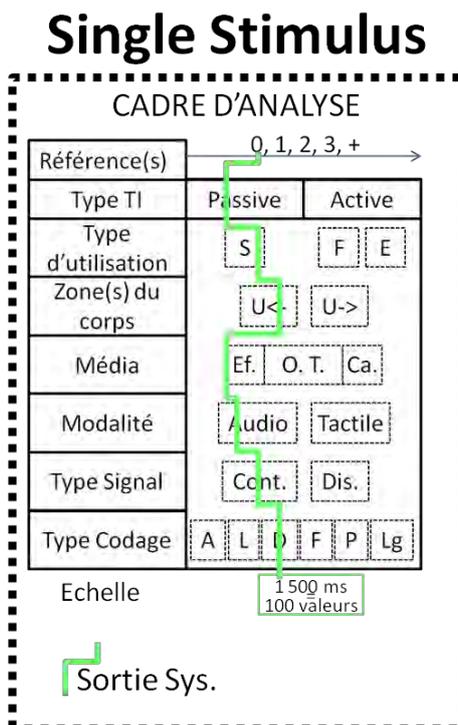


Figure 3-6 : illustration de l’analyse de la technique Single Stimulus.

**Matériel** : un casque audio [Plantronics audio 650 usb] (voir Figure 3-5). La formule pour traduire la durée en valeur est [durée = 15\*Valeur].

### 3.4.1.3 Le Dual Stimuli

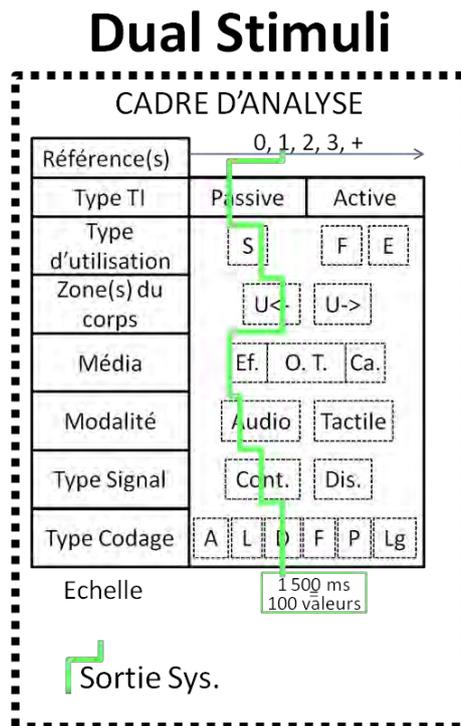
**Principe** : il y a deux bips. Le premier bip -invariable- d'une durée de 750ms, correspond à la valeur 50 et sert de référence. Juste après un court silence de quelques dizaines de millisecondes, est joué le second bip. Ce dernier bip a une durée variable et représente une valeur (ex. une distance) entre 1 et 100.

**Illustration** :



Figure 3-7 : illustration de la technique dual stimuli. Exemple d'utilisation : « je perçois un premier son d'une durée fixe, puis un second d'une durée variable, j'estime la valeur transmise par la partie variable ».

**Cadre d'analyse** : technique d'interaction passive avec une référence (son de durée fixe de 750 ms). Le signal, en sortie du système vers l'utilisateur, s'effectue avec un dispositif (effecteur) capable de transmettre des stimuli audio, de manière continue et dont la durée code une valeur entre 1 et 100 (1500 ms = 100 valeurs) (voir Figure 3-8).



**Figure 3-8 : illustration de l'analyse de la technique Dual Stimuli.**

**Matériel** : un casque audio [**Plantronics audio 650 usb**] (voir Figure 3-7). La formule pour traduire la durée en valeur est [durée = 15\*Valeur]. La durée totale d'un "Dual bip" est égale à la somme du premier bip, du silence, puis du second bip variable : [750 + 100 + 15\*Valeur] ms.

#### 3.4.1.4 Le ClicOnMe

**Principe** : similaire à un chronomètre. Une souris est tenue en main, à l'instar d'un chronomètre, et avec le pouce, le bouton du "chronomètre" est pressé pendant une certaine durée. Lorsque la durée est atteinte (ou dépassée), un feedback avverti l'utilisateur. L'utilisateur doit alors estimer la valeur (ex. une distance), en fonction de la durée qu'il a passée à appuyer sur le bouton.

**Remarque** : nous avons essayé de créer une affordance pour le dispositif. En effet, une souris n'est pas un chronomètre, mais la prise en main est similaire, et le bouton est sur le dessus. Il a été demandé aux utilisateurs, de prendre le dispositif comme un chronomètre, et d'appuyer sur l'un des boutons.

**Illustration** :



Figure 3-9 : Illustration du prototype "ClicOnMe" pris en main comme un chronomètre ; à gauche, le concept, à droite, le prototype tenu en main. Exemple d'utilisation : « J'appuis sur le bouton de la souris, j'attends. Je perçois un feedback au bout d'une certaine durée, j'estime la valeur transmise».

**Cadre d'analyse** : technique d'interaction active sans référence. Le signal, en entrée du système, s'effectue avec un objet tangible capable de transmettre un stimulus tactile de manière continue, et dont la durée code une valeur entre 1 et 100 (1500 ms = 100 valeurs). L'utilisateur recevra un feedback de l'accomplissement de son action, avec un effecteur capable de transmettre un signal auditif, discret basé sur une pulsation (ici, un bip) (voir Figure 3-10).

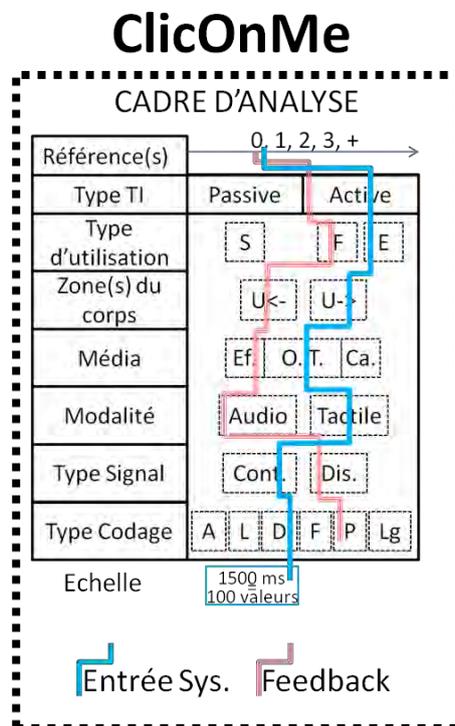


Figure 3-10 : illustration de l'analyse de la technique ClicOnMe.

**Matériel & Utilisation** : une [souris] (voir Figure 3-9). Il faut prendre la souris (le dispositif de pointage pour ordinateur) dans une main, et la tenir de manière à pouvoir, facilement, appuyer sur l'un de ses boutons. Il faut, ensuite, maintenir l'appui pendant un certain

temps. Tant que l'utilisateur appuie sur l'un des boutons de la souris, le "chronomètre tourne". Un feedback l'avertira quand la durée sera suffisante (valeur atteinte). Nous n'avons pas mis de « tick » pour représenter la seconde, car il s'agit d'une technique sans référence et que le temps maximal est de 1,5 secondes.

### 3.4.1.5 Le Pipe

**Principe** : il faut faire tourner un tube vers l'avant. L'utilisateur tient le tube avec ses deux mains et le fait pivoter vers l'avant. Lorsque l'angle de rotation est atteint (ou dépassé), un feedback averti l'utilisateur. L'utilisateur doit alors estimer la valeur (ex. une distance), en fonction de l'angle de rotation qu'il a appliqué au tube. Pour des raisons d'implémentation de la technique, un cube a été fixé sur le tube ; chacune de ses faces est reconnue par le système via une caméra et un logiciel d'analyse d'image. L'inclinaison des faces donne au système l'angle de rotation courant du tube.

**Illustration** :

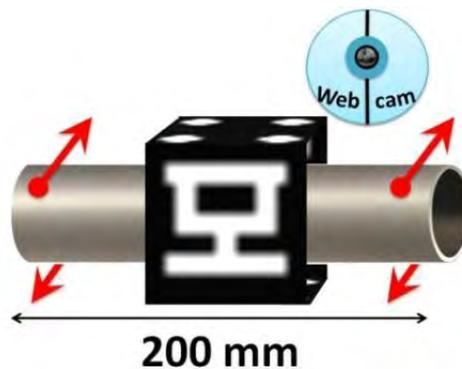


Figure 3-11 : illustration de la conception de la technique « pipe ». Exemple d'utilisation : « Je tourne doucement le cube vers l'avant. Lors de ma rotation, j'entends un bip, j'estime la valeur transmise ».

**Cadre d'analyse** : technique d'interaction active sans référence. Le signal, en entrée du système, s'effectue avec un objet tangible capable de transmettre un stimulus tactile, de manière continue, et dont l'angle code une valeur entre 1 et 100 ( $90^\circ = 100$  valeurs). L'utilisateur recevra un feedback de la fin de son action, avec un effecteur capable de transmettre un signal auditif discret, basé sur une pulsation (ici, un bip) (voir Figure 3-12).

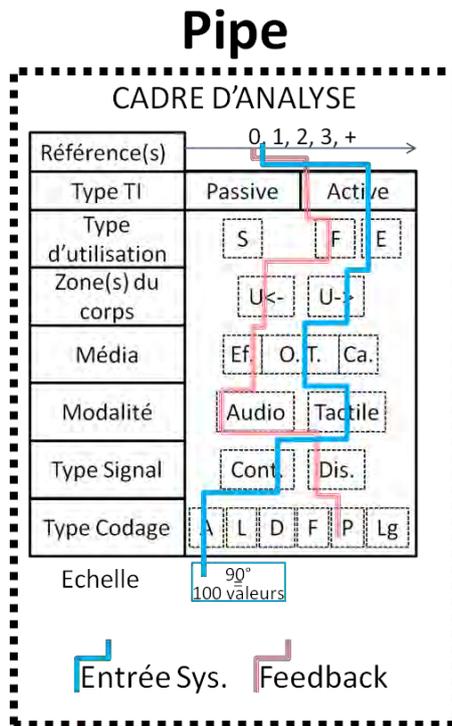


Figure 3-12 : illustration de l'analyse de la technique « Pipe ».

**Matériel :** Un [cube en polystyrène] traversé par un [tube en acier de 200 mm de long et de 20 mm de diamètre] (voir Figure 3-11). Les quatre faces -non percées- du cube sont couvertes par des [motifs] différents les uns des autres, pour permettre leur détection *via* une webcam [Logitech Quickcam pro 9000] et un logiciel de reconnaissance de forme basé sur la [library Nyartoolkitcs-4.1.1].

### 3.4.1.6 Le Volant

**Principe :** L'utilisateur doit tourner un volant vers la gauche ou vers la droite. L'angle de rotation du volant désigne une valeur. Lorsque l'angle de rotation est atteint (ou dépassé), un feedback averti l'utilisateur. L'utilisateur doit alors estimer la valeur (ex. une distance), en fonction de l'angle de rotation qu'il a appliqué au volant. Notons que le volant se remet en position initiale (0), lorsque l'utilisateur le lâche, ce qui fait de la position « zéro » une référence. De plus, le volant a deux butées physiques (gauche et droite) marquant la fin de la rotation maximale. L'utilisateur peut ainsi chercher ces butées, après avoir eu le signal de fin de tâche, et estimer par différence (angle max - angle courant) la valeur à estimer.

**Illustration :**



Figure 3-13 : illustration de la technique « volant ». Exemple d'utilisation « Je tourne doucement le volant à droite ou à gauche. Lors de ma rotation, j'entends un bip, j'estime la valeur transmise ».

**Cadre d'analyse** : technique d'interaction active avec trois références (position « zéro », Butée de Fin-gauche, Butée de Fin-droite). Le signal, en entrée du système, s'effectue avec un objet tangible capable de transmettre un stimulus tactile, de manière continu, et dont l'angle code une valeur entre 1 et 100 (90° = 100 valeurs). L'utilisateur recevra un feedback de l'accomplissement de son action, avec un effecteur capable de transmettre un signal auditif discret, basé sur une pulsation (ici, un bip) (voir Figure 3-14).

## Volant

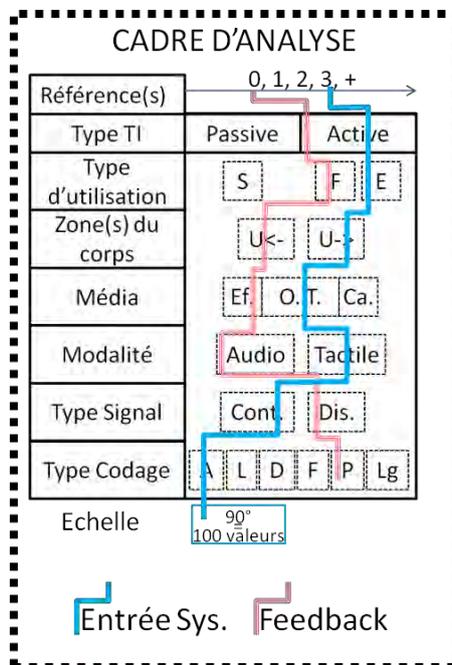


Figure 3-14 : illustration de l'analyse de la technique Volant.

**Matériel** : le Game Wheel [Logitech precision vibration feedback wheel], voir Figure 3-13.

### 3.4.1.7 La Pédale

**Principe** : L'utilisateur doit appuyer sur une pédale comme il le ferait avec son pied. L'angle d'inclinaison de la pédale désigne une valeur. Lorsque l'angle d'inclinaison est atteint (ou dépassé), un feedback averti l'utilisateur. L'utilisateur doit alors estimer la valeur (ex. une distance), en fonction de l'angle d'inclinaison, qu'il a appliqué à la pédale. Notons que la pédale se remet en position initiale (0), lorsque l'utilisateur la lâche, ce qui fait de la position « zéro » une référence. De plus, la pédale a une butée physique marquant la fin de l'inclinaison maximale. L'utilisateur peut ainsi chercher la butée, après avoir eu le signal de fin de tâche, et par différence, (angle max- angle courant) estimer la valeur.

**Illustration** :

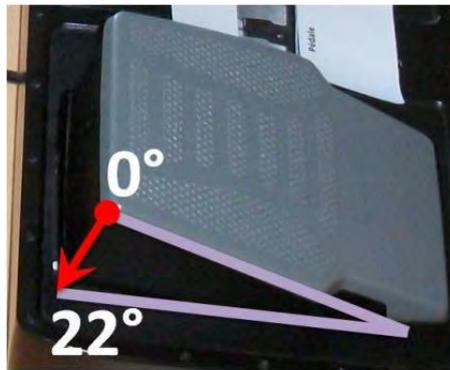


Figure 3-15 : illustration de la technique Pédale. Exemple d'utilisation : « J'incline doucement la pédale vers le bas. Lors de mon inclinaison, j'entends un bip, j'estime la valeur transmise ».

**Cadre d'analyse** : technique d'interaction active avec deux références (position « zéro » et butée de fin). Le signal, en entrée du système, s'effectue avec un objet tangible capable de transmettre un stimulus tactile, de manière continu, et dont l'angle code une valeur entre 1 et 100 ( $22^\circ = 100$  valeurs). L'utilisateur recevra un feedback de l'accomplissement de son action, avec un effecteur capable de transmettre un signal auditif discret, basé sur une pulsation (ici, un bip) (voir Figure 3-16).

# Pédale

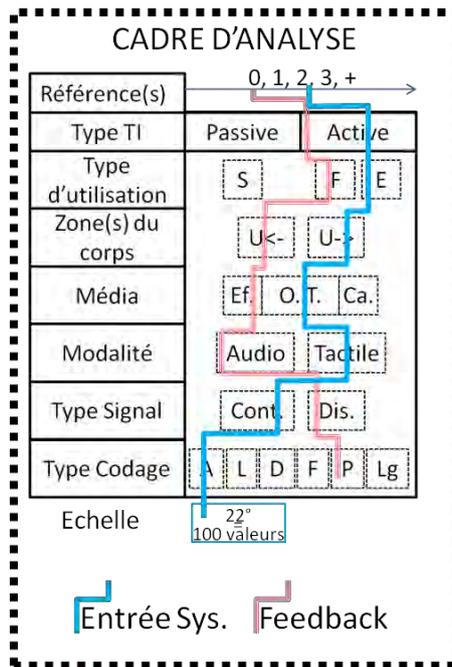


Figure 3-16 : illustration de l'analyse de la technique Pédale.

**Matériel** : le game pedal [Logitech precision vibration feedback pedal], voir Figure 3-15.

### 3.4.1.8 Le Slider

**Principe** : Le déplacement du «picot» de la position initiale crée un écart (longueur). La longueur correspond à la valeur (ex. une distance). Lorsque la valeur est atteinte (ou dépassée), un feedback averti l'utilisateur. La plage de valeur est comprise entre 1 et 100. De plus, le Slider a deux butées physiques, l'une marquant la position initiale (0), et l'autre marquant la position maximale (100). L'utilisateur peut ainsi chercher les butées, après avoir eu le signal de fin de tâche, et par différence, estimer la valeur.

**Illustration** :



Figure 3-17 : illustration de la technique « Slider ». Exemple d'utilisation : « Je déplace doucement le slider vers l'extrémité opposée du rail. Lors de mon déplacement, j'entends un bip, j'estime la valeur transmise ».

**Cadre d'analyse** : technique d'interaction active avec deux références (butée de début de rail, butée de fin de rail). Le signal, en entrée du système, s'effectue avec un objet tangible capable de transmettre un stimulus tactile, de manière continue, et dont la longueur code une valeur entre 1 et 100 (54 mm = 100 valeurs). L'utilisateur recevra un feedback de l'accomplissement de son action, avec un effecteur capable de transmettre un signal auditif discret, basé sur une pulsation (ici, un bip) voir Figure 3-18.

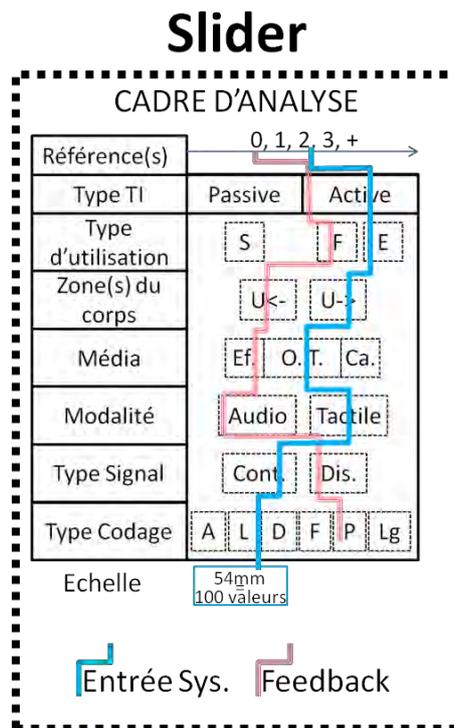


Figure 3-18 : illustration de l'analyse de la technique Slider.

**Matériel** : slider [1112\_1 - Slider 60] (voir Figure 3-17). Une petite précision pour le matériel [1112\_1 - Slider 60], la longueur maximale du rail est de 60 mm, mais la connectique du dispositif commence à partir de 6 mm, donc la valeur du slider change à partir de 6 mm du départ. Pour résoudre le problème, nous avons mis une butée de 6 mm au début du slider. Ainsi les 100 valeurs sont codées sur 54 mm seulement.

### 3.4.2 Discussion et conclusion sur la génération des techniques

#### 3.4.2.1 Synthèse des solutions explorées

La Figure 3-19 ci-dessous montre les huit techniques selon les facteurs « technique passive » et « technique active ». Chaque trait décrit une technique, et traverse toutes les lignes, et seulement une colonne (case) par ligne. Les cases « non traversées » montrent clairement les possibilités non explorées ici. Cependant, d'autres travaux ont déjà exploré certaines de ses possibilités non étudiées/explorées :

- Pour les techniques passives :
  - avec du son utilisant un angle : (Hartmann and Raked 1989; Mills 1958), dans une tâche d'estimation de l'azimut d'une source sonore.
  - avec du son utilisant une longueur : (Mershon and Bowers 1979; Zahorik 2002), dans une tâche d'estimation de distance égocentrique, à l'aide d'une source sonore dont la distance avec l'utilisateur varie.
  - avec du son utilisant une force : (Coleman 1963; Stevens 1957), dans une tâche d'estimation de valeur où l'intensité du stimulus varie.
  - utilisant un langage : les earcons (Garzonis *et al.* 2008) dans une tâche d'association de son à des commandes ou les tactons (Brunet *et al.* 2013)(Brewster and Brown 2004)
  - un exemple d'une autre possibilité : l'utilisation des battements de cœur d'un utilisateur (U->, passif)
- Pour les techniques actives :
  - avec du tactile utilisant la force : (Witt *et al.* 2004), dans une tâche d'estimation d'une distance ;
  - un exemple d'une autre possibilité : l'utilisation de la voix d'un utilisateur avec un capteur (ex. micro).

Il apparaît très clairement les possibilités explorées et leurs différences (voir Figure 3-19). Les différences entre les techniques s'articulent autour de trois facteurs principaux : le type d'interaction (active ou passive), le nombre de références et la modalité (auditive ou tactile). Nous pouvons également voir que toutes les techniques utilisent un signal continu, mais aussi que toutes les techniques actives utilisent le même feedback : un bip audio discret sans référence. Dans le chapitre suivant, nous présenterons une expérimentation visant à évaluer et à comparer ces techniques, dans une tâche d'estimation de valeur.

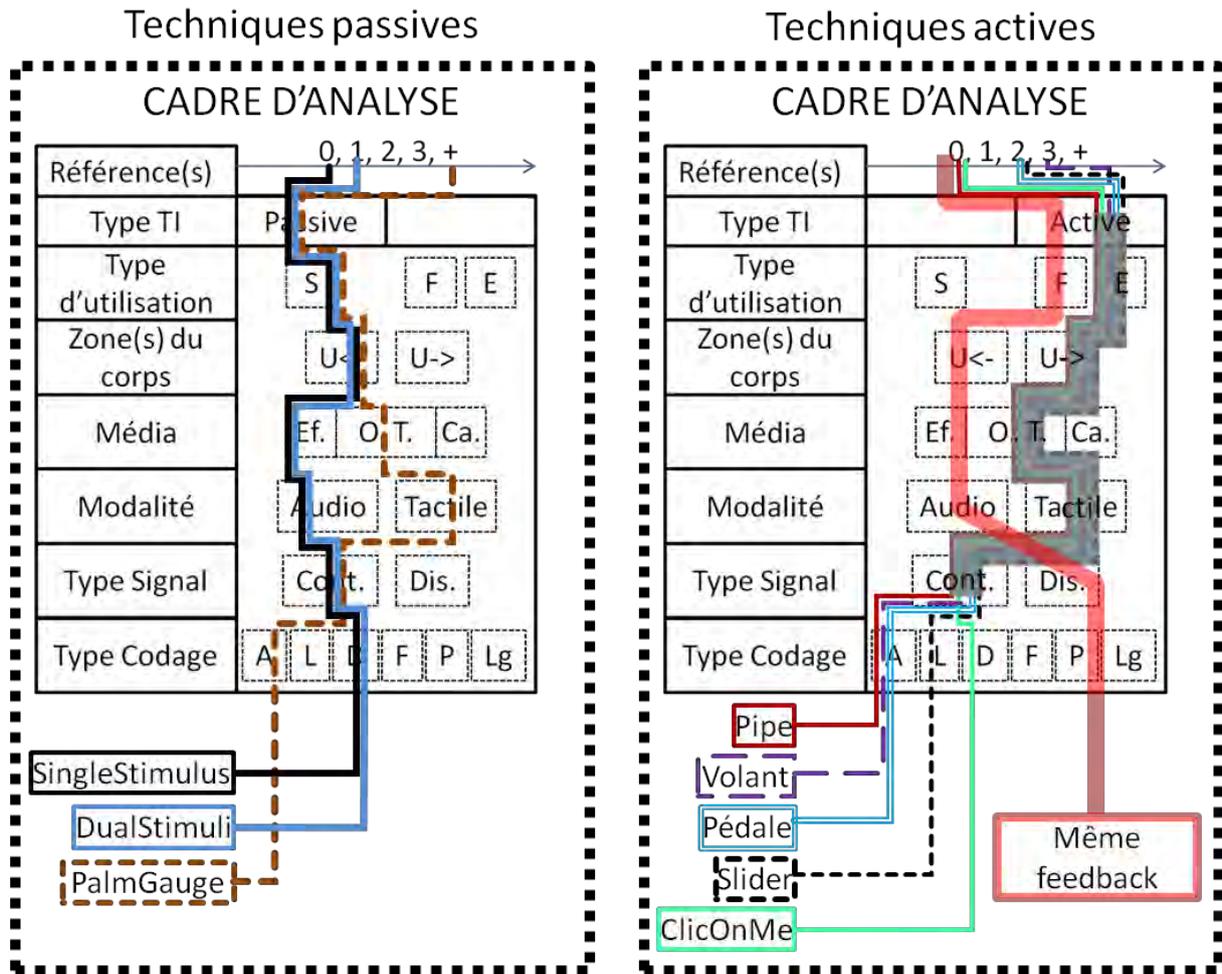


Figure 3-19 : Récapitulatif des huit techniques d'interaction générées. A gauche les passives, à droite les actives. Les traits représentent les designs de conception des techniques ; le trait épais met en valeur le fait que les techniques actives ont un design similaire sauf pour la présence de référence et le type de codage ; les cases non traversées représentent les possibilités de conception non explorées ici ; nous pouvons également voir les facteurs communs et spécifiques aux techniques.

### 3.4.2.2 Limitations

Nous avons choisi de ne pas utiliser le codage FORCE, car il existe de nombreux travaux sur des tâches de « magnitude estimation », notamment ceux de (Stevens 1957) qui ont relié l'intensité physique d'un stimulus à la perception de celui-ci par l'Homme. Ces relations sont regroupées dans ce qui est appelé la « Stevens' power law »,  $S = k.I^a$  (où l'exposant de Stevens (<sup>a</sup>) est lié à l'intensité (I) de la stimulation avec k une constante et où S est l'intensité de la perception). Le Tableau 3-6 ci-après montre quelques unes des ces lois :

Exposants de la loi de puissance de Stevens

Sensation	Exposant	Nom de l'unité	Publication
Sonie	0,3	Sone	Stevens 1957 d'après Piéron 1959
Hauteur (musique)	—	Mel	Stevens, Volk et Newman
Luminosité (blancheur)	1,2	—	Stevens 1957 d'après Piéron 1959
Luminosité (sensation lumineuse)	3,3	—	Stevens d'après (Le Grand 1972) et Piéron 1959, p. 8
Épaisseur	1,30	—	Stevens, non publié, d'après Piéron 1959
Lourdeur	1,45	Veg	Stevens 1957, d'après Piéron 1959
Choc électrique	3,5	Samp	Stevens 1957, d'après Piéron 1959

Tableau 3-6 : lois pour diverses sensations/perceptions :  $S = k.I^n$ , (voir [Loi de Stevens](#)).

Dans cette étude, nous n'avons pas utilisé le codage de type LANGAGE, comme par exemple la parole, les earcons, les tactons, etc., car nous voulions nous abroger, le plus possible, des limites liées aux langages (ex. barrière de la langue, apprentissage, etc.) d'une part, et d'autre part, nous souhaitions évaluer les limites de la perception humaine, spécifiquement, pour la capacité à estimer une valeur entre 1 et 100 à partir de stimuli sonifiés (ex. DUREE d'un son == 100 valeurs), vibrotactiles (DUREE d'une vibration == 100 valeurs) et haptiques (DUREE d'un effort physique, mouvements ANGLES, mouvements LONGUEURS == 100 valeurs).

Enfin, nous n'avons pas utilisé d'échelle basée sur le nombre où par exemple, un nombre de pulsations encode une valeur (PULSATION == NOMBRE, où « X pulses » = nombre X). De plus, la conception et l'évaluation d'une technique, où les utilisateurs doivent compter entre 1 et 100 pulses, dans une tâche d'estimation d'une valeur entre 1 et 100, ne nous a pas semblé être une contribution pertinente dans notre contexte conception de technique d'interaction non visuelle et optimisée pour transmettre une information de type valeur entre 1 et 100.

### 3.4.2.3 Précision sur les techniques basées sur le type de codage ANGLE

Nous avons quatre techniques utilisant le codage ANGLE, mais avec des échelles différentes. Cela dépend des objets tangibles utilisés dans les techniques, mais aussi des parties du corps requises pour manipuler lesdits objets : 180° pour un compteur tactile sous la main (palm gauge), 22° pour une pédale (de simulateur de conduite), 2x90° pour un volant et 90° pour un cylindre (« pipe »). Le palm gauge a 180° d'amplitude et va de la gauche vers la droite. La pédale a 22° d'amplitude et va de haut en bas. Le volant a 180° d'amplitude, mais pour des raisons d'habitudes, nous avons fait en sorte que les utilisateurs puissent tourner le volant soit vers la droite soit vers la gauche, ainsi le volant a 2x90° d'amplitude. Le « pipe » a été

bridé à 90°, car nous voulions que les utilisateurs effectuent un mouvement de rotation vers l'avant, sans qu'ils aient besoin de lâcher le tube pendant la tâche. De plus, un mouvement de 90° vers l'avant, avec les poignets et les doigts, est possible sans forcer sur les articulations.

#### 3.4.2.4 Précision sur les techniques basées sur le type de codage DURÉE

En nous appuyant sur les travaux de (Fraisie 1984), nous avons souhaité mettre en application l'estimation d'une DURÉE. La première technique est donc la plus simple, un son est joué pendant un certain temps et l'utilisateur doit traduire cette durée en une valeur entre 1 et 100. En nous basant sur (Fraisie 1984), nous savons que la perception d'une durée est optimale lorsque celle-ci se situe entre 2 et 3 secondes et indiscernable pour une durée inférieure à 100 ms. De plus, dans le cas d'une technique avec référence (Fraisie 1984) recommande de donner une référence juste avant ou après le stimulus variable, afin de mieux pouvoir percevoir sa durée. Afin de tester cette recommandation, nous avons donc augmenté la difficulté en mappant 100 valeurs dans 1 500 ms (15 ms = la valeur 1). Notons que cela rend théoriquement indiscernable les valeurs de 1-6 car elles sont mappées par des sons inférieurs à 100ms. Ainsi, nous avons conçu deux techniques : un son simple et un duo de sons. Dans le son simple, le stimulus varie entre 0 et 1 500 ms. Dans le duo de sons, l'un est fixe à 750 ms, comme une référence de valeur 50, et l'autre est variable (entre 1 et 100). Enfin, nous avons également conçu une troisième technique, où le stimulus sonore était remplacé par un stimulus proprioceptif ; ici cela signifie que l'utilisateur appuie consciemment sur un bouton pendant un certain temps (même codage que le son simple).

Ainsi, nous avons défini trois techniques :

- ✚ SingleStimulus : un son simple, 100 valeurs = 1 500 ms ;
- ✚ DualStimuli : un son de référence = valeur 50 = 750 ms, joué juste avant le son variable ;
- ✚ ClicOnMe : où l'utilisateur estime le temps qu'il passe à appuyer sur un bouton avant un feedback, 100 valeurs = 1 500 ms.

### 3.5 Conclusion sur l'apport du cadre d'analyse pour concevoir nos techniques

Dans notre contexte de transmission d'information, ce cadre d'analyse organise le champ immense des possibilités de conception, en regroupant les propriétés pertinentes en huit

facteurs de conception. Les choix de conception des techniques sont décrits dans deux niveaux de détails : le premier se veut simple et s'illustre par une description graphique ; le second est textuel et s'appuie sur la version graphique, tout en y ajoutant des détails. La description graphique de plusieurs techniques, pouvant se superposer, permet ainsi de les comparer, rapidement, et en même temps. La forme textuelle de la description permet une vision plus détaillée des techniques.

Ce cadre d'analyse est notre première contribution. Nous l'avons utilisé pour générer huit techniques. Nous sommes allés plus loin en évaluant et en comparant certains facteurs, afin de tester leur influence sur l'efficacité des techniques. Ceci fait l'objet de notre seconde contribution.

Notre seconde contribution est d'avoir conçu et évalué un panel de plus d'une vingtaine de techniques dans des tâches de perception et d'estimation. Ce panel nous a servi de base pour trois expérimentations, visant à évaluer des techniques d'interaction non visuelle optimisées pour de la transmission d'information(s) et ce, dans trois cas d'utilisation distincts :

- Cas 1. Perception et estimation d'une valeur, chapitre 4 ;
- Cas 2. Perception et estimation d'une séquence de couples de valeurs (direction-distance) et mémorisation-restitution des informations, chapitre 5 ;
- Cas 3. Perception et estimation d'un couple de valeurs (coordonnées) et intégration-localisation d'une position cible dans une portion d'espace, chapitre 6.

Ce chapitre a décrit les techniques de base, utilisées dans le Cas1 (Première étude : techniques d'interaction non visuelle pour estimer une valeur), et conçues en s'appuyant sur notre cadre d'analyse pour des techniques de transmission d'information de type une valeur. Dans le Cas2 (Deuxième étude : interactions non visuelles pour faciliter le rendu d'une séquence de couples direction-distance) et le Cas3 (Troisième étude : assistance tactile à la localisation de cibles périphériques pour des personnes à vision tubulaire), nous montrerons comment utiliser ce cadre d'analyse, afin de transmettre un couple de valeurs et une séquence de couples de valeurs. Ces trois chapitres montrent aussi comment tester spécifiquement certains facteurs de conception et reconnaître leur influence sur l'efficacité des techniques dans différentes tâche d'estimation.

# **Chapitre 4**

## **Première étude :**

**techniques d'interaction non  
visuelle pour estimer une  
valeur**



## 4 Première étude : techniques d'interaction non visuelle pour estimer une valeur

Nous avons généré avec notre cadre d'analyse, huit techniques de base d'interaction non visuelle que nous avons décrites dans le chapitre précédent. Ce chapitre présente notre première expérimentation (exp. 1.1) visant à évaluer ces huit techniques et leurs alternatives (techniques conçues à partir d'une des techniques de base et présentées dans ce chapitre) et à observer l'impact de trois facteurs de conception (techniques actives ou passives, présence ou non de références, type de codage ANGLE, LONGUEUR ou DURÉE) sur l'efficacité des techniques et la satisfaction des utilisateurs. A la suite de cette expérience, nous avons souhaité ordonner les facteurs en termes de « puissance d'impact », c'est-à-dire, lequel parmi les trois facteurs améliore-t-il le plus l'efficacité des techniques et la satisfaction des utilisateurs. Pour ce faire, nous avons mené une étude complémentaire (exp. 1.2) présentée à la suite de notre première expérimentation.

### 4.1 Motivation

Dans l'état de l'art (cf. chapitre 2), nous avons présenté un ensemble de techniques permettant de transmettre une ou des informations par des interactions non visuelles. Cependant, il est difficile de comparer entre elles ces techniques en termes d'efficacité et de satisfaction car elles ont été évaluées dans des conditions différentes ; notamment la plage de valeurs et le type d'information sont rarement les mêmes d'une technique à une autre. Par exemple, les techniques d'interaction de (Fujimoto and Turk 2014) transmettent un couple direction-distance où la direction était un angle entre  $0^\circ$  et  $180^\circ$  et la distance était une valeur entre 0 et 1, tandis que les techniques d'interactions de (Hajnal *et al.* 2011) transmettent une inclinaison parmi seize, réparties dans un intervalle  $0^\circ$ - $90^\circ$ . La technique de (Ellard and Shaughnessy 2003) permet d'estimer une distance en marchant entre 4 et 10 mètres. Ainsi, nous avons décidé d'évaluer nos techniques dans les mêmes conditions, et c'est pourquoi, dans cette première étude, nous comparons dix-neuf techniques d'interaction dans une même tâche d'estimation d'une valeur entre 1 et 100.

De plus, nos techniques sont conçues à partir de notre cadre d'analyse, dont trois facteurs cités précédemment ainsi qu'un quatrième facteur : la modalité sensorielle auditive ou vibrotactile. Ainsi, nous avons étudié, en parallèle, l'impact de quatre facteurs de conception sur l'efficacité des techniques et la satisfaction des utilisateurs, ainsi que l'impact de quatre combinaisons de facteurs (exp. 1.2). *In fine*, nous avons classé les techniques par puissance

d'impact et avons analysé pourquoi elles présentaient des différences entre elles, en expliquant quels facteurs ou combinaison de facteurs étaient en cause.

Le paragraphe suivant illustre des scénarios d'utilisation pour nos techniques non visuelles.

#### **SCENARIOS D'UTILISATION DE TECHNIQUES D'INTERACTION NON VISUELLE**

De toute évidence, il y a beaucoup de situations auxquelles, les gens ayant une déficience visuelle sont confrontés, et où les techniques d'interaction non visuelle leur seraient très utiles. Il pourrait être pertinent dans le contexte de la mobilité, de fournir à l'utilisateur des alertes pour transmettre la position d'un danger (Fujimoto and Turk 2014). Il pourrait également être utile dans un contexte d'exploration de transmettre le plan d'un bâtiment ou d'un quartier. Ces techniques peuvent être intégrées à des cartes interactives tactiles (Brock 2013; Brock *et al.* 2015) et servir de fonctionnalité avancée pour fournir à l'utilisateur des repères de distance discrets et privés entre les différents points de repère de la carte.

Mais les personnes voyantes peuvent également profiter de ces techniques dans des contextes où la vue n'est pas utilisable. Il a été démontré, par exemple, que les pompiers et les personnes aveugles peuvent nécessiter des techniques d'orientation non visuelle similaires (Reaudin and Gilliéron 2011). Des tâches, où la vision est déjà très monopolisée (conduite, pilotage, etc.) représentent un autre contexte tout à fait approprié. En effet, les techniques « eyes-free » peuvent fournir à l'utilisateur des indices de distance très utiles, concernant son environnement, sans avoir à changer son regard. Par exemple, pendant l'exploration d'une carte par des personnes voyantes. Avec la quantité croissante de données disponibles, les cartes sont de plus en plus utilisées comme un support structuré de visualisation d'information contextualisée : photos prises par les visiteurs, un restaurant, des itinéraires, événements en cours, rapport de dommages ou d'accidents, etc. Par conséquent, l'affichage visuel de toutes ces données simultanément devient impossible ou du moins, surcharge l'affichage (Keim *et al.* 2008). Enrichir une carte interaction avec des modalités non visuelles, devrait fournir à l'utilisateur différents types d'informations, d'une manière synergique (Trocka-Leszczynska and Jablonska 2013).

La section suivante présente la première partie de l'expérimentation.

### **4.2 Expérience 1.1 : techniques d'interaction non visuelle pour estimer une valeur**

Dans le chapitre 3, nous avons présenté nos huit techniques de base conçues pour étudier trois facteurs de conception : <Active/Passive> (A/P), <AvecRéférences/SansRéférence>

(AR/SR) et type de codage <ANGLE/LONGUEUR/DUREE>. Dans la section suivante, nous présentons les sept techniques alternatives afin de présenter notre quatrième facteur : modalité <Auditive/Vibrotactile>.

#### 4.2.1 Huit techniques et sept « alternatives »

Celles que nous appelons « techniques alternatives », utilisent une modalité alternative à leur technique de base. Par exemple, un stimulus auditif d'une certaine durée, peut-être remplacé par un stimulus vibrotactile d'une certaine durée. De même, un feedback auditif, peut-être remplacé par un feedback vibrotactile. Le Tableau 4-1 présente les différences entre les techniques de base et leur alternative. Puis, la Figure 4-1, montre ces différences avec la description graphique des techniques au travers de notre cadre d'analyse. Notons que seule, la technique PalmGauge n'a pas d'alternative, sa version auditive s'approchant trop des techniques d'estimation d'un azimut qui sont détaillées dans notre état de l'art (cf. L'azimut).

<b>Récapitulatif des 8 TI et des 7 TI alternatives</b>	
<b>5 TI Haptiques + feedback sonifié</b>	<b>5 TI alt. vibrotactile</b>
ClicOnMe ; Slider ; Pipe ; Volant ; Pédale	feedback vibrotactile
<b>2 TI avec sonification</b>	<b>2 TI alt. vibrotactile</b>
Bip continu ; Dual bip	Vibre continu ; Dual Vibre
<b>1 TI tangible</b>	
Palm Gauge	

**Tableau 4-1 : Synthèse des différences entre les techniques de base et leur alternative. Notons que seule PalmGauge n'a pas d'alternative.**

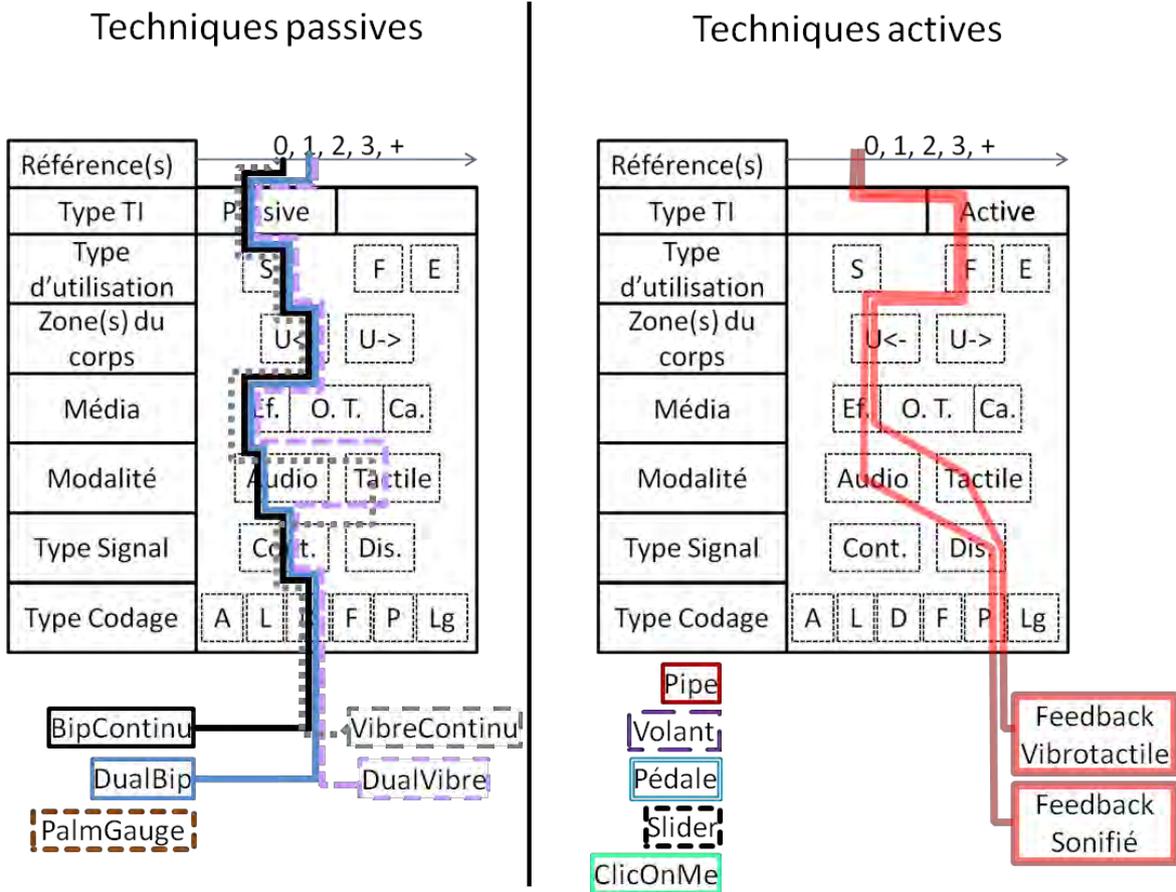


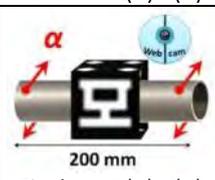
Figure 4-1 : illustration des différences entre les techniques de base et leur alternative. A gauche les « Passive », la différence se situe au niveau de la modalité, à droite les « Active », la différence se situe au niveau de la modalité du feedback.

**MATERIEL ET CALIBRATION SPECIFIQUE**

Pour les techniques alternatives utilisant la modalité vibrotactile, le dispositif est un bracelet avec un [Moteur Vibreur en Disque VPM2 Solarbotics] dont la fréquence vibration était fixée à 250 Hz, ce qui est la fréquence optimale pour la perception d'une vibration chez l'Homme (Makous *et al.* 1995).

Le Tableau 4-2 suivant récapitule les techniques.

**TABLEAU DE SYNTHÈSE DES TECHNIQUES**

Récapitulatif des techniques d'interactions utilisées dans cette expérience. Nous y retrouverons pour chaque technique : le matériel nécessaire, les modalités utilisées, la plage de valeur des grandeurs physiques (angle, longueur, durée), une courte description illustrée et leur appartenance à la catégorie passive ou active.		
Techniques passives	Single Stimulus	 <ul style="list-style-type: none"> <li>- Matériel : casque audio ou bracelet vibrant</li> <li>- Stimulation : auditive ou vibrotactile</li> <li>- Distance (1-100) = durée d'un seul signal (15-1500 ms)</li> </ul>
	Dual Stimulus	 <ul style="list-style-type: none"> <li>- Matériel : casque audio ou bracelet vibrant</li> <li>- Le premier stimulus est la référence (distance 50=750 ms)</li> <li>- Stimulation : auditive ou vibrotactile</li> <li>- Distance (1-100) = durée d'un seul signal (15-1500 ms)</li> </ul>
	Palm Gauge	 <ul style="list-style-type: none"> <li>- Matériel : Aiguille qui tourne via un servomoteur</li> <li>- Stimulation : tactile</li> <li>- Distance (1-100) = la position de l'aiguille sous la main (1=1,8° ; 100=180°)</li> </ul>
Techniques actives	ClicOnMe	 <ul style="list-style-type: none"> <li>- Matériel : Souris avec un bouton</li> <li>- Distance (1-100) = durée de la pression du bouton jusqu'au feedback de fin (1=15ms ; 100=1500ms)</li> <li>- Feedback de fin : audio ou vibrotactile</li> </ul> <p>Final cue A(●) V(●)</p>
	Pipe	 <ul style="list-style-type: none"> <li>- Matériel : 4 faces avec 4 tag vidéo sur un tube en acier (200 mm, 20mm Ø) ; reconnaissance d'image via une webcam</li> <li>- Distance (1-100) = angle de rotation du tube jusqu'au feedback de fin (1=1/400<sup>ième</sup> de tour ; 100=¼ tour)</li> <li>- Feedback de fin : audio ou vibrotactile</li> </ul> <p>Final cue A(●) V(●)</p>
	Pédale	 <ul style="list-style-type: none"> <li>- Matériel : game pedal</li> <li>- Distance (1-100) = l'inclinaison de la pédale par la main jusqu'au feedback de fin (1=0,2° ; 100=22°)</li> <li>- Feedback de fin : audio ou vibrotactile</li> </ul> <p>Final cue A(●) V(●)</p>
	Slider	 <ul style="list-style-type: none"> <li>- Matériel : slider</li> <li>- Distance (1-100) = translation du slider avec la main jusqu'au feedback de fin (1=0,54 et 100=54 mm)</li> <li>- Feedback de fin : audio ou vibrotactile</li> </ul> <p>Final cue A(●) V(●)</p>
	Volant	 <ul style="list-style-type: none"> <li>- Matériel : Game Wheel</li> <li>- Distance (1-100) = rotation du volant avec les bras jusqu'au feedback de fin (1=0,90° ; 100=90°)</li> <li>- Feedback de fin : audio ou vibrotactile</li> <li>- Remarque : le volant peut se tourner vers la gauche ou vers la droite</li> </ul> <p>Final cue A(●) V(●)</p>

**Tableau 4-2 : tableau de synthèse des 15 techniques d'interaction de l'expérience 1.1**

La section suivante présente le protocole de la première expérience de cette première étude.

### 4.2.2 Protocole expérimental (exp. 1.1)

Un total de 12 sujets les yeux bandés ont participé à l'expérience 1.1. Aucun des sujets n'avaient participé à une étude connexe. Le groupe était composé de neuf hommes et trois femmes, avec un âge moyen de  $28,7 \pm 7,2$ . Le dispositif expérimental a été installé dans une pièce silencieuse et comprenait huit prototypes (voir Tableau 4-2) connectés à un PC via un hub USB2 [**Mobility Lab Hub USB 2.0 10 ports**]. L'expérience a été contrôlée par un logiciel *ad-hoc* écrit en langage C#.

Les sujets ont passé l'expérience en deux sessions d'une heure à au moins 24h d'intervalle, l'une avec des signaux auditifs et l'autre avec des signaux tactiles. Cette dichotomie est valable pour les techniques « Passive » où le stimulus est soit sonifié, soit vibrotactile (sauf pour le « PalmGauge » qui n'a pas d'alternative). Elle est également valable pour des techniques « Active » où le feedback est soit vibrotactile (courte vibration) ou sonifié (court son pur).

L'ordre des techniques et les sessions ont été contrebalancés entre les sujets. Au cours d'une session, les sujets sont assis en face d'une table sur laquelle sont posés les dispositifs. Nous leur avons expliqué que le but de la tâche était d'estimer une valeur - entre 1 et 100 - avec huit techniques d'interaction différentes. Chaque session comprenait huit blocs correspondant aux huit techniques d'interaction.

Au début de chaque bloc, le but de la tâche et le fonctionnement de la technique étaient rappelés. Puis, les sujets passaient une **phase de familiarisation** de dix essais avec des valeurs pseudo-aléatoires entre 1 et 100. Pendant cette phase, pour chaque essai, les sujets devaient d'abord estimer la valeur avec le dispositif avant d'avoir la correction. Cette phase était reproductible et permettait aux sujets de se construire une représentation interne de l'échelle des valeurs. **La phase de test** incluait 30 essais avec des valeurs pseudo-aléatoires entre 1 et 100. Après chaque essai, les sujets estimaient verbalement la valeur mais ne recevaient pas de correction ou de commentaires concernant la valeur cible ou l'exactitude de leur réponse.

Après chaque bloc, les sujets répondaient à un questionnaire de satisfaction (SUS, (Bangor, Kortum, and Miller 2008)) concernant la technique du bloc. **A la fin des huit blocs**, les sujets ont utilisé des cartes illustrant les techniques afin de les classer en fonction de leur préférence.

Les sujets ont réalisé les huit blocs selon un ordre différent. L'ordre des blocs était défini par les règles suivantes : 1 / une technique n'a jamais la même position dans la série d'un sujet à l'autre ; 2 / une technique passive (« PalmGauge », « DualStim. » et « SingleStim. ») n'est jamais suivie par une autre technique passive. Enfin, aucun sujet n'a eu le même ordre des huit blocs dans la deuxième session expérimentale.

Nous avons mesuré une variable dépendante, l'erreur relative entre les valeurs estimées et la cible. La moyenne de l'erreur représente la précision et l'écart-type de l'erreur représente sa dispersion. Nous avons également mesuré le score SUS et l'ordre de préférence comme variables qualitatives. Les variables indépendantes sont les facteurs de conception et les techniques.

Cette première expérimentation (1.1) avait pour but de comparer, dans les mêmes conditions, huit techniques non visuelles dans deux grandes familles de modalités (auditive et vibrotactile) permettant de transmettre une information de type valeur entre 1 et 100. Notre cadre d'analyse nous a permis d'expliquer les différences observées.

La section suivante présente les résultats de cette expérience (exp. 1.1).

### 4.2.3 Résultats de l'expérience 1.1

---

Pour rappel, nous avons 12 sujets, 30 essais par sujet par technique par session, 1 score SUS par sujet par technique par session et 1 ordre de préférence par sujet par session.

Lorsque cela n'est pas spécifiquement mentionné, nous avons utilisé un seuil de significativité de 0,05 pour les tests statistiques. Nous avons également appliqué la correction de Bonferroni pour les comparaisons multiples.

Nous présentons les résultats de manière à répondre dans l'ordre à nos hypothèses ci-dessous :

1. « modalité sensorielle » : les techniques de bases (version auditive) sont meilleures que les techniques alternatives (version vibrotactile) en termes d'efficacité et de satisfaction ;
2. Catégorie de techniques « Active/Passives » : les techniques actives présentent de meilleurs résultats que les techniques passives en termes d'efficacité et de satisfaction.

3. Catégorie de techniques « AvecRéférences / SansRéférence » : les techniques avec des références présentent de meilleurs résultats que les techniques sans référence en termes d'efficacité et de satisfaction.

#### 4.2.3.1 Évaluation d'un apprentissage potentiel entre les deux sessions

---

Tous nos sujets ont passé l'expérience 1.1 en deux sessions. Les sessions sont différentes, l'une contient la série de techniques de base (version auditive), l'autre contient la série des techniques alternatives (version vibrotactile) et l'ordre de présentation était contrebalancé entre les sujets. Cependant, la session présentée en première était la session numéro 1 et celle présentée en seconde était la session numéro 2 et ce, indépendamment du fait que ce soit la série de techniques de bases ou celle des alternatives. Ainsi, il est possible que les utilisateurs présentent des résultats différents, entre la première et la seconde session (effet d'apprentissage ?). Nous avons donc réparti les résultats en deux groupes : « session1 » et « session2 ». Les deux échantillons de données ne suivaient pas la loi normale, nous avons donc testé s'il existait une différence significative entre les deux sessions avec un test non-paramétrique : **Wilcoxon signed rank test**. Le résultat n'est pas significatif ni pour l'efficacité ( $V = 21$ ,  $p$ -value = n.s.), ni pour la satisfaction (Score SUS,  $V = 16$ ,  $p$ -value = n.s.).

Ces résultats confirment qu'il n'y a pas d'apprentissage observé entre la première et la deuxième session. La section suivante présente les résultats selon la session des techniques de bases (version auditive) et la session des techniques alternatives (version vibrotactile).

#### 4.2.3.2 Facteur 1 : la modalité

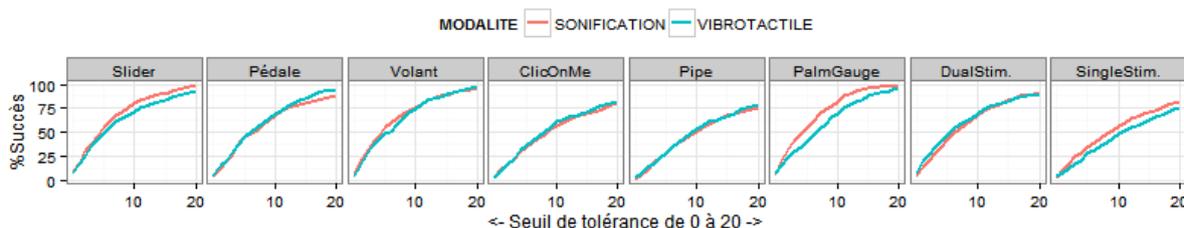
---

Les utilisateurs commençaient soit par la session modalité audio, soit par la session modalité vibrotactile, mais ils ont tous passé les deux sessions. Afin de montrer que les techniques de bases (modalité auditive, sonification) sont meilleures que les techniques alternatives (modalité vibrotactile), nous avons regroupé les résultats en deux groupes : « Sonification » et « Vibrotactile ». Les deux échantillons de données ne suivaient pas la loi normale, nous avons donc testé s'il existait une différence significative entre les deux modalités avec un test non-paramétrique : **Wilcoxon signed rank test**. Le résultat n'est pas significatif ni pour l'efficacité ( $V = 33,5$ ,  $p$ -value = n.s.), ni pour la satisfaction (Score SUS,  $V = 49$ ,  $p$ -value = n.s.).

Nous avons décidé d'illustrer ce résultat à l'aide d'une nouvelle mesure basée sur l'erreur : le pourcentage de succès. Un succès correspond à un essai où la différence entre l'estimation de l'utilisateur et la valeur cible est inférieure ou égale à un seuil de tolérance donné, c'est-à-dire une différence nulle pour une tolérance de 0 et une différence maximale

de 20 pour une tolérance de 20. La Figure 4-2, montre l'évolution des techniques de base (version Sonifiée) et de leur alternative (version Vibrotactile) en fonction d'un seuil de tolérance entre 1 et 20. Nous pouvons observer que les courbes pour chacune des techniques de base se superposent à celle représentant leur technique alternative. Cela explique graphiquement que les techniques sont équivalentes en termes d'efficacité (%succès) et ce, peu importe la tolérance à l'erreur. Il est intéressant de noter que le « PalmGauge », technique sans alternative, à une légère variation d'évolution entre sa session «modalité audio » et sa session «modalité vibrotactile ». La différence, faible, non significative mais observée, peut s'expliquer par le fait que les vibrations, perçues lors du passage des autres techniques dans la session «modalité vibrotactile » impacteraient la sensibilité de la main de l'utilisateur et donc la précision des estimations lors du passage de la technique « PalmGauge ». Une observation similaire peut être faite pour le « Slider » qui requière la main de l'utilisateur. Pour la technique « SingleStim. », la version « modalité audio » semble plus efficace que la version «modalité vibrotactile », alors que la technique « DualStim. » ne présente pas de différence, i.e. les deux versions de « DualStim. » sont aussi efficaces.

Nous pouvons voir également que les techniques « Pipe », « ClicOnMe » et « SingleStim. » ont un pourcentage de succès relativement faible comparé à celui des autres techniques. En effet, ces techniques « SansRéférence » avoisinent les 50% de succès avec une tolérance de 10 alors que les autres obtiennent, pour la même tolérance, plus de 75% de succès.



**Figure 4-2 : les techniques sont équivalentes en termes d'efficacité (%succès) et ce, peu importe la tolérance à l'erreur.**

De notre point de vue, cela montre que les techniques de base et leur alternative sont équivalentes (cf. Les propriétés CARE), même s'il existe de légères différences observables entre les techniques de base et leur alternative. La modalité n'impacte pas significativement les résultats en termes d'efficacité et de satisfaction. Nous pouvons exclure cette variable

indépendante de notre analyse. La section suivante présente les résultats selon les techniques d'interaction.

#### 4.2.3.3 « Facteur 2 » : les techniques

Ni la session, ni la modalité n'impactent significativement les résultats, nous ne faisons plus la distinction entre les techniques de base (version auditive) et leur alternative (version vibrotactile). Nous avons donc regroupé les résultats de chaque paire de techniques (base+alternative) en un seul groupe, soit huit groupes : « PalmGauge », « Slider », « Pédale », « Volant », « ClicOnMe », « Pipe », « DualStim. » et « SingleStim. ».

##### EFFICACITE DES HUIT TECHNIQUES

Les huit échantillons de données ne suivaient pas la loi normale, nous avons donc testé s'il existait une différence significative entre les huit techniques avec un test non-paramétrique : **Friedman rank sum test**. Le résultat est significatif pour l'efficacité ( $\chi^2=58,9$ ,  $df=7$ ,  $p < 0,001$ ). Nous avons également fait le test post-hoc associé : **Pairwise comparisons using Wilcoxon signed rank test**. Le Tableau 4-3 montre les résultats pour l'efficacité pour le post-hoc et la Figure 4-3 montre les moyennes et écarts-types des techniques.

	Slider	Pédale	Volant	ClicOnMe	Pipe	PalmGauge	DualStim.	SingleStim.
Slider	-	0.2721	0.0060	0.0022	0.0022	0.0022	0.0022	0.0022
Pédale	0.2721	-	0.0229	0.0029	0.0022	0.0029	0.0022	0.0022
Volant	0.0060	0.0229	-	0.0060	0.0619	0.1579	0.0414	0.0022
ClicOnMe	0.0022	0.0029	0.0060	-	0.3078	0.0995	0.4802	0.0414
Pipe	0.0022	0.0022	0.0619	0.3078	-	0.5829	0.5303	0.0029
PalmGauge	0.0022	0.0029	0.1579	0.0995	0.5829	-	0.2721	0.0029
DualStim.	0.0022	0.0022	0.0414	0.4802	0.5303	0.2721	-	0.0029
SingleStim.	0.0022	0.0022	0.0022	0.0414	0.0029	0.0029	0.0029	-

Tableau 4-3 : résultats du test post-hoc pour l'efficacité. Notons que nous effectuons nos tests statistiques à partir de la moyenne de l'erreur relative (signée), or, il se trouve que le « Pipe » et le « PalmGauge » ont une moyenne similaire et des écarts-types très différents, pourtant dans ce tableau, ils ne sont pas significativement différents. Ce test ne montre donc pas complètement les différences entre les techniques.

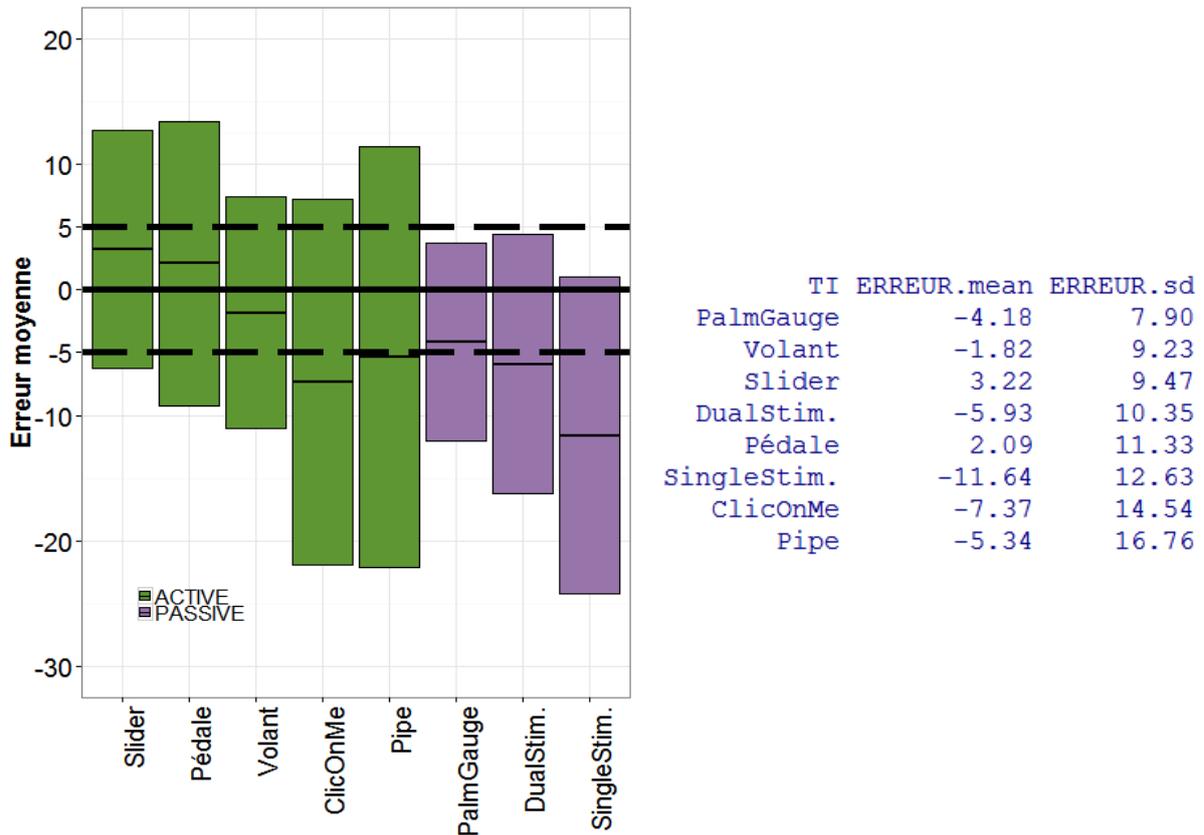


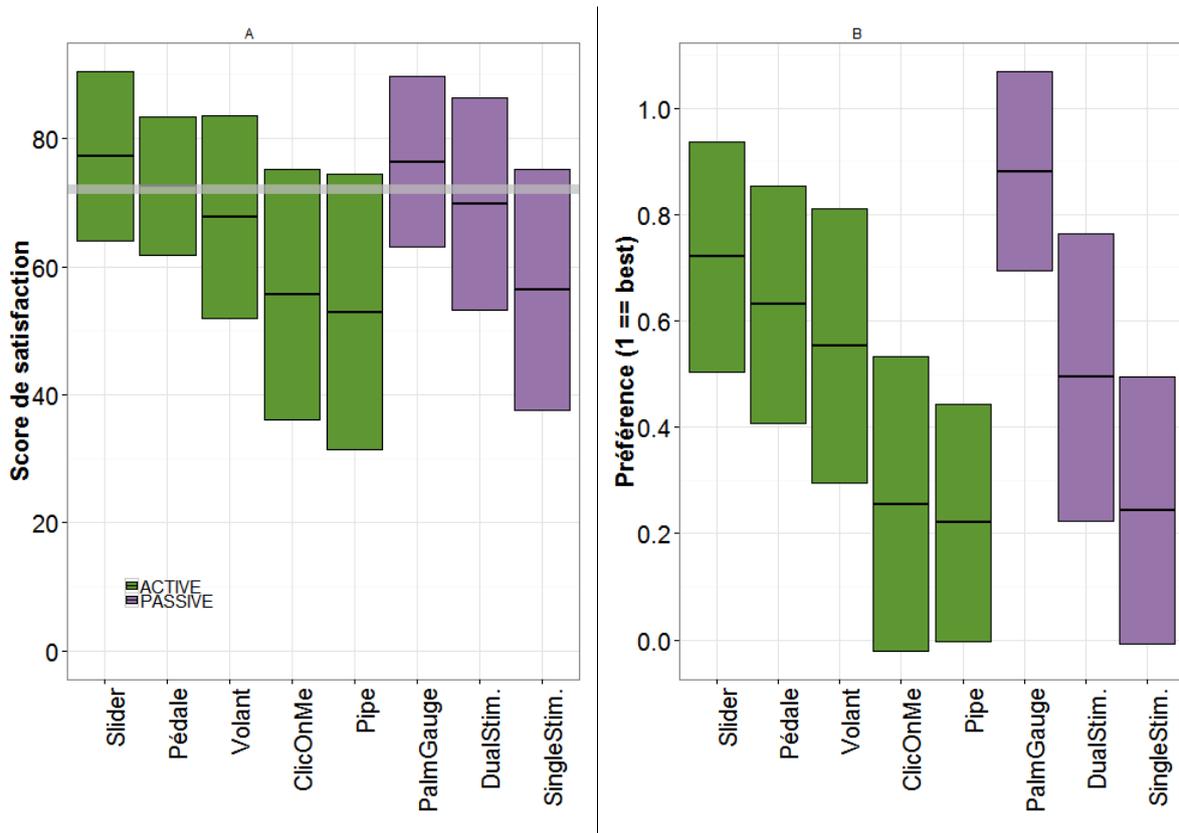
Figure 4-3 : A gauche, l'erreur (précision) et l'écart-type (dispersion) pour chaque technique. Les techniques « ClicOnMe », « Pipe » et « SingleStim. » semblent graphiquement moins précises et plus dispersées que les autres techniques. Les lignes en tiret montrent les limites -5 et +5. A droite, les valeurs des moyennes et des écarts-types, pour chacune des techniques, ordonnés de manière croissante selon les écarts-types.

#### SATISFACTION DES UTILISATEURS VIS-A-VIS DES TECHNIQUES

Les huit échantillons de données ne suivaient pas la loi normale, nous avons donc testé s'il existait une différence significative entre les huit techniques avec un test non-paramétrique : **Friedman rank sum test**. Le résultat est significatif pour la satisfaction (Score SUS,  $X^2=42,4$ ,  $df=7$ ,  $p < 0,001$ ). Nous avons également fait le test post-hoc associé : **Pairwise comparisons using Wilcoxon signed rank test**. Le Tableau 4-4 montre les résultats pour le score SUS pour le post-hoc et la Figure 4-4 montre les moyennes et écarts-types des Scores SUS (A) et des préférences des utilisateurs (B). Ces deux mesures sont cohérentes et montrent que les trois techniques ont été considérées comme étant bonnes ou acceptables (Bangor *et al.* 2008), et ont été aussi les trois préférées des utilisateurs. Voici dans l'ordre de préférence les techniques et leurs score SUS : « PalmGauge » (76,4, bonne), « Slider » (77,2, bonne) et « Pédale » (72,6, OK) puis les techniques restantes considérées comme non acceptables « Volant » (67,8), « DualStim. » (69,8), « ClicOnMe » (55,6), « SingleStim. » (56,5) et le « Pipe » (52,9).

	Slider	Pédale	Volant	ClicOnMe	Pipe	PalmGauge	DualStim.	SingleStim.
Slider	-	0.0827	0.0059	0.0037	0.0022	0.7986	0.0165	0.0058
Pédale	0.0827	-	0.0411	0.0047	0.0047	0.2286	0.5558	0.0093
Volant	0.0059	0.0411	-	0.0084	0.0029	0.0259	0.9283	0.0232
ClicOnMe	0.0037	0.0047	0.0084	-	0.8136	0.0047	0.0339	0.7533
Pipe	0.0022	0.0047	0.0029	0.8136	-	0.0022	0.0042	0.5337
PalmGauge	0.7986	0.2286	0.0259	0.0047	0.0022	-	0.1358	0.0058
DualStim.	0.0165	0.5558	0.9283	0.0339	0.0042	0.1358	-	0.0596
SingleStim.	0.0058	0.0093	0.0232	0.7533	0.5337	0.0058	0.0596	-

**Tableau 4-4 : résultats du test post-hoc pour la satisfaction (score SUS). Notons, par exemple, que le « Pipe » n'est pas différent du « ClicOnMe », ni du « SingleStim. », ces trois techniques sont « SansRéférence ».**



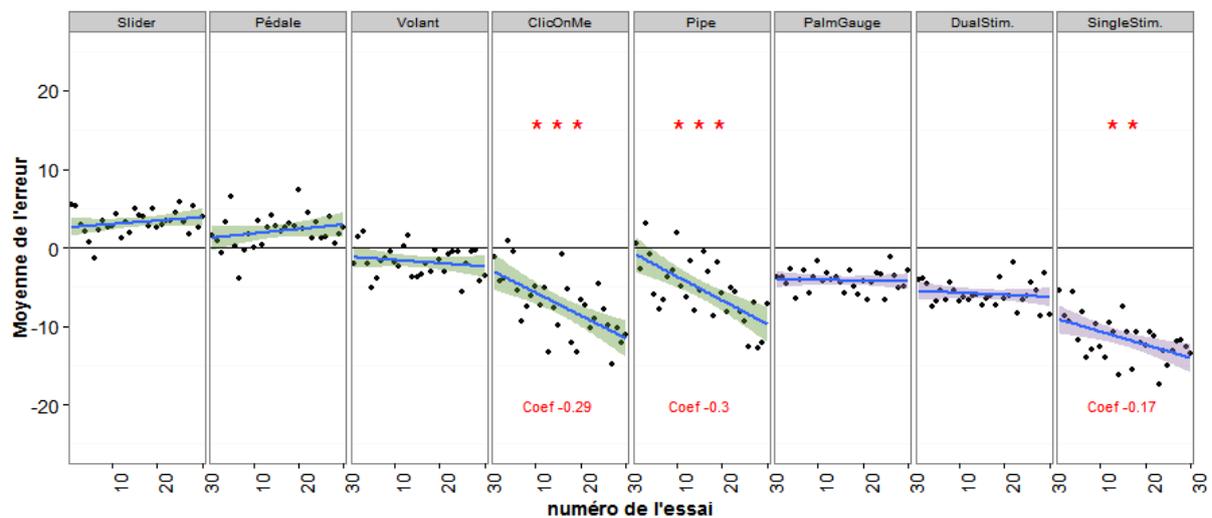
**Figure 4-4 : A gauche (A), Scores de satisfaction pour chaque technique. La ligne grise à 72 sur l'axeY, représente trois catégories SUS : bonne (au-dessus), acceptables (sur), et non acceptables (en dessous). Trois techniques sont considérées comme étant bonnes ou acceptables : « Slider », « Pédale » et « PalmGauge ». A droite (B), l'ordre de préférence –normalisé– des techniques par les utilisateurs, sachant qu'une préférence de 1 est la meilleure et 0 la pire. Notons que les trois meilleures techniques (« PalmGauge », « Slider » et « Pédale ») sont sur le podium de la satisfaction (A) et de la préférence (B), tandis que les pires techniques sont « ClicOnMe », « SingleStim. » et « Pipe » pour (A) et (B). Rappelons que ces trois dernières techniques sont « SansRéférence ».**

Afin de compléter notre analyse, nous avons observé la dérive de l'erreur au cours du « temps ».

#### ÉTUDE DE LA DERIVE DE L'ERREUR

Nous avons fait une régression linéaire de l'erreur moyenne au cours du « temps » (ici, nous parlons des essais qui sont passés les uns après les autres) pour chaque technique (cf. Figure 4-5). La dérive de l'erreur signifie que la précision diminue de manière continue entre le

premier et le dernier essai. Nous avons observé deux groupes de techniques, dans l'un, les techniques dérivent de manière significative, et dans l'autre, les techniques sont « stables » au cours du temps. Dans le groupe « dérive », nous retrouvons les techniques « SingleStim. » (Prédiction ( $b=-0,169$   $t(28)= -3,27$  avec  $p < 0,01$ ). Force du test (adjusted  $R^2= 25\%$   $F(1,28)= 10,679$ ,  $p.value < 0,01$ )), « ClicOnMe » (Prédiction ( $b=-0,291$   $t(28)= -4,382$  avec  $p < 0,001$ ). Force du test (adjusted  $R^2= 38.6\%$   $F(1,28)= 19,21$ ,  $p.value < 0,001$ )) et le « pipe » (Prédiction ( $b=-0,304$   $t(28)= -4,74$  avec  $p < 0,001$ ). Force du test (adjusted  $R^2= 42.5\%$   $F(1,28)= 22,45$ ,  $p.value < 0,001$ )).



**Figure 4-5 : Régression linéaire de la moyenne de l'erreur (12 sujets) par technique. Les points noirs représentent la moyenne de l'erreur de chaque essai pour tous les sujets pour une technique. Notons que les techniques (« ClicOnMe », « Pipe » et « SingleStim. ») qui dérivent significativement n'ont pas de références explicites. Les [\*] montrent les dérives significatives (\*\*  $p < 0,01$  ; \*\*\*  $p < 0,001$ ) et les chiffres en dessous des régressions linéaires montrent leur coefficient.**

#### **CONCLUSION DE L'ANALYSE DES DIFFERENCES ENTRE LES TECHNIQUES**

En résumé, les techniques montrent, entre elles, des différences significatives en termes d'efficacité et de satisfaction. Nous avons vu que les techniques « SansRéférences » étaient les pires techniques et que celles « AvecRéférences » étaient les meilleures. Toutefois, le facteur de conception <Active/Passive> devrait jouer un rôle lui aussi. La section suivante expose les résultats selon deux groupes : techniques actives et techniques passives.

#### **4.2.3.4 Facteur 3 : la catégorie Active/Passive (A/P)**

Nous avons regroupé les résultats en deux groupes, soit cinq techniques pour « Active » et trois techniques pour « Passive ».

#### **EFFICACITE DES GROUPES A/P**

Les deux échantillons de données ne suivaient pas la loi normale, nous avons donc testé s'il existait une différence significative entre les deux groupes avec un test non-paramétrique : **Wilcoxon signed rank test**. Le résultat est significatif pour l'efficacité ( $V = 78$ ,  $p\text{-value} < 0,01$ ). La Figure 4-6 montre les moyennes et écarts-types des groupes « Active » et « Passive ».

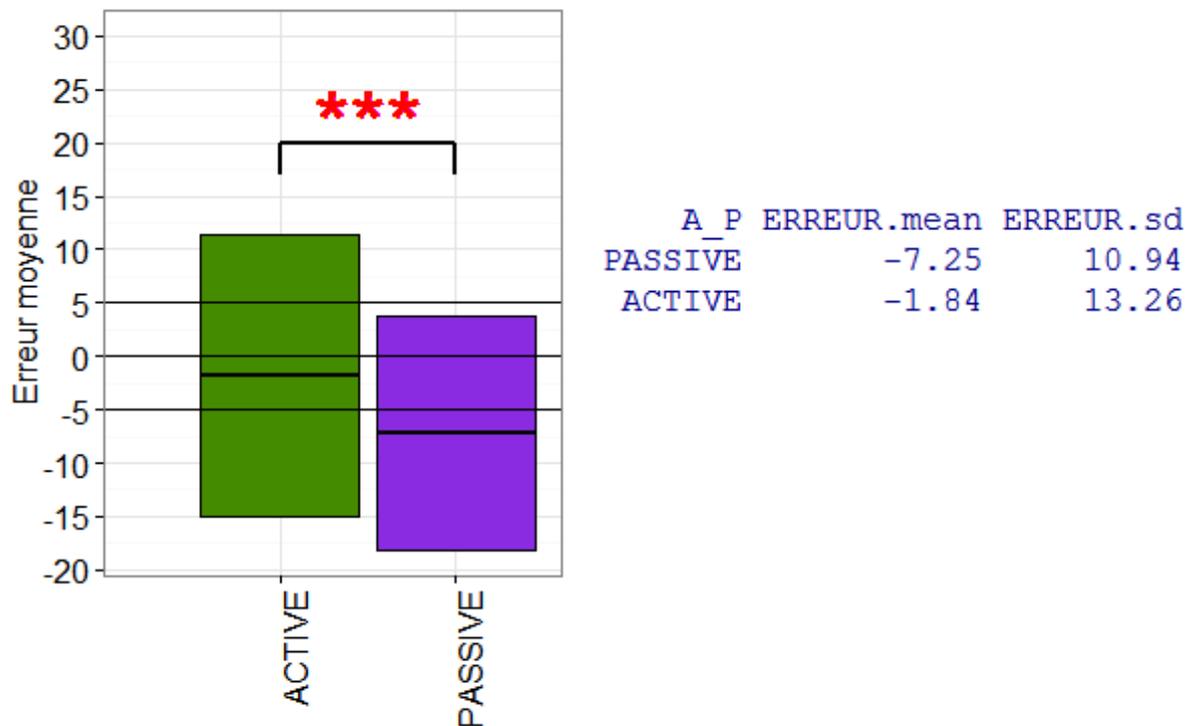


Figure 4-6 : A gauche, l'erreur (précision) et l'écart-type (dispersion) pour chaque groupe de techniques (Active/Passive). Les lignes montrent les limites -5,0 et +5 ; les [\*] montrent les différences significatives (\*\*\*,  $p < 0,001$ ). A droite, les valeurs des moyennes et des écarts-types, pour les groupes « Active » et « Passive », ordonnés de manière croissante selon les écarts-types.

#### SATISFACTION DES UTILISATEURS VIS-A-VIS DES GROUPES A/P

Les deux échantillons de données ne suivaient pas la loi normale, nous avons donc testé s'il existait une différence significative entre les deux groupes avec un test non-paramétrique : **Wilcoxon signed rank test**. Le résultat n'est pas significatif pour la satisfaction (Score SUS,  $V = 25$ ,  $p\text{-value} = \text{n.s.}$ ).

#### ÉTUDE DE LA DERIVE DE L'ERREUR DES GROUPES A/P

Nous avons fait une régression linéaire de l'erreur moyenne au cours du « temps » pour chaque catégorie de techniques (cf. Figure 4-7). Il se trouve que les deux groupes de techniques dérivent de manière significative, « Active » (Prédiction ( $b = -0,107$   $t(28) = -4,078$  avec  $p < 0,001$ )). Force du test (adjusted  $R^2 = 35\%$   $F(1,28) = 16,629$ ,  $p\text{-value} < 0,001$ ) et « Passive » (Prédiction ( $b = -0,068$   $t(28) = -2,995$  avec  $p < 0,01$ )). Force du test (adjusted  $R^2 = 21.6\%$   $F(1,28) = 8,971$ ,  $p\text{-value} < 0,01$ )).

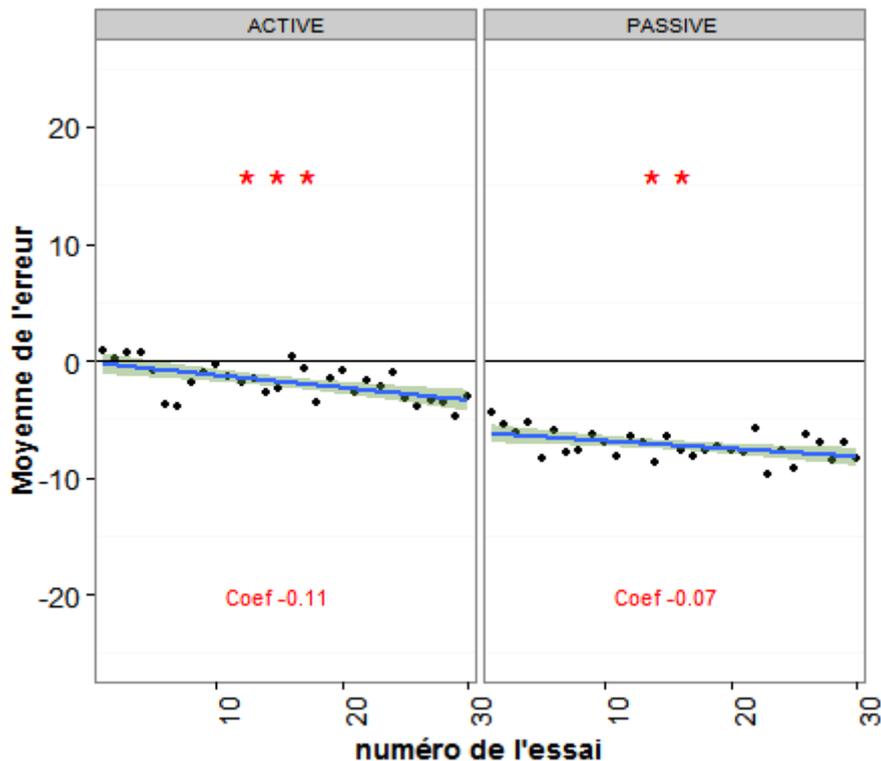


Figure 4-7 : Régression linéaire de la moyenne de l'erreur (12 sujets) par catégorie de technique Active/Passive. Les points noirs représentent la moyenne de l'erreur de chaque essai pour tous les sujets pour un groupe de techniques. Les [\*] montrent la p.value (\*\*  $p < 0,01$  ; \*\*\*,  $p < 0,001$ ) et les chiffres en dessous des régressions montrent leur coefficient.

#### CONCLUSION DE L'IMPACT DU FACTEUR ACTIVE/PASSIVE SUR LES RESULTATS

En résumé, le groupe des techniques « Active » est significativement meilleur que celui des « Passive » en termes d'efficacité moyenne. Cependant, il n'y a pas de différence significative en termes de satisfaction et les deux groupes dérivent au cours du temps. La section suivante expose les résultats selon deux groupes : techniques AvecRéférences et techniques SansRéférence.

#### 4.2.3.5 Facteur 4 : la catégorie AvecRéférences/SansRéférence (AR/SR)

De la même manière que pour l'analyse des groupes « Active/Passive », nous avons réparti les résultats en deux nouveaux groupes : soit cinq techniques pour « AvecRéférences » et trois techniques pour « SansRéférence ».

#### EFFICACITE DES GROUPES AR/SR

Les deux échantillons de données ne suivaient pas la loi normale, nous avons donc testé s'il existait une différence significative entre les deux groupes AvecRéférences/SansRéférence avec un test non-paramétrique : **Wilcoxon signed rank test**. Le résultat est significatif pour l'efficacité ( $V = 78$ ,  $p\text{-value} < 0,01$ ), ce qui est identique aux résultats de ce même test pour la

catégorie Active/Passive. La Figure 4-8 montre les moyennes et écarts-types des catégories de technique AvecRéférences/SansRéférence.

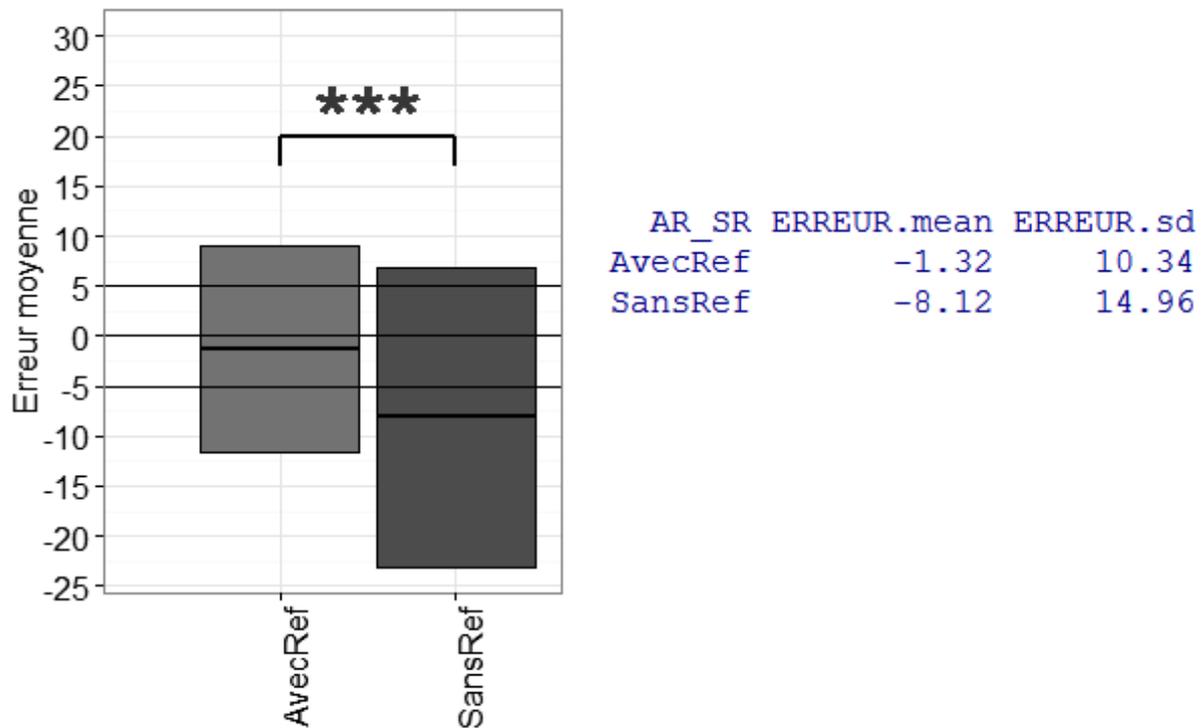


Figure 4-8 : A gauche, l'erreur (précision) et l'écart-type (dispersion) pour chaque catégorie de technique AvecRéférences/SansRéférence. Les lignes montrent les limites -5,0 et +5 ; les [\*] montrent les différences significatives (\*\*\*,  $p < 0,001$ ). A droite, les valeurs des moyennes et des écarts-types, pour les groupes de techniques AvecRéférences et SansRéférence, ordonnés de manière croissante selon les écarts-types.

#### SATISFACTION DES UTILISATEURS VIS-A-VIS DES GROUPES AR/SR

Les deux échantillons de données ne suivaient pas la loi normale, nous avons donc testé s'il existait une différence significative entre les deux groupes AvecRéférences/SansRéférence avec un test non-paramétrique : **Wilcoxon signed rank test**. Le résultat est significatif pour la satisfaction (Score SUS,  $V = 77$ ,  $p\text{-value} < 0,01$ ). Le groupe de techniques « AvecRéférences » à un score SUS moyen de 72,75, tandis que le groupe « SansRéférence » à un score SUS moyen de 55,0.

#### ÉTUDE DE LA DERIVE DE L'ERREUR DES GROUPES AR/SR

Nous avons fait une régression linéaire de l'erreur moyenne au cours du « temps » pour chaque catégorie de techniques (cf. Figure 4-9). Il se trouve que seul le groupe de techniques « SansRéférence » dérive (prédiction ( $b = -0,255$   $t(28) = -6,85$  avec  $p < 0,001$ ). Force du test (adjusted  $R^2 = 61.3\%$   $F(1,28) = 46,927$ ,  $p\text{-value} < 0,001$ ).

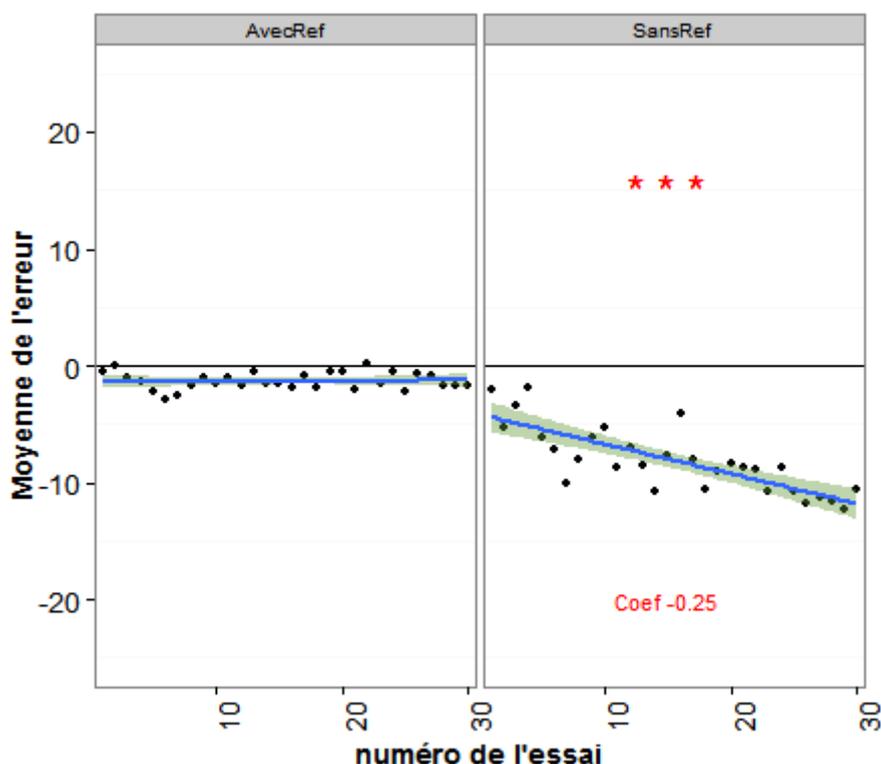


Figure 4-9 : Régression linéaire de la moyenne de l'erreur (12 sujets) par catégorie de technique AvecRéférences/SansRéférence. Les points noirs représentent la moyenne de l'erreur de chaque essai pour tous les sujets pour un groupe de techniques. Les [\*] montrent la p.value (\*\*\*,  $p < 0,001$ ) et les chiffres en dessous des régressions montrent leur coefficient.

#### CONCLUSION DE L'IMPACT DU FACTEUR AVECREFERENCES/SANSREFERENCE SUR LES RESULTATS

En résumé, le groupe des techniques « Avec références » est significativement meilleur que celui des « SansRéférence » en termes d'efficacité et de satisfaction. De plus, le groupe de techniques « SansRéférence » dérive significativement au cours du temps avec un coefficient de 0,25, c'est-à-dire que l'erreur augmente de 1 tous les 4 essais à partir du moment où la phase de familiarisation est finie.

#### 4.2.4 Discussion de l'expérience 1.1

Dans cette expérimentation, nous avons conçu quinze techniques d'interaction non visuelle que nous avons évaluées dans une tâche d'estimation de valeurs entre 1 et 100. Le premier résultat montre qu'il n'y a pas de différences significatives, en termes d'efficacité et de satisfaction, entre les techniques de base (« modalité audio ») et leur alternative « modalité vibrotactile ». Cela signifie que les utilisateurs ne sont pas meilleurs avec l'une des deux modalités et qu'ils n'en préfèrent pas l'une plus que l'autre. Dans notre contexte de transmission d'information(s), des modalités sont **équivalentes** si chacune permet de transmettre la même information (cf. Les propriétés CARE). Or d'après les résultats de cette expérience 1.1 visant à transmettre une information de type valeur, il n'y a pas de

différence significative entre nos modalités d'interaction auditive et nos modalités d'interaction tactile, de plus nous avons comparé ces techniques d'interaction dans les mêmes conditions (sauf ce facteur). Donc nous pouvons dire, que de notre point de vue, les techniques de base sont équivalentes à leur technique alternative. C'est un résultat important car les utilisateurs ont le choix entre de la modalité audio ou tactile pour réaliser une tâche d'estimation de valeur entre 1 et 100 et peuvent ainsi choisir la modalité sensorielle la plus adéquate en fonction du contexte. Par exemple, il est connu que les utilisateurs aveugles utilisent les sons naturels de l'environnement pendant leurs déplacements. Les techniques basées tactiles permettraient de leur fournir, sans parasiter leur canal auditif, des informations discrètes et fiables sur des distances à parcourir, des propriétés d'objets (ex. la taille, position, etc.), sur les itinéraires, etc. Cela est également vrai pour la conception d'interfaces dédiées aux utilisateurs voyants qui ont besoin d'alternative afin de ne pas surcharger (voir décharger) leurs canaux sensoriels visuels et auditifs, comme par exemple, les pilotes, les pompiers, etc.

Mise à part le facteur modalité, le Tableau 4-5 récapitule les différents résultats obtenus dans l'expérience 1.1. Il montre clairement que trois techniques sont certainement utilisables, avec une bonne précision et un taux de satisfaction élevé. Lorsque nous considérons l'ensemble de ces résultats, les trois techniques les plus utilisables sont le « PalmGauge », le « Slider » et la « Pédale ».

technique (ligne)-résultat (colonne)	Précision (Mean)	Dispersion (SD)	Dérive	SUS	Préférence	A/P	AR/SR
PalmGauge	<5	<12	«stable»	Bonne	1	Passive	AR
Slider	<5	<12	«stable»	Bonne	2	Active	AR
Pédale	<5	<12	«stable»	Acceptable	3	Active	AR
Volant	<5	<12	«stable»	Non Accept.	4	Active	AR
DualStim.	>5	<12	«stable»	Non Accept.	5	Passive	AR
SingleStim.	>5	>12	OUI	Non Accept.	7	Passive	SR
ClicOnMe	>5	moins	OUI	Non Accept.	6	Active	SR
Pipe	>5	moins	OUI	Non Accept.	8	Active	SR

**Tableau 4-5 : synthèse des résultats pour les huit techniques que nous avons conçu et évalué.**

**TECHNIQUES ACTIVES OU TECHNIQUES AVEC REFERENCES ?**

Nous avons vu que les techniques « Active » avaient des résultats supérieurs à celles « Passive ». Notre hypothèse que le fait d'être actif dans une tâche d'estimation améliore les résultats s'en trouve confortée et montre que le facteur de conception « Active/Passive » impacte l'utilisabilité des techniques. D'ailleurs, comme mentionné par (Hutchins *et al.* 1985), il est probable que la manipulation directe des objets aide les utilisateurs dans leur perception et donc pour leur estimation. De plus, le feedback immédiat favorise la traduction naturelle des intentions de l'utilisateur en actions et rend la tâche plus facile. Cependant, il semble que la technique la plus utilisable (« PalmGauge ») soit « Passive » ; elle est toutefois suivie par deux techniques « Active » (« Slider » et « Pédale »). Encore plus surprenant, nous avons observé que certaines des techniques « Active » (« ClicOnMe » et « Pipe ») ne sont absolument pas efficaces et non satisfaisantes. Alors même si nous avons remarqué un impact positif du facteur « Active » sur l'efficacité des techniques, il semble que ce facteur ne suffit pas à expliquer l'utilisabilité des techniques.

La présence ou l'absence de références explicites serait un meilleur facteur pour expliquer la dichotomie entre les techniques utilisables et non-utilisables. Pour le « PalmGauge », le signal tactile continu sous la main transmis par la rotation de l'aiguille, semble transmettre efficacement une valeur. En outre, la densité des récepteurs tactiles dans la main est très élevée (Johansson and Vallbo 1979) ce qui fournit probablement à l'utilisateur une bonne sensation de la position de l'aiguille sous la main et donne confiance à l'utilisateur dans son estimation. Cette combinaison de facteurs explique sans doute, l'excellente précision de cette technique (-4,18) et sa faible dispersion (7,90 seulement), ainsi que son classement comme technique préférée des utilisateurs avec un très grand score SUS (76,4). En outre, bien que nous ne vérifiions pas spécifiquement ce point dans l'étude, nous devinons que l'efficacité du « PalmGauge » résulte du fait que les utilisateurs pouvaient associer une position de l'aiguille à un point de contact sous la main, c'est-à-dire que les utilisateurs créaient eux-mêmes leurs références explicites. En effet, les utilisateurs se servaient de leurs doigts comme repères et se servaient de ceux-ci afin d'affiner leur estimation. Cette méthode est potentiellement très puissante car elle permet d'augmenter le nombre de références de la technique.

Plus généralement, les résultats montrent très clairement que les techniques « SansRéférence » ne permettaient pas aux utilisateurs de se re-calibrer, ce qui entraînait des estimations de plus en plus fausses et provoquait une dérive de l'erreur. A l'inverse, les

techniques « AvecRéférences » permettaient aux utilisateurs d'avoir une estimation très stable des valeurs au cours du temps (Figure 4-5 et Figure 4-9). Fait intéressant, la référence était de type différent selon les techniques. Nous avons vu que la technique « PalmGauge » permettait aux utilisateurs de créer eux-mêmes leurs références explicites avec leur main et ce, pendant la phase de familiarisation. Dans la technique « DualStim. », la référence explicite était logicielle et indiquait la valeur médiane (50) avant la valeur variable, ce qui permettait aux utilisateurs de comparer les durées et d'améliorer leur estimation. D'ailleurs, si nous comparons « DualStim. » et « SingleStim. », nous pouvons voir que la présence d'une référence améliore grandement l'efficacité (bonne précision, pas de dérive), ce résultat justifie la recommandation de (Fraisse 1984) pour des tâches d'estimation de durée. Dans les techniques « Slider », « Volant » et « Pédale », les références explicites étaient physiques (tangibles). Par exemple, ces trois techniques avaient toutes une butée (physique) pour marquer la fin de la plage de valeur. Nous avons observé que les utilisateurs se basaient souvent sur la butée de fin, afin d'estimer une valeur intermédiaire. Par exemple : pour le « Slider » les utilisateurs mesuraient avec leurs doigts, la distance entre la position du slider et la fin du rail ; pour le « volant », les utilisateurs une fois la valeur trouvée, tournaient le volant « à fond » pour mesurer l' « angle » manquant pour atteindre la valeur 100, à ce moment, les utilisateurs pouvaient se représenter la position de la valeur cible sur l'échelle des valeurs et affiner leur estimation ; pour la « pédale », les utilisateurs ont adopté une technique mixte entre celle utilisée pour le « Slider » et celle utilisée pour le « Volant ». Par conséquent, là encore ce que nous appelons la référence a été très utile pour estimer la cible par méthode de comparaison et ce, avec une grande stabilité dans le temps (cf. Figure 4-5).

#### ***PRECISION, DISPERSION, DERIVE ET SATISFACTION SONT COHERENTS***

Fait intéressant, les techniques plus précises, les moins dispersées (faible écart-type) et les plus stables ont également reçu les meilleurs scores de satisfaction et de classement (Tableau 4-5). Cette observation n'est pas triviale car les sujets n'avaient aucun retour sur leur propre efficacité, ce qui signifie que les sujets ont perçu, d'eux-mêmes, que leurs estimations étaient bonnes, et cela leur donnait confiance. En ce qui concerne les techniques « Active », basées sur la modalité haptique, il a été montré précédemment l'importance de la proprioception dans des tâches d'estimation de valeur (Altenhoff *et al.* 2012). Nous avons déjà préconisé pourquoi les techniques plus précises étaient aussi les plus fiables et les plus satisfaisantes. Par opposition, le « Pipe », « ClicOnMe » et

« SingleStim. » étaient les pires techniques utilisables. Les techniques « ClicOnMe » et « SingleStim. » étaient similaires en ce que la seule façon d'estimer la valeur cible était d'intégrer le temps écoulé. Il se trouve que, sans référence, l'intégration de temps est très difficile. De plus, nos résultats ont montré que cela ne faisait pas une grande différence lorsque le sujet estimait la durée à partir d'un seul stimulus auditif ou vibratoire, ou à partir d'une pression sur un bouton.

Le « Pipe » a été la pire technique en termes de dispersion (16,8), de stabilité (coefficient de 0,30), de satisfaction (SUS = 52,9), et de classement (8<sup>ième</sup> place). Une explication de ces observations repose sur la conception de la technique elle-même. Le « Pipe » ne présente aucune référence, ni physique, ni logicielle et les utilisateurs n'ont pu en créer eux-mêmes. En effet, le « Pipe » peut tourner indéfiniment. Bien que cette conception offre de meilleures possibilités de contrôle, il n'y a pas de positions de départ ou de fin qui puissent aider l'utilisateur à prendre une référence. De plus, le matériau présentait une surface polie sans aucun relief, seul le mouvement de plusieurs doigts des deux mains pour faire tourner le « Pipe » permettait transmettre l'angle de rotation. Nous faisons l'hypothèse que l'utilisation des toutes ces parties du corps, hypersensibles d'une part et symétriques d'autre part, aie provoqué l'effet escompté inverse, c'est-à-dire que l'utilisateur n'a pas pu intégrer autant d'information.

***VERS UNE VERSION INTEGREE DES TECHNIQUES : « PALMGAUGE », « SLIDER » ET « PEDALE »***

Nous croyons fermement que de telles techniques d'interaction non visuelle peuvent être facilement intégrées dans de nombreux appareils personnels. De toute évidence, cela ne représente pas un défi technique important de concevoir des dispositifs de bureau qui intègrent ces techniques. Un clavier ou une souris peuvent facilement intégrer un artéfact similaire à la technique « PalmGauge » par exemple. Cependant, nous pensons en particulier que les Smartphones représentent une plate-forme prometteuse pour offrir de telles techniques d'interaction non visuelle. Concrètement, le récent « Galaxy Note Edge<sup>40</sup> » offre désormais un écran arrondi pour un accès rapide aux menus spécifiques. Il est clair que la partie arrondie, recouvrant le bord droit de l'appareil, pourrait servir de support pour la mise en œuvre de la technique « Slider ». Enfin, les dispositifs à « forme changeante » sont actuellement étudiés au niveau académique (Hemmert 2010; Kim *et al.* 2008), mais sont aussi en train de devenir des concepts industriels et des produits commerciaux (Kildal,

<sup>40</sup> <http://www.lesnumeriques.com/telephone-portable/galaxy-note-edge-p21576/test.html>

Paasovaara, and Aaltonen 2012). Ceci est bien sûr encore exploratoire, mais la technique de « PalmGauge » que nous avons conçu pourrait facilement être transposée sur ce type de dispositif.

#### 4.2.5 Conclusion de l'expérience 1.1

---

Dans cette étude, nous avons conçu et évalué quinze techniques d'interaction non visuelle qui permettent à l'utilisateur d'estimer une valeur. Ces quinze techniques sont basées sur huit techniques de base, lesquelles sont conçues en fonction des connaissances préalables issues principalement de la recherche psychologique sur la perception de la distance et de la recherche sur les dispositifs auditifs et tactiles permettant de transmettre des informations. Parmi ces huit solutions, trois émergent clairement en termes d'efficacité (précision et la stabilité des estimations) et de satisfaction utilisateur (SUS, classement) : le « PalmGauge », le « Slider » et la « Pédale ». En parallèle de l'évaluation des techniques, nous avons testé l'impact de quatre facteurs de conception : techniques actives ou passives, présence ou non de références, le type de codage ANGLE, LONGUEUR ou DURÉE et la modalité principale auditive ou tactile. Dans cette première expérience, nous n'avons pas traité les résultats selon le type de codage des techniques, nous nous y emploierons dans la conclusion générale de ce chapitre. Notre expérience a montré l'équivalence de techniques de base et de leur alternative, ce qui offre la possibilité aux utilisateurs de choisir selon leur préférence de modalité ou alors, selon le contexte d'utilisation. Nous avons également vérifié que le fait d'être « actif » dans une tâche d'estimation facilitant la tâche, bien que les résultats aient révélé un effet significatif plutôt faible du facteur « Actif/Passif ». Enfin, notre troisième facteur de conception « AvecRéférences/SansRéférence » a eu l'impact le plus important en termes d'efficacité et de satisfaction. Dans de futurs travaux, nous nous concentrerons sur le déploiement *in situ* de ces techniques. En effet, elles pourraient être utilisées à des tâches beaucoup plus exigeantes que l'estimation d'une seule valeur. Par exemple, il peut être intéressant d'étudier l'utilisation de ces techniques dans le cadre d'une mémorisation d'itinéraire, c'est à dire, d'un ensemble de couples direction-distances. En outre, il peut être intéressant de vérifier si elles peuvent alléger la charge perceptive lorsqu'elles sont utilisées au cours de tâches visuelles très exigeantes.

L'expérience 1.1 nous a permis de connaître un peu plus l'impact des facteurs sur les résultats. Toutefois, nous souhaitons connaître la relation entre deux facteurs de conception qui ont significativement impacté les résultats de cette première étude. A la fin de cette

étude, nous montrerons les différences entre ces deux facteurs en termes de puissance d'impact, et proposerons un classement des techniques de nos deux expériences (1.1 et 1.2) en termes de puissance d'impact.

### 4.3 Expérience 1.2 : étude spécifique des facteurs de conception A/P et AR/SR

Dans l'expérience 1.1, nous avons étudié l'impact de trois facteurs de conception sur l'efficacité de quinze techniques d'interaction dans une tâche d'estimation d'une seule valeur. Nous avons vu que le facteur « AvecRéférences/SansRéférence » avait un impact significatif sur les techniques en termes d'efficacité et de satisfaction. Ce facteur semble également plus « puissant » que celui « Actif/Passif » et que celui de la modalité « Audio/Tactile ». Cependant, notre expérience 1.1 est une étude exploratoire et le protocole n'équilibrait pas les groupes de techniques (5 actives/ 3 passives ; 5 (autres) AvecRéférences / 3 SansRéférence). De plus, le nombre de références était disparate selon les techniques, l'encodage n'était pas équivalent (ex. 22° == 100 valeurs vs. 180° == 100 valeurs). Afin de conforter les résultats observés dans l'expérience 1.1, nous proposons une exp. 1.2 complémentaire pour approfondir l'impact de ces deux facteurs de conception empruntés aux sciences cognitives. De plus, dans l'expérience 1.1, nous avons fait l'hypothèse que la présence de références, avant ou après le stimulus, améliorerait l'estimation, nous avons vérifié cette hypothèse, afin d'aller un peu plus loin, nous proposons une version où les références seront « dans » le stimulus.

#### 4.3.1 Mise en place de l'expérimentation

Nous avons pris la meilleure des techniques « Active » (A) et la meilleure des techniques « Passive » (P) en excluant le « PalmGauge ». Puis, pour chacune de ces techniques, nous avons conçu deux alternatives : l'une avec références (AR), l'autre sans référence (SR) :

1. meilleure TI Active : Slider → 2 nouvelles conceptions : AR et SR
2. meilleure TI Passive : dual stimulus → 2 nouvelles conceptions : AR et SR

Ce qui nous donne quatre techniques : « Règle graduée », « Règle lisse », « Bip Multiple » et « Bip Solo », selon deux facteurs à deux niveaux : « Actif/Passif » et « AvecRéférences / SansRéférence ». Cela nous donne quatre conditions : « A+AR » ; « A+SR » ; « P+AR » ; « P+SR ».

Dans la littérature nous avons vu qu'être actif améliore les performances (Gibson 1962; Hutchins *et al.* 1985), nous savons que la présence de références y contribue également (Fraisie 1984). Cependant, ces facteurs n'ont pas été testés dans un contexte de transmission d'information, plus précisément dans une tâche d'estimation d'une valeur entre 1 et 100. D'ailleurs, dans l'expérience précédente, nous avons vu que la meilleure technique était « Passive ». Cela nous amène à nous demander lequel des deux facteurs est le plus « impactant » en termes d'efficacité. D'une manière générale, nous souhaitons ordonner ces facteurs selon leur « puissance d'impact » et répondre à la question : est-ce le fait d'être une technique « Active » ou le fait d'être une technique « AvecRéférences » est un gage d'efficacité ? Quel facteur est le plus « impactant » en termes d'efficacité ? Afin de répondre ces questions, nous posons les deux inégalités suivantes :

$$\text{Active (A) ; Passive (P) ; AvecRéférences (AR) ; SansRéférence (SR)}$$

$$\text{« A+AR > A+SR > P+AR > P+SR » ou « A+AR > P+AR > A+SR > P+SR »}$$

(nous pensons que c'est celle de droite)

## 4.3.2 Présentation des quatre nouvelles techniques

### 4.3.2.1 La Règle-Lisse (A+SR)

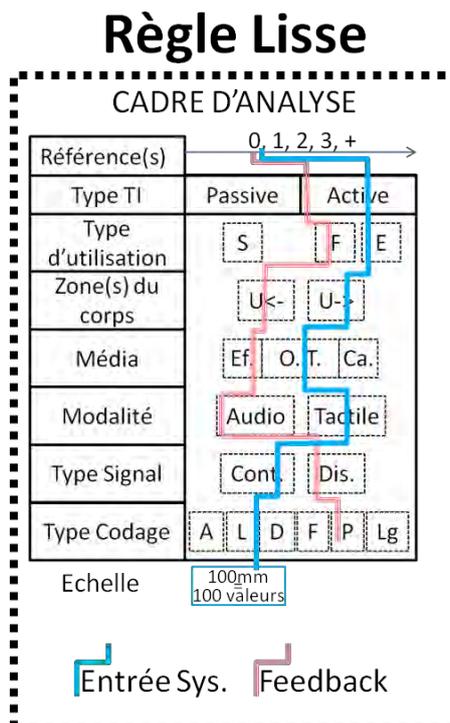
**Principe** : Tracer un trait avec son doigt sur un écran tactile à l'aide d'une règle lisse (objet tangible). Il faut effectuer une double tape sur la gauche de l'écran pour commencer à tracer un trait vers la droite. Le déplacement du doigt crée un écart (longueur). La longueur correspond à la valeur (ex. une distance). Lorsque la valeur est atteinte (ou dépassée), un feedback (retour d'information) audio averti l'utilisateur. La plage de valeur est comprise entre 1 et 100. De plus, la règle n'a aucune référence (voir Figure 4-10).

**Illustration :**



**Figure 4-10 : illustration de la technique « Règle lisse ». Exemple d'utilisation : « je fais une double tape, j'entends un bip, je déplace doucement mon doigt vers la droite. Lors de mon déplacement, j'entends un bip, j'estime la valeur transmise».**

**Cadre d'analyse :** technique d'interaction active sans référence. Le signal, en entrée du système, s'effectue avec un objet tangible (écran tactile + règle sans référence) capable de transmettre un stimulus tactile de manière continue, et dont la longueur code une valeur entre 1 et 100 (100 mm = 100 valeurs). L'utilisateur recevra un feedback de l'accomplissement de son action, avec un effecteur capable de transmettre un signal auditif discret basé sur une pulsation (ici, un bip) (voir Figure 4-11).



**Figure 4-11 : illustration de l'analyse de la technique Règle-Lisse.**

**Matériel** : Tablette tactile avec Windows Seven 64bits, un casque audio [Plantronics audio 650 usb] et une règle lisse (voir Figure 4-10).

#### 4.3.2.2 La Règle-Graduée (A+AR)

**Principe** : Tracer un trait avec son doigt sur un écran tactile à l'aide d'une règle graduée (objet tangible). Il faut effectuer une double tape sur la gauche de l'écran pour commencer à tracer un trait vers la droite. Le déplacement du doigt crée un écart (longueur). La longueur correspond à une valeur (ex. une distance). Lorsque la valeur est atteinte (ou dépassée), un feedback audio averti l'utilisateur. La plage de valeur est comprise entre 1 et 100. De plus, la règle a six références (clou dans la planche) : 0, 20, 40, 60, 80 et 100 (voir Figure 4-12).

**Illustration** :



Figure 4-12 : illustration de la technique « Règle graduée ». Exemple d'utilisation : « je fais une double tape, j'entends un bip, je déplace doucement mon doigt vers la droite en passant sur les clous (réf. physiques). Lors de mon déplacement, j'entends un bip, j'estime la valeur transmise, m'aidant des références sur la règle».

**Cadre d'analyse** : technique d'interaction active avec six références. Le signal, en entrée du système, s'effectue avec un objet tangible (écran tactile + règle avec références) capable de transmettre un stimulus tactile de manière continue, et dont la longueur code une valeur entre 1 et 100 (100 mm = 100 valeurs). L'utilisateur recevra un feedback de l'accomplissement de son action, avec un effecteur capable de transmettre un signal auditif discret basé sur une pulsation (ici, un bip) (voir Figure 4-13).

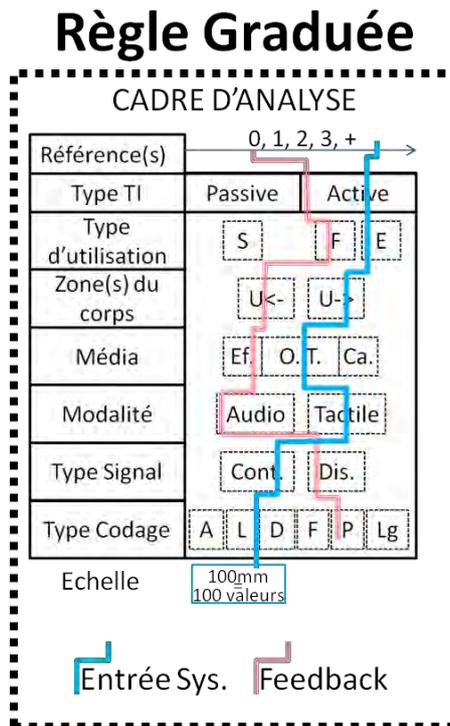


Figure 4-13 : illustration de l'analyse de la technique Règle-Graduée.

**Matériel :** Tablette tactile avec Windows Seven 64bits, un casque audio [Plantronics audio 650 usb] et une règle graduée de 100 millimètres où six clous ont été plantés, chacun espacé de 20 millimètres pour marquer les vingtaines : 0, 20, 40, 60, 80 et 100 (voir Figure 4-12).

### 4.3.2.3 Le BipSolo (P+SR)

**Principe :** (identique à SingleStimulus de l'exp. 1.1) la durée d'un bip -sonore- représente une valeur entre 1 et 100 (voir Figure 4-14).

**Illustration :**



Figure 4-14 : Illustration de la technique BipSolo. Exemple d'utilisation : « J'entends un son pendant une certaine durée, puis plus rien. J'estime la valeur transmise »

**Cadre d'analyse** : technique d'interaction passive sans référence. Le signal, en sortie du système, vers l'utilisateur s'effectue avec un dispositif (effecteur) capable de transmettre un stimulus audio de manière continue, et dont la durée code une valeur entre 1 et 100 (1500 ms = 100 valeurs) (voir Figure 4-15).

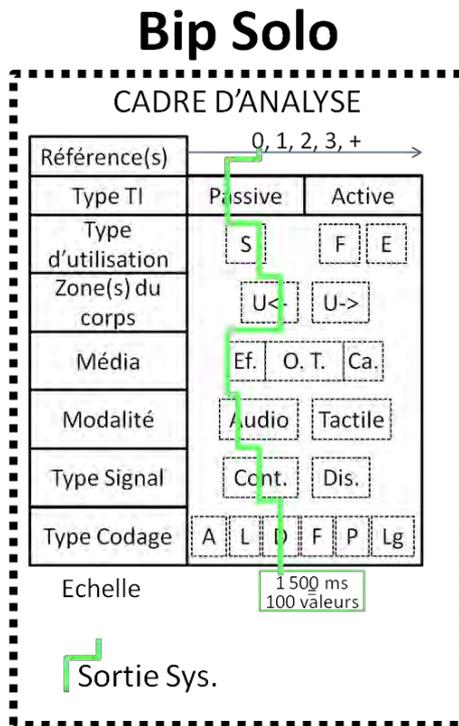


Figure 4-15 : illustration de l'analyse de la technique BipSolo.

**Matériel** : un casque audio [Plantronics audio 650 usb] (voir Figure 4-14). La formule pour traduire la durée en valeur est [durée = 15\*Valeur].

#### 4.3.2.4 Le BipMultiple (P+AR)

**Principe** : la durée d'un son représente une valeur entre 1 et 100. Cependant, le son comporte six repères temporels (bips) : le début du son et les vingtaines. Il y a seul son, dans lequel, sont joués plusieurs bips. Les bips d'une durée de 105 ms correspondent aux vingtaines : 20, 40, 60, 80, 100. Le son, lui-même, a une durée variable entre 15 ms et 2025 ms (voir Figure 4-16).

**Illustration** :



Figure 4-16 : illustration de la technique « BipMultiple ». Exemple d'utilisation : « j'entends un son, dans lequel il y a plusieurs bips, chaque bip marque une vingtaine, j'estime la valeur transmise (dans l'exemple, 100, car 5 bips) ».

**Cadre d'analyse** : technique d'interaction passive avec 6 références (repères temporels : début du son et les bips marqueur de vingtaine). Le signal, en sortie du système, vers l'utilisateur s'effectue avec un dispositif (effecteur) capable de transmettre des stimuli audio de manière continue, et dont la durée code une valeur entre 1 et 100 (1950 ms = 100 valeurs) (voir Figure 4-17).

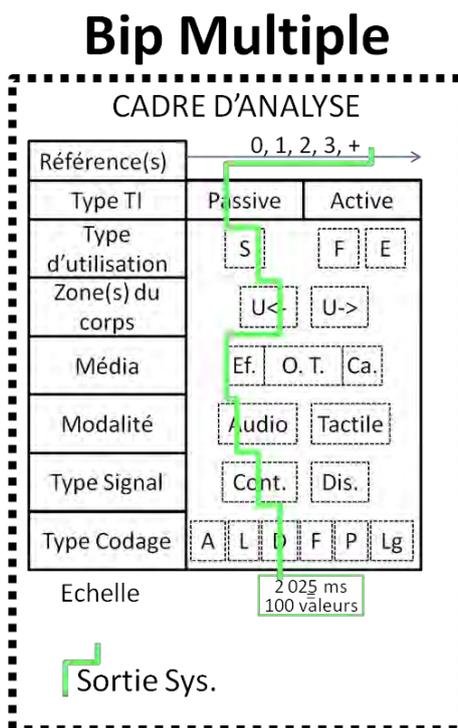


Figure 4-17 : illustration de l'analyse de la technique BipMultiple.

**Matériel** : un casque audio [Plantronics audio 650 usb] (voir Figure 4-16). Un son continu de 285 ms code 19 valeurs [15 ms x 19], puis une variation de ce son à partir de 285ms, pendant 105 ms marque la 20<sup>aine</sup>. Ainsi la valeur 100 est codée par 5 répétitions [15 ms x 19 + 105], soit 1950 ms. La valeur 50 est codée par 2 répétitions [15 ms x 19 + 105] plus un bout de [15 ms x 10], soit 930 ms. Le son est à 440 Hz, les bips sont à 2 000 Hz.

### 4.3.3 Protocole expérimental (exp. 1.2)

---

Un total de 12 sujets les yeux bandés ont participé à l'expérience 1.2. Le dispositif expérimental a été installé dans une pièce silencieuse. L'expérience a été contrôlée par un logiciel ad-hoc écrit en langage C# sur une tablette tactile avec un système d'exploitation Windows Seven 64bits.

Les sujets ont passé l'expérience en une session d'une heure et demie. L'ordre des techniques a été contrebalancé entre les sujets. Au cours d'une session, les sujets sont assis en face d'une table où ont été posés les dispositifs et portent un casque audio. Nous leur avons expliqué que le but de la tâche était d'estimer une valeur, entre 1 et 100, avec quatre techniques d'interaction différentes. Chaque session comprenait quatre blocs correspondant aux quatre techniques d'interaction.

Au début de chaque bloc, le but de la tâche et le fonctionnement de la technique étaient rappelés par synthèse vocale. Puis, les sujets passaient une phase de familiarisation de dix essais avec des valeurs pseudo-aléatoires entre 1 et 100. Pendant cette phase, pour chaque essai, les sujets devaient d'abord estimer la valeur avec le dispositif, avant d'avoir la correction. Cette phase était reproductible et permettait aux sujets de se construire une représentation interne de l'échelle des valeurs. La phase de test incluait 50 essais avec des valeurs pseudo-aléatoires entre 1 et 100. Après chaque essai, les sujets estimaient verbalement la valeur mais ne recevaient pas de correction ou de commentaire concernant la valeur cible ou l'exactitude de leur réponse.

Après chaque bloc, les sujets donnaient leurs commentaires. A la fin des quatre blocs, les sujets répondaient à un questionnaire de satisfaction (SUS, (Bangor *et al.* 2008)) pour chacune des techniques, et ont classé les quatre techniques en fonction de leur préférence.

Les mesures observées sont : l'erreur (estimée-cible), le score SUS et un classement. Nous avons deux facteurs : « Active/Passive » et « AvecRéférences/SansRéférence ». Le but de cette expérience complémentaire était d'étudier la puissance d'impact des facteurs en termes d'efficacité et de satisfaction.

### 4.3.4 Résultats de l'expérience 1.2

---

#### 4.3.4.1 Facteur 1 : la catégorie Active/Passive (A/P)

---

Nous avons regroupé les résultats en deux groupes, soit deux techniques pour « Active » et deux techniques pour « Passive ».

**EFFICACITE DES GROUPES A/P**

Les deux échantillons de données ne suivaient pas la loi normale, nous avons donc testé s'il existait une différence significative entre les deux groupes avec un test non-paramétrique : **Wilcoxon signed rank test**. Le résultat est significatif pour l'efficacité ( $V = 75$ ,  $p$ -value  $< 0,01$ ). La Figure 4-18 montre les moyennes et écarts-types des groupes « Active » et « Passive ».

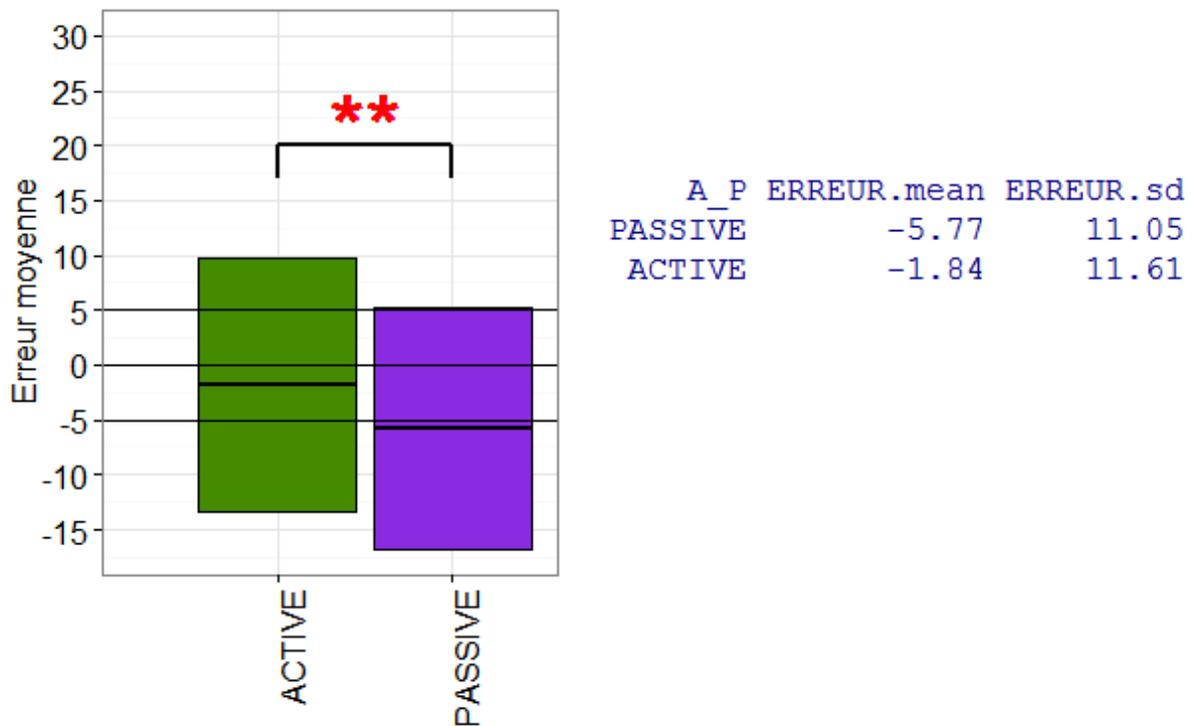


Figure 4-18 : A gauche, l'erreur (précision) et l'écart-type (dispersion) pour chaque groupe de techniques (Active/Passive). Les lignes montrent les limites -5,0 et +5 ; les [\*] montrent les différences significatives (\*\*,  $p < 0,01$ ). A droite, les valeurs des moyennes et des écarts-types, pour les groupes « Active » et « Passive », ordonnées de manière croissante selon les écarts-types.

**SATISFACTION DES UTILISATEURS VIS-A-VIS DES GROUPES A/P**

Les deux échantillons de données ne suivaient pas la loi normale, nous avons donc testé s'il existait une différence significative entre les deux groupes avec un test non-paramétrique : **Wilcoxon signed rank test**. Le résultat n'est pas significatif pour la satisfaction (Score SUS,  $V = 58$ ,  $p$ -value = n.s.).

**ÉTUDE DE LA DERIVE DE L'ERREUR DES GROUPES A/P**

Nous avons fait une régression linéaire de l'erreur moyenne au cours du « temps » pour chaque catégorie de techniques (cf. Figure 4-19). Il se trouve que les deux groupes de techniques dérivent de manière significative, « Active » Prédiction ( $b = -0,055$   $t(48) = -2,89$  avec  $p < 0,01$ ). Force du test (adjusted  $R^2 = 13\%$   $F(1,48) = 8,349$ ,  $p$ .value  $< 0,01$ ) et « Passive »

(Prédiction ( $b=-0,108$   $t(48)= -6,253$  avec  $p < 0,001$ ). Force du test (adjusted  $R^2= 43.7\%$   $F(1,48)= 39,1$ ,  $p.value < 0,001$ )).

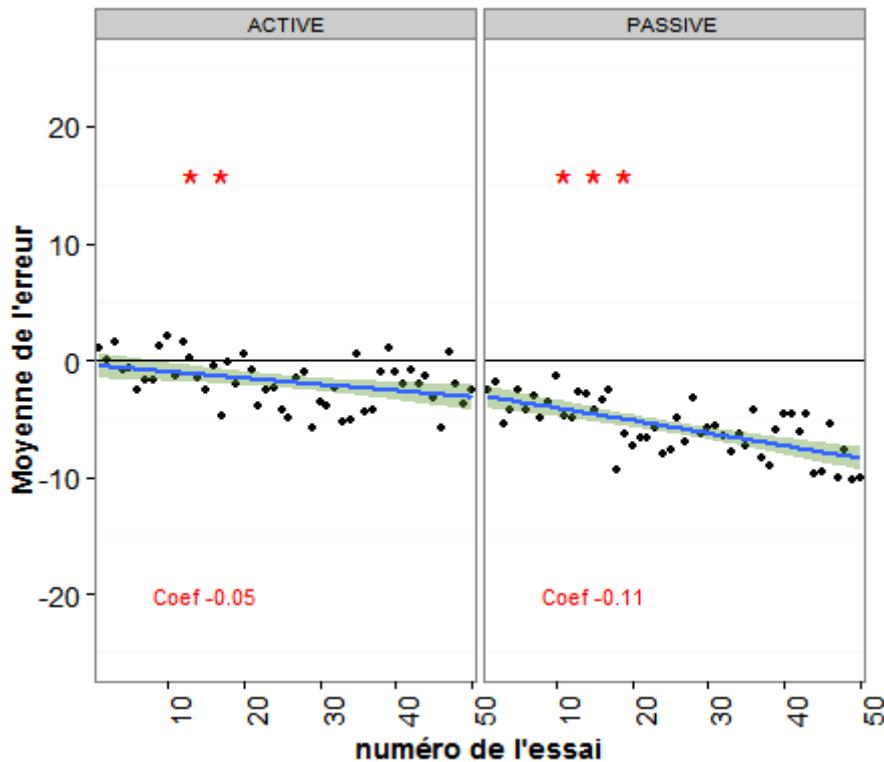


Figure 4-19 : Régression linéaire de la moyenne de l'erreur (12 sujets) par catégorie de technique Active/Passive. Les points noirs représentent la moyenne de l'erreur de chaque essai pour tous les sujets pour un groupe de techniques. Les [\*] montrent la p.value (\*\*  $p < 0,01$  ; \*\*\*,  $p < 0,001$ ) et les chiffres en dessous des régressions montrent leur coefficient.

#### CONCLUSION DE L'IMPACT DU FACTEUR ACTIVE/PASSIVE SUR LES RESULTATS

En résumé, le groupe des techniques « Active » est significativement meilleur que celui des « Passive » en termes d'efficacité moyenne, bien que la dispersion des deux groupes soient élevée ( $>11$ ). Cependant, il n'y a pas de différence significative en termes de satisfaction et les deux groupes dérivent au cours du temps.

#### 4.3.4.2 Facteur 2 : la catégorie AvecRéférences/SansRéférence (AR/SR)

De la même manière que pour l'analyse des groupes « Active/Passive », nous avons réparti les résultats en deux nouveaux groupes : soit deux techniques pour « AvecRéférences » et deux techniques pour « SansRéférence ».

#### EFFICACITE DES GROUPES AR/SR

Les deux échantillons de données ne suivaient pas la loi normale, nous avons donc testé s'il existait une différence significative entre les deux groupes AvecRéférences/SansRéférence

avec un test non-paramétrique : **Wilcoxon signed rank test**. Le résultat est significatif pour l'efficacité ( $V = 75$ ,  $p\text{-value} < 0,01$ ), ce qui est identique aux résultats de ce même test pour la catégorie Active/Passive. La Figure 4-20 montre les moyennes et écarts-types des catégories de technique AvecRéférences/SansRéférence.

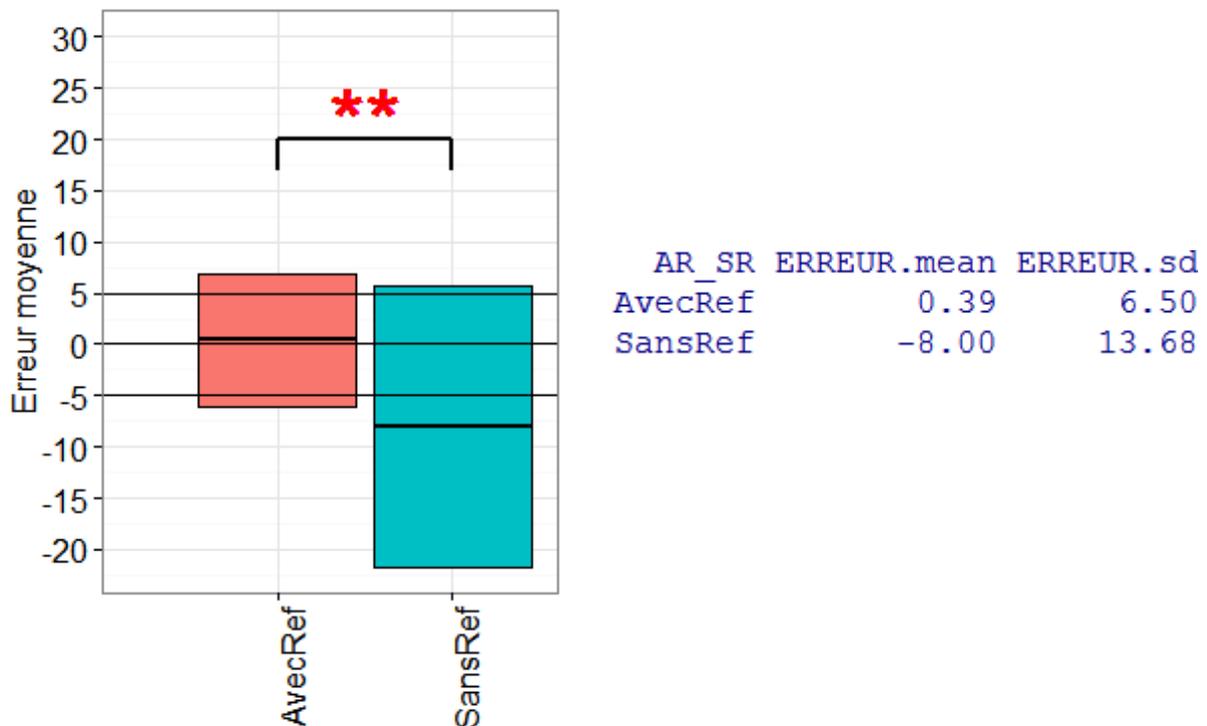


Figure 4-20 : A gauche, l'erreur (précision) et l'écart-type (dispersion) pour chaque catégorie de technique AvecRéférences/SansRéférence. Les lignes montrent les limites -5,0 et +5 ; les [\*] montrent les différences significatives (\*\*,  $p < 0,01$ ). A droite, les valeurs des moyennes et des écarts-types, pour les groupes de techniques AvecRéférences et SansRéférence, ordonnés de manière croissante selon les écarts-types.

#### SATISFACTION DES UTILISATEURS VIS-A-VIS DES GROUPES AR/SR

Les deux échantillons de données ne suivaient pas la loi normale, nous avons donc testé s'il existait une différence significative entre les deux groupes AvecRéférences/SansRéférence avec un test non-paramétrique : **Wilcoxon signed rank test**. Le résultat n'est pas significatif pour la satisfaction (Score SUS,  $V = 59$ ,  $p\text{-value} = \text{n.s.}$ ).

#### ÉTUDE DE LA DERIVE DE L'ERREUR DES GROUPES AR/SR

Nous avons fait une régression linéaire de l'erreur moyenne au cours du « temps » pour chaque catégorie de techniques (cf. Figure 4-21). Il se trouve que, seul, le groupe de techniques « SansRéférence » dérive (Prédiction ( $b = -0,147$   $t(48) = -6,644$  avec  $p < 0,001$ ). Force du test (adjusted  $R^2 = 46.8\%$   $F(1,48) = 44,144$ ,  $p\text{-value} < 0,001$ ).

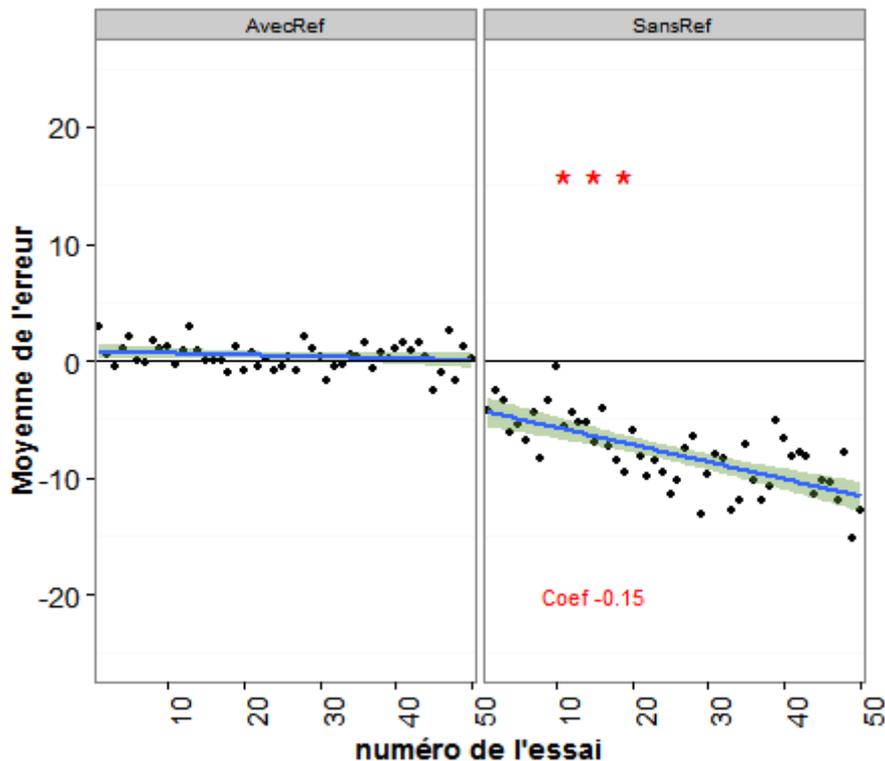


Figure 4-21 : Régression linéaire de la moyenne de l'erreur (12 sujets) par catégorie de technique AvecRéférences/SansRéférence. Les points noirs représentent la moyenne de l'erreur de chaque essai pour tous les sujets pour un groupe de techniques. Les [\*] montrent la p.value (\*\*\*,  $p < 0,001$ ) et les chiffres en dessous des régressions montrent leur coefficient.

#### CONCLUSION DE L'IMPACT DU FACTEUR AVECREFERENCES/SANSREFERENCE SUR LES RESULTATS

En résumé, le groupe des techniques « Avec références » est significativement meilleur que celui des « SansRéférence » en termes d'efficacité, avec une précision moyenne 8 fois plus grande et une dispersion deux fois moins importante. De plus, le groupe de techniques « SansRéférence » dérive significativement au cours du temps avec un coefficient de 0,15, c'est-à-dire que l'erreur augmente de 1 tous les 6 essais à partir du moment où la phase de familiarisation est terminée.

#### 4.3.4.3 Interaction des facteurs A/P et AR/SR

Pour rappel, nous avons quatre groupes d'une technique : « A+AR », « A+SR », « P+AR » et « P+SR ».

#### EFFICACITE DES « TECHNIQUES » (INTERACTION)

Les quatre échantillons de données ne suivaient pas la loi normale, nous avons donc testé s'il existait une différence significative entre les huit techniques avec un test non-paramétrique : **Friedman rank sum test**. Le résultat est significatif pour l'efficacité ( $\chi^2=23,9$ ,  $df=3$ ,  $p < 0,001$ ). Nous avons également fait le test post-hoc associé : **Pairwise comparisons**

using Wilcoxon signed rank test. Le Tableau 4-6 montre les résultats pour l'efficacité pour le post-hoc et la Figure 4-22 montre les moyennes et écarts-types des techniques.

	ACTIVE_AvecRef	ACTIVE_SansRef	PASSIVE_AvecRef	PASSIVE_SansRef
ACTIVE_AvecRef	-	0.015	0.015	0.020
ACTIVE_SansRef	0.015	-	0.124	1.000
PASSIVE_AvecRef	0.015	0.124	-	0.052
PASSIVE_SansRef	0.020	1.000	0.052	-

Tableau 4-6 : résultats du test post-hoc pour l'efficacité avec la correction de Bonferroni. Seul le groupe « A+AR » est significativement différent des autres.

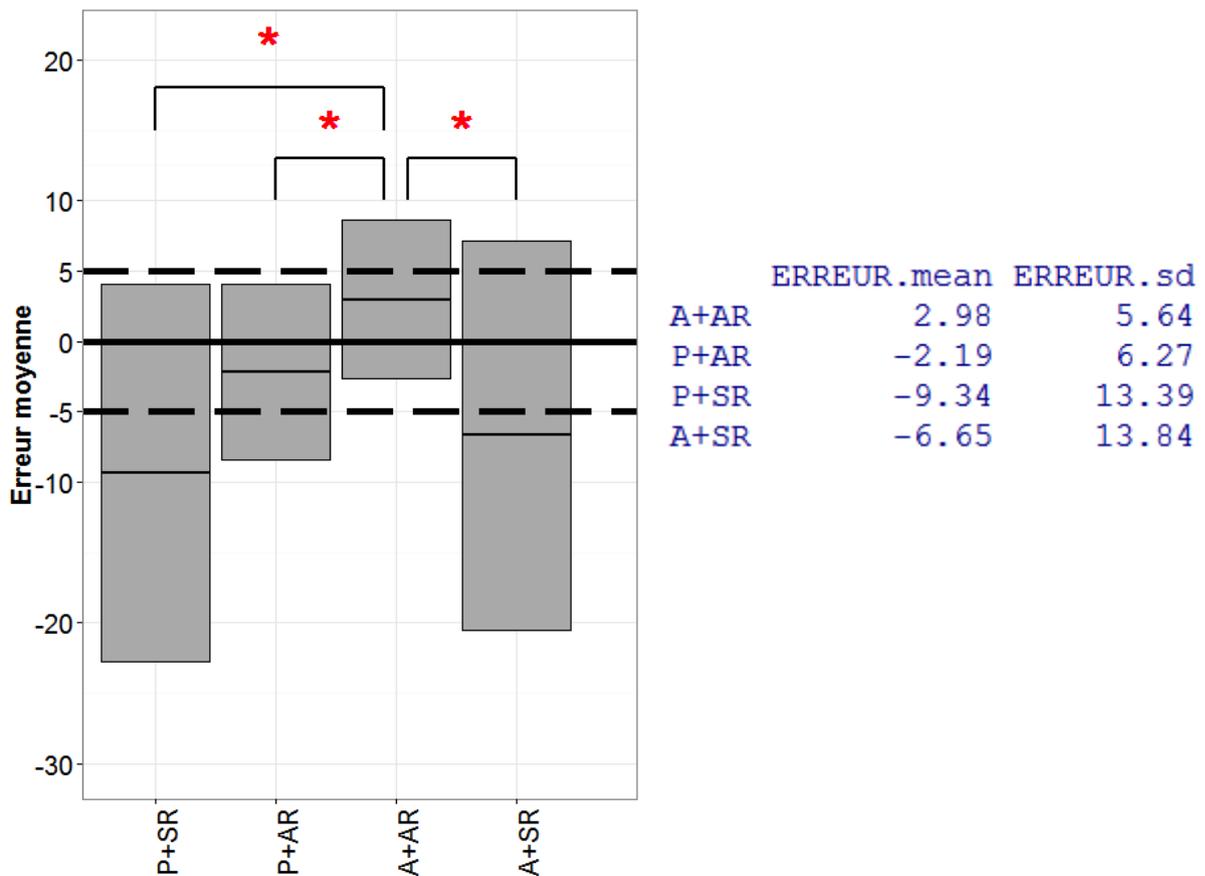


Figure 4-22 : A gauche, l'erreur (précision) et l'écart-type (dispersion) pour chaque technique. Les techniques « +SR » semblent graphiquement moins précises et plus dispersées que les techniques « +AR ». Les lignes en pointillées montrent les limites -5 et +5. A droite, les valeurs des moyennes et des écarts-types, pour chacune des conditions, ordonnés de manière croissante selon les écarts-types.

#### SATISFACTION DES UTILISATEURS VIS-A-VIS DES TECHNIQUES

Les huit échantillons de données ne suivaient pas la loi normale, nous avons donc testé s'il existait une différence significative entre les huit techniques avec un test non-paramétrique : **Friedman rank sum test**. Le résultat est significatif pour la satisfaction (Score SUS,  $X^2=10,4$ ,  $df=3$ ,  $p < 0,05$ ) et la préférence utilisateur (Classement,  $X^2=14,5$ ,  $df=3$ ,  $p < 0,01$ ). Nous avons également fait les tests post-hocs associés : **Pairwise comparisons using Wilcoxon signed rank test**. Le Tableau 4-7 montre les résultats pour le score SUS pour le post-hoc (et Tableau

4-8 pour la préférence utilisateur) et la Figure 4-23 montre les moyennes et écarts-types des Scores SUS (A) et des préférences des utilisateurs (B). Ces deux mesures sont cohérentes et montrent que la techniques qui a été considérée comme étant bonne (Bangor *et al.* 2008), a été aussi la préférée des utilisateurs. Voici dans l'ordre de préférence les techniques et leurs score SUS : « A+AR » (81, bonne), « P+AR » (70,4, OK) puis les techniques restantes considérées comme non acceptables « A+SR » (65,2) et « P+SR » (60,0).

	ACTIVE_AvecRef	ACTIVE_SansRef	PASSIVE_AvecRef	PASSIVE_SansRef
ACTIVE_AvecRef	-	0.35	1.00	0.08
ACTIVE_SansRef	0.35	-	1.00	1.00
PASSIVE_AvecRef	1.00	1.00	-	1.00
PASSIVE_SansRef	0.08	1.00	1.00	-

Tableau 4-7 : résultats du test post-hoc pour la satisfaction (score SUS).

	ACTIVE_AvecRef	ACTIVE_SansRef	PASSIVE_AvecRef	PASSIVE_SansRef
ACTIVE_AvecRef	-	0.157	0.281	0.033
ACTIVE_SansRef	0.157	-	1.000	1.000
PASSIVE_AvecRef	0.281	1.000	-	0.565
PASSIVE_SansRef	0.033	1.000	0.565	-

Tableau 4-8 : résultats du test post-hoc pour la préférence utilisateur (classement). Notons que « A+AR » est significativement différent de « P+SR »

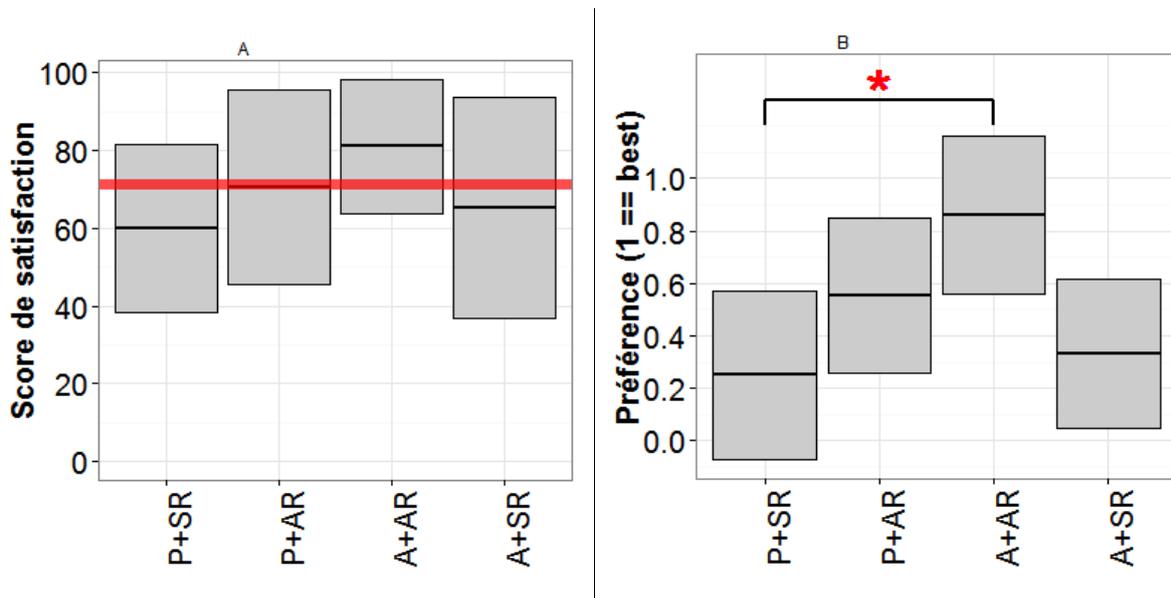
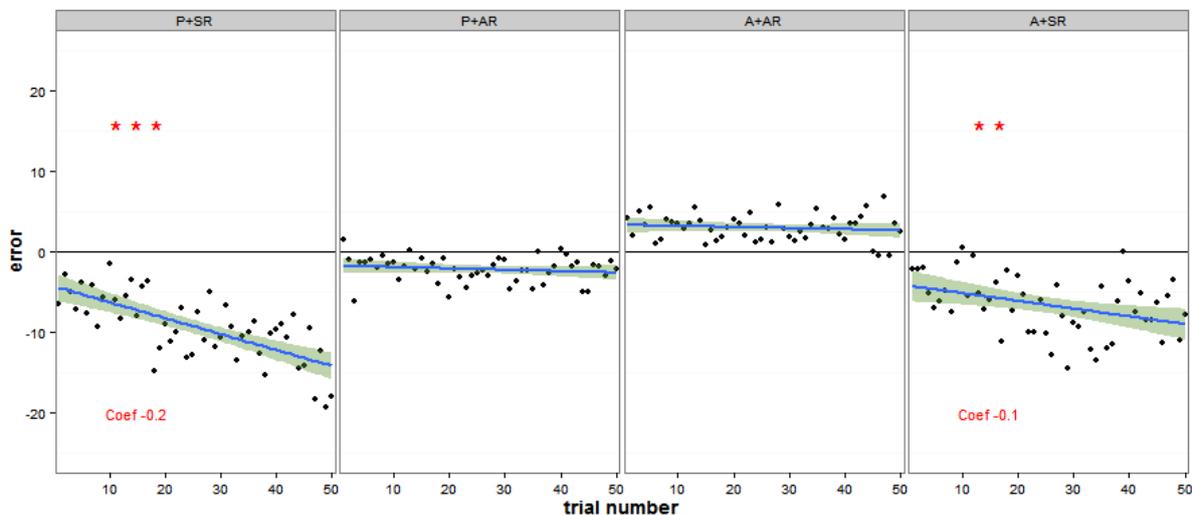


Figure 4-23 : A gauche (A), Scores de satisfaction pour chaque technique. La ligne rouge à 70 sur l'axe Y, représente trois catégories SUS : bonne (au-dessus), acceptables (sur), et non acceptables (en dessous). Une seule technique est considérée comme étant bonne : « A+AR ». A droite (B), l'ordre de préférence – normalisé – des techniques par les utilisateurs, sachant qu'une préférence de 1 est la meilleure et 0 la pire. Notons que la meilleure technique est « A+AR » est sur le podium de la satisfaction (A) et de la préférence (B), tandis que la pire technique est « P+SR » pour (A) et (B).

Afin de compléter notre analyse, nous avons observé la dérive de l'erreur au cours du « temps ».

### ÉTUDE DE LA DERIVE DE L'ERREUR DES TECHNIQUES

Nous avons fait une régression linéaire de l'erreur moyenne au cours du « temps » pour chaque technique (cf. Figure 4-24). La dérive de l'erreur signifie que la précision diminue de manière continue entre le premier et le dernier essai. Nous avons observé deux groupes de techniques, dans l'un les techniques dérivent de manière significative et dans l'autre les techniques sont « stables » au cours du temps. Dans le groupe « dérive », nous retrouvons les techniques SansRéférence « P+SR » (Prédiction ( $b=-0,198$   $t(48)= -6,81$  avec  $p<0,001$ ). Force du test (adjusted  $R^2= 48,1\%$   $F(1,48)= 46,374$ ,  $p.value <0,001$ ) et « A+SR » (Prédiction ( $b=-0,096$   $t(48)= -2,872$  avec  $p<0,01$ ). Force du test (adjusted  $R^2= 12,9\%$   $F(1,48)= 8,247$ ,  $p.value <0,01$ )).



**Figure 4-24 : Régression linéaire de la moyenne de l'erreur (12 sujets) par technique. Les points noirs représentent la moyenne de l'erreur de chaque essai pour tous les sujets pour une technique. Notons que les techniques sans référence dérivent significativement. Les [\*] montrent les dérives significatives (\*\*  $p<0,01$  ; \*\*\*,  $p<0,001$ ) et les chiffres en dessous des régressions linéaires montrent leur coefficient.**

### ÉTUDE DE L'ERREUR EN FONCTION D'UN SEUIL DE TOLERANCE DES TECHNIQUES

La dichotomie d'efficacité entre les deux facteurs semble évidente au vu des résultats précédents. Toutefois, nous avons décidé d'illustrer ce résultat à l'aide d'une mesure basée sur l'erreur : le pourcentage de succès (déjà utilisée dans exp. 1.1). Un succès correspond à un essai où la différence entre l'estimation de l'utilisateur et la valeur cible est inférieure ou égale à un seuil de tolérance donné, c'est-à-dire une différence nulle pour une tolérance de 0 et une différence maximale de 20 pour une tolérance de 20. La Figure 4-25, montre l'évolution des techniques (conditions) en fonction d'un seuil de tolérance entre 1 et 20. Nous pouvons observer que les courbes des techniques « AvecRéférences » sont très proches l'une de l'autre et se distinguent clairement des courbes des techniques « SansRéférence ». Cela montre graphiquement que le facteur « AvecRéférences /

SansRéférence » a plus de « puissance d'impact » sur les résultats que le facteur « Active/Passive » en termes d'efficacité (%succès) et ce, peu importe la tolérance à l'erreur. Il est intéressant de noter que la combinaison de facteur « A+AR » semble être meilleure que « P+AR » et que « A+SR » semble meilleure que « P+SR ». La différence, faible, non significative mais observée, peut s'expliquer par le fait qu'être « actif » améliore un peu les résultats. Ainsi, nous pouvons observer graphiquement :

$$A+AR > P+AR > A+SR > P+SR.$$

Nous pouvons voir également que les techniques « SansRéférence » nécessitent une tolérance presque trois fois plus importante afin d'obtenir le même pourcentage de succès que pour les techniques « AvecRéférences ». Par exemple, 70% de succès nécessite une tolérance de 5-6 pour le groupe « AvecRéférences » et une tolérance de 15-16 pour le groupe « SansRéférence ».

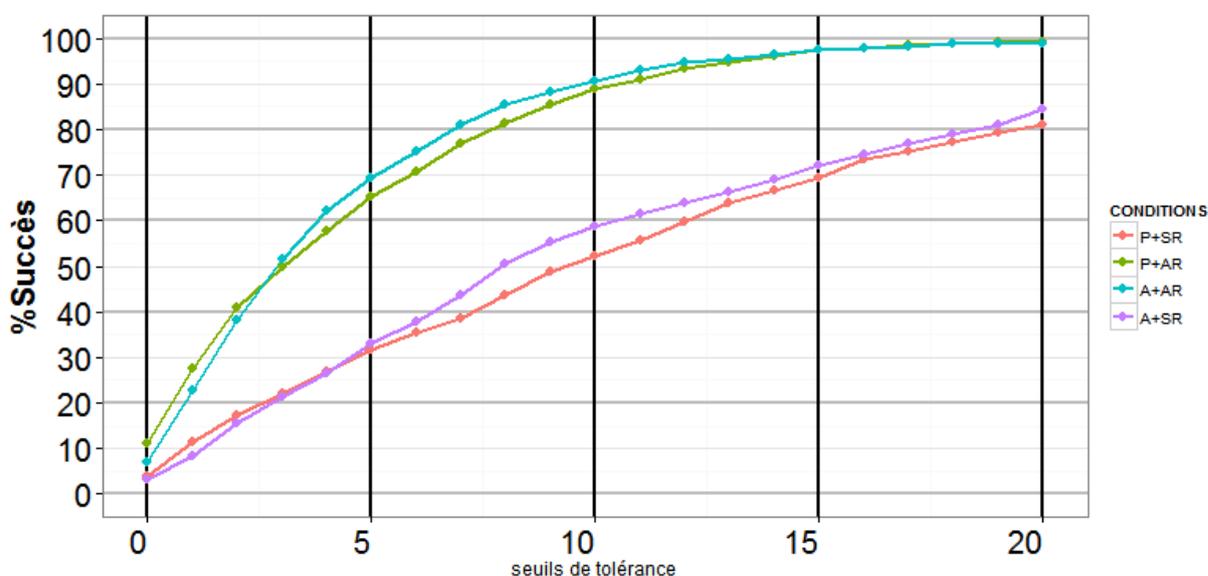


Figure 4-25 : les techniques « SansRéférence » nécessitent une tolérance trois fois plus grande pour obtenir un pourcentage de succès similaire à ceux de techniques « AvecRéférences ».

#### CONCLUSION DE L'IMPACT DE L'INTERACTION DES FACTEURS A/P ET AR/SR

En résumé, les conditions montrent, entre elles, des différences significatives en termes d'efficacité et de satisfaction. Nous avons vu que les techniques « SansRéférence » étaient les pires techniques (faible efficacité, grand écart type, dérive de l'erreur, faible satisfaction des utilisateurs, faible préférence des utilisateurs) et que celles « AvecRéférences » étaient les meilleures. Toutefois, le facteur de conception <Active/Passive> joue un rôle secondaire.

En effet, nous avons pu constater que l'inégalité :  $A+AR > P+AR > A+SR > P+SR$  était totalement respectée en termes de satisfaction utilisateur (score SUS moyen), de préférence (classement moyen) et de tolérance à l'erreur (pourcentage de succès). Et qu'elle était respectée en partie, en termes de précision (moyenne) et de dispersion de l'erreur (écart-type).

#### 4.3.5 Discussion de l'expérience 1.2

Dans cette expérimentation complémentaire, nous avons conçu quatre techniques d'interaction non visuelle à partir des meilleures techniques (« Active » et « Passive ») de l'exp. 1.1. Ces techniques ont été évaluées dans une tâche d'estimation de valeur entre 1 et 100. Les résultats montrent que les deux facteurs de conception « Active/Passive » et « AvecRéférences/SansRéférence » ont une puissance d'impact significative sur l'efficacité et sur la satisfaction des utilisateurs (seulement avec « AvecRéférences/SansRéférence »). Cela est cohérent avec la littérature, nos hypothèses et les résultats de l'exp. 1.1. Cependant, nous souhaitons ordonner ces deux facteurs et ainsi que leurs combinaisons. Ainsi, nous avons analysé les résultats une nouvelle fois en considérant l'interaction des deux facteurs, sachant que nous faisons l'hypothèse qu'ils soient ordonnés suivant l'inégalité suivante :  $A+AR > P+AR > A+SR > P+SR$ . Les résultats montrent que cette hypothèse est totalement validée en termes de satisfaction utilisateur (score SUS moyen), de préférence (classement moyen) et de tolérance à l'erreur (pourcentage de succès), mais partiellement validée, en termes de précision (moyenne) et de dispersion de l'erreur (écart-type).

#### **ET GIBSON 62 ?**

Dans l'expérience de (Gibson 1962), les sujets avaient la possibilité de manipuler leur moule à cookie librement (« active touch »). Dans notre expérience, et plus particulièrement pour les techniques Actives, les sujets pouvaient librement manipuler le dispositif et chercher les références explicites. Cependant, le feedback (retour d'information) de fin, symbolisant la fin de la tâche mais aussi la « position » de la valeur sur le dispositif, n'était joué qu'une seule fois. L'une des différences entre l'expérience de Gibson et la notre, du point de vue de l'exploration active, est que dans celle de Gibson, les sujets manipulaient activement le moule et percevaient la « valeur/forme » jusqu'à ce qu'ils donnent leur réponse, tandis que dans la notre, les sujets recevaient la valeur une seule fois pendant leur manipulation, et jusqu'à ce qu'ils donnent leur réponse. Il est probable que si nous avions laissé la possibilité

aux sujets de percevoir la valeur autant de fois qu'ils le désiraient avant de donner leur réponse, les résultats des techniques Actives seraient bien meilleurs, surtout pour les techniques « A+SR » de l'expérience 1.1 et 1.2. Cependant, cela aurait sans doute créé un écart encore plus important entre les techniques Passives et les techniques Actives, et aurait rallongé le temps de réponse des sujets, ce qui aurait potentiellement biaisé les résultats, car les techniques Passives seraient alors utilisées bien moins longtemps que les Actives.

#### **TECHNIQUES DE BASES VERSUS OPTIMISEES : AUDIO**

Dans ce paragraphe nous comparons les techniques de l'expérience 1.1 à leurs versions « 2.0 » de l'expérience 1.2, pour les techniques basées sur la modalité audio. Les techniques « SingleStimulus » et « BipSolo » sont identiques, leurs résultats sont similaires ( $-11,6 \pm 12,6$  ;  $-9,3 \pm 13,4$ ) ; la légère différence de précision peut être due au nombre d'essais (30 vs 50) et/ou au fait que la moitié des sujets ont passé la technique « BipMultiple » avant « BipSolo », ce qui a pu contribuer à entraîner les sujets à estimer la durée d'un son. Les techniques « DualStimuli » et « BipMultiples » sont différentes sur plusieurs points, deux sons dans l'une, un son avec des repères temporels dans l'autre. Toutefois, leur efficacité semble proche ( $-5,9 \pm 10,5$  ;  $-2,2 \pm 6,3$ ), le nombre de références cinq fois plus important dans « BipMultiple » a permis d'améliorer de près de 40% la précision et de réduire de 60% la dispersion. Le Tableau 4-9 compare les quatre techniques en se basant sur les résultats de « SingleStimulus » :

Techniques	Précision	Dispersion	Dérive
« SingleStimulus »	-11.6	12.6	0.17X-3.27
« BipSolo » + d'entraînement	Amélioration de 20%	<u>Aggravation</u> de 6%	0.19X-6.81
« DualStimulus » + 1 ref	Amélioration de 50%	Réduction de 16%	« stable »
« BipMultiple » + 5 ref + d'entraînement	Amélioration de 81%	Réduction de 50%	« stable »

**Tableau 4-9 : comparatif de quatre techniques en se basant sur les résultats de « SingleStimulus ». Les mots soulignés indiquent un effet négatif.**

Ce tableau, suggère que l'ajout de références dans un signal audio améliore grandement la précision et réduit la dispersion. Nous pouvons également constater que l'entraînement semble améliorer la précision, mais pas la dérive. Cependant, seule la présence de référence(s) « stabilise » l'estimation d'une durée d'un stimulus audio, au cours du temps, et ce, que la référence soit jouée avant le signal ou pendant le signal.

**TECHNIQUES DE BASES VERSUS OPTIMISEES : TACTILE**

Dans ce paragraphe nous comparons la technique de l'expérience 1.1 « Slider » à ses versions « 2.0 » de l'expérience 1.2. Rappelons que le « Slider » a deux références et sa longueur totale pour coder 100 valeurs est de 54 mm, la « Règle-Lisse » n'a aucune référence et sa longueur est de 100 mm et la « règle-graduée » a six références et sa longueur est de 100 mm. Le fait que le nombre d'essais soit différent (30 vs 50) et/ou le fait que la moitié des sujets aie passé la technique « Règle-Lisse » avant « Règle-graduée » a pu contribuer à entrainer les sujets à estimer la longueur d'un trait virtuel. Le Tableau 4-10 comparatif pour les trois techniques en se basant sur les résultats de « Slider » :

Techniques	Précision	Dispersion
« Slider », 54 mm, 2 ref	3.22	9.47
« Règle-Lisse », sans ref, 100 mm + d'entraînement	<u>Aggravation</u> de 106%	<u>Aggravation</u> de 46%
« Règle-Graduée » 2 ref + 4 ref, 100 mm + d'entraînement	Amélioration de 7.5%	Réduction de 40%

**Tableau 4-10 : comparatif de trois techniques en se basant sur les résultats de « Slider.». Les mots soulignés indiquent un effet négatif.**

Ce tableau, suggère qu'un nombre important de références, dans des techniques basées sur la modalité tactile n'impacterait pas la précision, mais seulement la dispersion. Toutefois, la présence de références semble nécessaire afin de garantir une bonne précision.

#### 4.3.6 Conclusion de l'expérience 1.2

Dans cette étude complémentaire, nous avons conçu et évalué quatre techniques d'interaction qui représentaient quatre conditions d'une interaction de deux facteurs « Active/Passive » et « AvecRéférences/SansRéférence ». Ces techniques sont basées sur les meilleures techniques de l'expérience 1.1, à savoir « DualStimuli » pour les passives et « Slider » pour les actives, pour lesquelles nous avons conçu une alternative « SansRéférence » et une « AvecRéférences ». Ces quatre techniques représentent quatre conditions, pour lesquelles nous avons fait l'hypothèse suivante : **A+AR > P+AR > A+SR > P+SR**. Les résultats de cette étude complémentaire montrent que cette hypothèse est totalement validée en termes de satisfaction utilisateur (score SUS moyen), de préférence (classement moyen) et de tolérance à l'erreur (pourcentage de succès). Cependant, les résultats de ladite expérience 1.2, montrent que cette même hypothèse est partiellement validée, en termes de précision (moyenne) et de dispersion de l'erreur (écart-type). Ainsi, cette étude permet de valider les hypothèses de la littérature (« Active > Passive » et « AvecRéférences > SansRéférence »), nos hypothèses de l'exp. 1.1 (« ça marche aussi pour

une tâche d'estimation de valeur entre 1 et 100 »). Nous avons vu également que le facteur de conception « AvecRéférences » est très impactant sur les résultats en termes d'efficacité et de satisfaction, que le facteur de conception « Active » impacte les résultats en termes d'efficacité et enfin, que la combinaison des deux, « Active AvecRéférences », a un impact global bénéfique sur les résultats en termes d'efficacité et de satisfaction. Au vu des résultats, il paraît clair que la combinaison de facteurs « Passive SansRéférence » n'est absolument pas recommandée afin de concevoir des techniques d'interaction non visuelle optimisées pour de la transmission d'information(s).

#### 4.4 Conclusion de la première étude

Ce chapitre a présenté la première étude de cette thèse. La première contribution est la comparaison de près d'une vingtaine de techniques dans une même tâche d'estimation de valeur entre 1 et 100. Tout d'abord, les résultats de l'exp. 1.1 montrent que les techniques Passives exclusivement basées sur la modalité audio (« BipContinu » et « DualBip ») sont équivalentes, en termes d'efficacité et de satisfaction, à leur alternative vibrotactile (« VibreContinu » et « DualVibre ») (cf. Figure 4-1). Il en est de même pour les techniques Actives, dont le feedback (retour d'information) est basé sur la modalité audio, qui sont équivalentes à leur alternative dont le feedback est basé sur la modalité vibrotactile (« Slider », « Pédale », « Volant », « ClicOnMe » et « Pipe »). Puis les résultats des exp. 1.1 et 1.2, montrent que les techniques « Active » sont plus efficaces que les techniques « Passive », bien que, la différence reste faible. Toutefois, les résultats montrent que les techniques « AvecRéférences » sont meilleures que les techniques « SansRéférence » en termes d'efficacité et de satisfaction. De plus, les techniques « SansRéférence » présentent une dérive de l'erreur significative au cours du temps alors que les techniques « AvecRéférences » sont « stables ». Dans l'exp. 1.2, les résultats permettent d'ordonner les quatre combinaisons des deux facteurs de conceptions, « Active/Passive » et « AvecRéférences/SansRéférence », selon leur « puissance d'impact » (cf. Tableau 4-11). La puissance d'impact d'une technique résulte d'un calcul, dans lequel nous faisons la moyenne des rangs normalisés des classements ordonnés pour cinq mesures : précision, dispersion, score SUS, préférence, coefficient de dérive. Ainsi, les résultats de l'exp. 1.2 montrent que l'ordre suivant **A+AR** > **P+AR** > **A+SR** > **P+SR** est également validé en termes de « puissance d'impact ». De plus, nous continuons d'observer la faible différence entre « Active/Passive » et la grande différence entre « AvecRéférences/SansRéférence ».

Combinaisons	rang selon classement par Précision, dispersion, Score SUS, préférence et coef. de dérive	« Puissance d'impact » (moyenne des rangs)
A+AR	2 ; 1 ; 1 ; 1 ; 1 « stable »	0,93
P+AR	1 ; 2 ; 2 ; 2 ; 1 « stable »	0,80
A+SR	3 ; 4 ; 3 ; 3 ; 3 « -0.09X »	0,26
P+SR	4 ; 3 ; 4 ; 4 ; 4 « -0.20X »	0,07

**Tableau 4-11 : Synthèse de l'exp. 1.2. Pour chaque combinaison, son ordre selon cinq mesures est donné. La moyenne des ordres donne la puissance d'impact des combinaisons. Sachant que plus la « puissance d'impact » est grande, plus la technique sera efficace et satisfaisante comparée à une technique d'une puissance inférieure.**

Cette nouvelle mesure permet de comparer les techniques entre elles, en fonction de leur classement selon cinq mesures. Si nous utilisons cette nouvelle mesure pour classer les huit techniques de l'exp. 1.1 et les quatre techniques de l'exp. 1.2, nous obtenons le Tableau 4-12 suivant :

Classement	Techniques	Puissance d'impact	Forte Dichotomie
1	« Règle-Graduée », A+AR, 6 Réf.	0,93	AR
2	« PalmGauge » ; P+AR, X Réf.	0,84	
3	« Slider » ; A+AR, 2 Réf.	0,80	
4	« BipMultiple », P+AR, 6 Réf.	0,80	
5	« Pédale » ; A+AR, 2 Réf.	0,76	
6	« Volant » ; A+AR, 3 Réf.	0,75	
7	« DualStimuli » ; P+AR, 1 Réf.	0,58	
8	« Règle-Lisse », A+SR	0,31	SR
9	« BipSolo », P+SR	0,20	
10	« SingleStimulus » ; P+SR	0,18	
11	« ClicOnMe » ; A+SR	0,14	
12	« Pipe » ; A+SR	0,09	

**Tableau 4-12 : Classement des 12 techniques en fonction de leur « puissance d'impact ». Notons que les techniques de l'exp. 1.1 gardent leur hiérarchie.**

Il est intéressant de noter pour les techniques « A+SR » et en termes de « puissance d'impact », que le codage LONGUEUR (« Règle-Lisse ») est supérieur au codage DURÉE (« ClicOnMe »), lui-même supérieur au codage ANGLE (« Pipe »). De même pour les techniques « A+AR », LONGUEUR > ANGLE.

En conclusion, les meilleures techniques pour estimer une valeur entre 1 et 100 sont celles avec plusieurs références et ce, peu importe le type (logiciel, physique, « choisi par l'utilisateur ») et de préférence « Active » pour satisfaire l'utilisateur. Nous avons vu que la main permet une meilleure efficacité afin de percevoir et donc d'estimer une valeur. Nous avons vu également qu'il était possible d'estimer efficacement une valeur avec la durée d'un son, mais uniquement lorsque des références explicites sont mises à disposition de l'utilisateur. Rappelons enfin que notre meilleure technique, « Règle-Graduée », est « Active AvecRéférences » pour une précision inférieure à trois et une dispersion inférieure à six. Notre deuxième meilleure technique, « PalmGauge », est « Passive AvecRéférences » pour une précision de moins quatre et une dispersion inférieure à huit. Et notre troisième (*ex æquo*) meilleure technique, « BipMultiple », est « Passive AvecRéférences » pour une précision de moins deux et une dispersion de six. Ces trois techniques sont efficaces et satisfaisantes, la première est haptique, la deuxième est tactile (ici, cutané) et la troisième est auditive.

# **Chapitre 5**

## **Deuxième étude :**

**interactions non visuelles  
pour faciliter le rendu d'une  
séquence de couples  
direction-distance**



## 5 Deuxième étude : interactions non visuelles pour faciliter le rendu d'une séquence de couples direction-distance

Dans le chapitre précédent, nous avons conçu et évalué des techniques d'interaction, non visuelles et sans parole, dans une tâche d'estimation d'une valeur variant entre 1 et 100. Certaines techniques ont été efficaces et jugées satisfaisantes par les utilisateurs, notamment les techniques avec références. Dans ce chapitre, nous nous sommes intéressés au transfert d'informations multiples par des techniques utilisant certaines propriétés de la multimodalité (cf. section Multimodalité) et par une technique de contrôle qui se base exclusivement sur de la parole. Nous avons mesuré l'impact des techniques en termes de perception et de mémorisation des informations, par leur restitution immédiate puis différée desdites informations. Ainsi, nous avons conçu, évalué et comparé entre elles quatre techniques d'interaction, dont trois utilisent la parole. Toutefois, la tâche sera plus complexe que d'estimer une valeur car dans cette deuxième étude, les techniques sont utilisées pour transmettre et restituer un itinéraire composé d'une séquence de couples direction-distance.

### 5.1 Introduction à la deuxième étude

Dans des situations où la perception visuelle est fortement contrainte ou impossible (ex. déficience visuelle), il est nécessaire de rendre perceptible l'information par une autre modalité sensorielle. Dans ces travaux, nous nous intéressons plus particulièrement à la transmission non visuelle d'une séquence de couples de valeurs de type direction-distance, qui peut avoir un rôle informationnel (liste d'instructions), de prévention ou de guidage. Cette deuxième étude (exp. 2.1) s'inspire du cas particulier de la transmission d'une feuille de route pour piéton non-voyant, générée à partir d'applications facilement disponibles et gratuites mais non dédiées aux déficients visuels (ex. Google Maps). La tâche générique consiste alors en la transmission sous forme non visuelle d'une séquence de couples direction-distance (voir Figure 5-1).

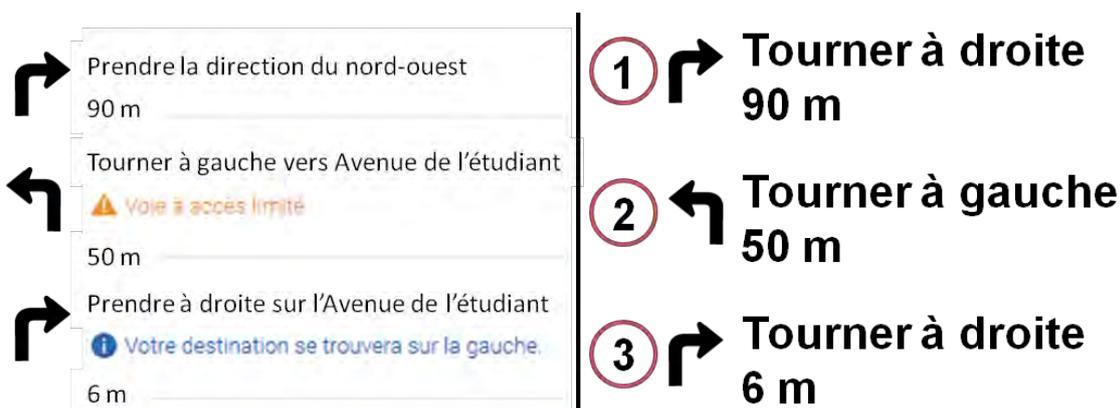


Figure 5-1 : Exemple d'un itinéraire de 3 segments (couples) ; à gauche la version Google Maps et à droite la version simplifiée transmise par nos techniques, soit 6 valeurs : Droite-90 ; Gauche-50 ; Droite-6

Dans la majorité des cas, la lecture de la feuille de route s'effectue au moyen d'un lecteur d'écran. Celui-ci permet de parcourir les informations de façon séquentielle et de les restituer à l'utilisateur sous forme non visuelle : il s'agit le plus fréquemment d'une forme audio, au moyen d'une synthèse vocale, ou d'une forme tactile, à l'aide d'une plage braille. Cependant, la mise en œuvre de ces solutions alternatives demeure confrontée aux limites de l'utilisateur en termes de mémorisation : un parcours est, en effet, constitué d'une séquence de couples de valeurs de plus de sept, plus ou moins deux, mnèmes (limite bien identifiée dans la littérature (Miller 1956)). Or, les travaux de Dubois (Dubois, Gyselinck, and Choplin 2003) ont établi l'intérêt de la multimodalité sur la mémoire de travail dans le cadre de la perception de données de différentes natures. Nous proposons donc d'explorer l'apport des techniques d'interaction multimodale non visuelle pour optimiser la perception et la mémorisation d'une séquence de couples de données de type propriété-mesure et plus précisément, de couples direction-distance.

A ces fins, nous avons conçu différentes techniques d'interaction non visuelle et multimodale directement inspirées de celles parmi les plus performantes dans la première étude (cf. Première étude : techniques d'interaction non visuelle pour estimer une valeur) et, avec l'objectif de favoriser la mémorisation de couples direction-distance, en répartissant la quantité d'information sur plusieurs modalités complémentaires. Ce chapitre présente les techniques évaluées et synthétise les résultats préliminaires obtenus suite à l'évaluation de ces différentes techniques.

Dans le chapitre 2<sup>41</sup> de ce manuscrit, nous avons vu de nombreuses techniques pour transmettre une ou des informations. Dans ce contexte de transmission d'informations

<sup>41</sup> Etat de l'art sur la transmission d'information basée sur les modalités tactiles et/ou auditives

multiples, nous avons sélectionné quelques exemples de références que nous présentons ci-après.

## 5.2 Rappel sur l'état de l'art : transmission de direction et de distance

Compenser un déficit de perception visuelle conduit à utiliser les modalités sensorielles auditive et/ou tactile pour transmettre les informations. Par exemple, il est possible de concevoir des techniques d'interaction basées sur la sonification d'informations (Barrass and Kramer 1999; Hermann *et al.* 2011) pour transmettre des informations de distance et de direction (Yamano *et al.* 2012). L'intensité ou la fréquence d'un son peuvent également être utilisées pour encoder et transmettre une mesure (Donald H. Mershon and King 1975; Zahorik, Brungart, and Bronkhorst 2005). D'autres attributs du son ont, également, été exploités dans la littérature comme véhicules d'une information à transmettre, en particulier à l'aide d'un *panning* (balance des hautparleurs gauche et droit), d'une réverbération ou de la spatialisation du son (Fontana and Rocchesso 2008; Larsen *et al.* 2013; Mills 1958). Certaines études ont aussi cherché à combiner ces différentes techniques pour transmettre un couple de valeurs (Nigay and Coutaz 1993), notamment un couple direction-distance, dans des contextes où l'information n'a pas besoin d'être mémorisée car elle doit être utilisée immédiatement, comme dans le cas du guidage (Rümelin *et al.* 2011) ou d'une alerte auditive ou tactile (Fujimoto and Turk 2014).

Ces travaux issus de la littérature s'inscrivent donc dans un usage immédiat des informations perçues. Notre situation d'étude dépasse ce contexte : le piéton doit se souvenir de l'itinéraire à parcourir pour pouvoir, par exemple, choisir le plus optimal, limiter le stress du piéton aveugle face à un itinéraire inconnu, ou transmettre l'itinéraire à une autre personne (aveugle). Nos travaux visent donc à explorer les apports et les limitations de la multimodalité pour la perception et la mémorisation à moyen termes d'une séquence d'informations de type propriété-mesure.

## 5.3 Pré-études et choix de design

Dans cette étude, nous avons utilisé le cadre d'analyse, afin de concevoir de nouvelles interactions permettant de transmettre une séquence de couples direction-distance. Pour ce faire, nous avons spécifié l'aspect temporel et le partage de dispositif dans la description.

**PLUSIEURS COMBINAISONS POSSIBLES POUR CONSTRUIRE UN COUPLE DIRECTION-DISTANCE**

Premièrement, nous avons repris nos huit techniques de base décrites dans le chapitre 3<sup>42</sup> et évaluées dans le chapitre 4 dans une tâche d'estimation de valeur. Cependant, nous avons ajouté des techniques se basant sur le type de codage LANGAGE, c'est-à-dire : soit sur un langage naturel oral comme la parole, soit sur un langage abstrait tactile comme les tactons.

Deuxièmement, dans le cas d'une transmission séquentielle de chacune des informations, c'est-à-dire l'énumération d'une liste d'informations, nous avons décidé de fixer l'ordre de présentation tels que « Google Maps » ou « Mappy » le proposent : chaque couple est présenté les uns à la suite des autres et dans chaque couple, la direction sera donnée avant la distance (par ex. Droite-90 ; Gauche-50 ; Droite-6). Ainsi, nos techniques seront compatibles avec ce type d'application.

Troisièmement, nous avons essayé plusieurs combinaisons <modalités - types de transmission> pour transmettre un couple direction-distance (voir Tableau 5-1) dans le cadre d'une étude comparative. Nous avons, tout d'abord, conçu une technique contrôle, monomodale, basée exclusivement sur la synthèse vocale (la parole). Celle-ci correspond au rendu des informations, grâce au lecteur d'écran et à la synthèse vocale, et servira de condition contrôle. A partir d'un ensemble de séances de brainstorming, impliquant des spécialistes en IHM, en cognition spatiale et en technologie de suppléances pour déficient visuel (DV), nous avons aussi conçu des techniques multimodales basées sur les deux grandes familles de modalités tactiles et auditives. Au cours de nombreuses itérations, nous avons testé avec trois sujets maximum, plusieurs techniques d'interaction multimodale, qui diffèrent notamment par le rendu de l'information et les associations de modalité. Le Tableau 5-1 synthétise les différentes options explorées et testées.

Direction (ligne) \ Distance (colonne)	TextToSpeech	Sonification	Tacton	Vibrotactile
<b>Palm Gauge</b>	PG-TTS (A)	PG-BipSolo (A)	PGSaccade-PG (A)	PG-Tactons (A)
<b>Slider</b>	Slider-TTS (A)			Slider-Tactons (A)
<b>TextToSpeech</b>	<u>TTS-TTS (E)</u>			
<b>Sonification</b>	TTS-Lat. (S)	Méloдие (S)		

<sup>42</sup> Illustration de génération de techniques avec notre cadre d'analyse

**Tableau 5-1 : synthèse de plusieurs combinaisons de techniques et aspects temporels pour transmettre un couple direction-distance. Les symboles entre parenthèses indiquent les différents aspects temporels de la transmission d'information : (E) = exclusif ; (A) = Alterné ; (C) = concurrent ; (S) = synergique. La technique « TextToSpeech » est monomodale et utilise exclusivement la parole pour transmettre le couple direction-distance, c'est notre technique contrôle. La modalité « sonification » repose sur l'utilisation des propriétés d'un son pour transmettre une information, telles que la durée (BipSolo) ou la fréquence (Mélodie) ou encore la spatialisation de la parole (TTS-Lat). Les tactons sont des messages abstraits vibrotactiles ou de mouvements saccadés (entrecoupés).**

Pendant cette phase de recherche de conception de techniques innovantes de transmission d'informations, nous avons aussi testé plusieurs ensembles de valeurs pour les directions (3 ou 5 valeurs) et d'échelles pour les distances (1-10, 1-100). En effet, dans les itinéraires piétons générés, via des applications tels que Google Maps, la première direction à suivre peut être une direction intermédiaire, par exemple la direction nord-ouest se trouve entre la direction nord et la direction ouest. Ainsi, nous avons testé cinq valeurs pour les directions, afin de prendre en compte ce cas particulier de première direction (gauche, mi-gauche, face, mi-droite et droite). Cependant, nous avons considéré que l'orientation initiale de l'utilisateur sur la bonne direction ne fait pas partie de la tâche de percevoir et mémoriser un itinéraire, mais fait partie d'une tâche de guidage en situation de navigation. Auquel cas, des techniques d'interaction tel que le *PocketNavigator* de (Pielot *et al.* 2012) ou la ceinture vibrante de (Tsukada and Yasumura 2004) peuvent remplir ce rôle. Nous avons, également, testé deux échelles différentes. La petite plage de valeur (1 à 10) se réfère aux multiples techniques d'interaction exposées dans l'état de l'art de cette thèse, où les distances estimées sont comprises dans cet ordre de grandeur. Toutefois, afin de rester cohérent avec nos techniques évaluées de l'expérience 1 et sur lesquelles nous nous basons ici, nous avons gardé la même plage de valeur ; c'est-à-dire des valeurs entre 1 et 100. Ainsi, nous avons gardé le format suivant du couple direction-distance à transmettre : trois directions dans un repère égocentré (gauche, droite et en face) et la distance sur une échelle allant de 1 à 100. Notons que ces deux choix sont également similaires aux générateurs d'itinéraire pour piéton de type Google itinéraire.

A l'issue de ces pré-tests, outre la technique monomodale de contrôle, trois techniques multimodales ont finalement été retenues pour transmettre une séquence de couples de valeurs direction-distance. Nous décrivons ces quatre techniques ci-après.

#### **5.4 Description des quatre techniques retenues**

Avant d'aborder cette section, nous proposons un rappel sur la description de techniques avec notre cadre d'analyse.

Le cadre d'analyse permet de décrire une technique selon huit paramètres de base et quatre spécifications (voir Quatre spécifications : l'échelle, la tâche utilisateur, l'aspect temporel et le dispositif partagé). Nous pouvons ainsi lister graphiquement des facteurs de conception d'une technique, mais aussi faire une spécification détaillée de celle-ci. Ainsi, pour chaque technique, nous donnerons son nom, son principe d'utilisation, une illustration d'utilisation, sa description graphique et le matériel nécessaire à sa mise en œuvre.

### 5.4.1.1 TextToSpeech (TTS)

La technique de contrôle est la technique TextToSpeech (TTS).

**Principe** : similaire à celui d'un lecteur d'écran, une synthèse vocale présente l'itinéraire comme une liste de segments (cf. Figure 5-1). Dans notre cas, cela signifie qu'un itinéraire est lu comme une succession de couple direction-distance où la direction est lue avant la distance (cf. Figure 5-2).

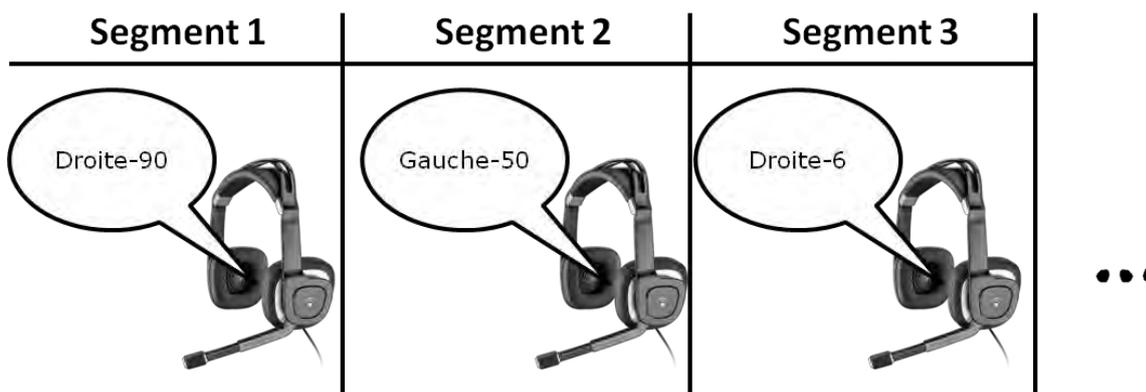


Figure 5-2 : illustration de la technique TTS. Exemple d'utilisation : liste des segments d'un itinéraire de taille trois prononcée par une synthèse vocale. Les «...» indiquent que la taille de l'itinéraire peut être augmentée.

**Cadre analytique** : technique d'interaction monomodale passive sans référence. Le signal, en sortie du système, vers l'utilisateur s'effectue avec un casque audio capable de transmettre un stimulus audio de manière continue, et avec le type de codage LANGAGE (parole « X » == valeur de X). Les informations des couples sont transmises en exclusif, et les couples sont transmis en séquence (cf. Figure 5-3).

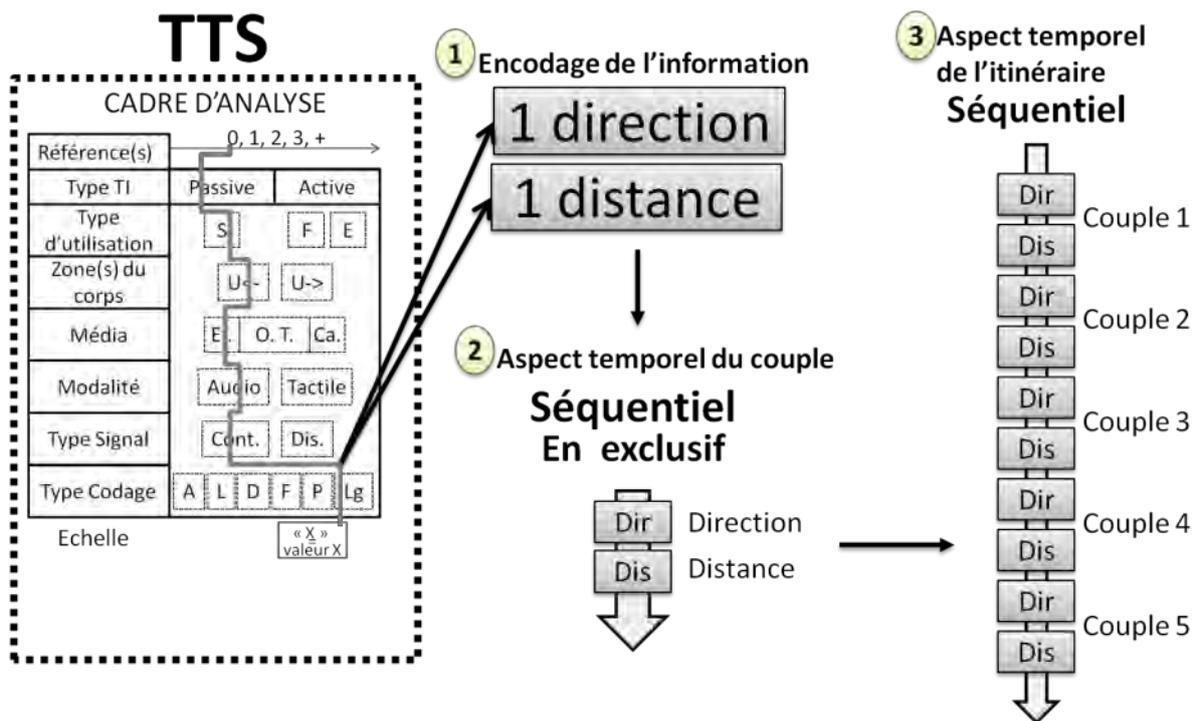


Figure 5-3 : illustration de l'analyse de la technique TextToSpeech (TTS).

**Matériel** : un casque audio [Plantronics audio 650 USB] et le logiciel de synthèse vocale ["Voix de Virginie" (ScanSoft Virginie\_Dri40\_16kHz), SAPI].

#### 5.4.1.2 TextToSpeech Latéralisé (TTS-Lat)

La technique TextToSpeech Latéralisé (TTS-Lat) repose sur la modalité de synthèse de parole et sur la modalité de spatialisation auditive. Seule la distance est vocalisée, mais son rendu est spatialisé pour indiquer la direction : la source sonore se trouve artificiellement à gauche, en face ou à droite. TTS-Lat est donc une technique multimodale avec fusion d'informations : la direction et la distance sont deux informations fusionnées en une parole spatialisée.

**Principe** : une synthèse vocale « lit » l'itinéraire comme une liste de distance, la spatialisation de la source sonore indique la direction (cf. Figure 5-4).

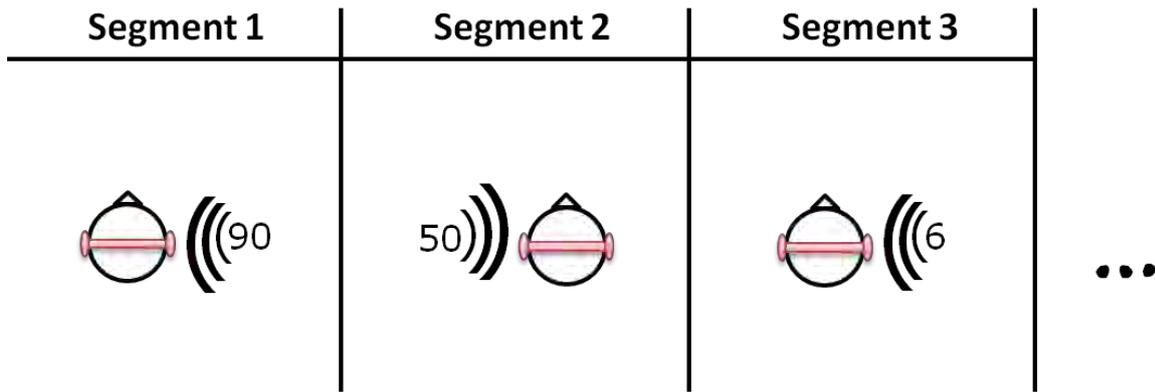


Figure 5-4 : illustration de la technique TTS-Lat. Exemple d'utilisation : liste des segments d'un itinéraire de taille trois (Droite-90 ; Gauche-50 ; Droite-6) prononcée par une synthèse vocale dont la source sonore est spatialisée. Les «...» indiquent que la taille de l'itinéraire peut être augmentée.

**Cadre analytique** : technique d'interaction multimodale passive. Le signal en sortie du système vers l'utilisateur s'effectue en synergie avec un effecteur partagé, capable de transmettre un stimulus audio de manière continue et dont l'information est composée de deux sous-unités : l'une avec le type de codage ANGLE (3 spatialisations pour les 3 directions) et l'autre avec le type de codage LANGAGE (parole « X » == valeur de X). La spatialisation du son donne trois références explicites : gauche, en face et droite. Les informations des couples sont transmises en synergie et les couples sont transmis en séquence (cf. Figure 5-5).

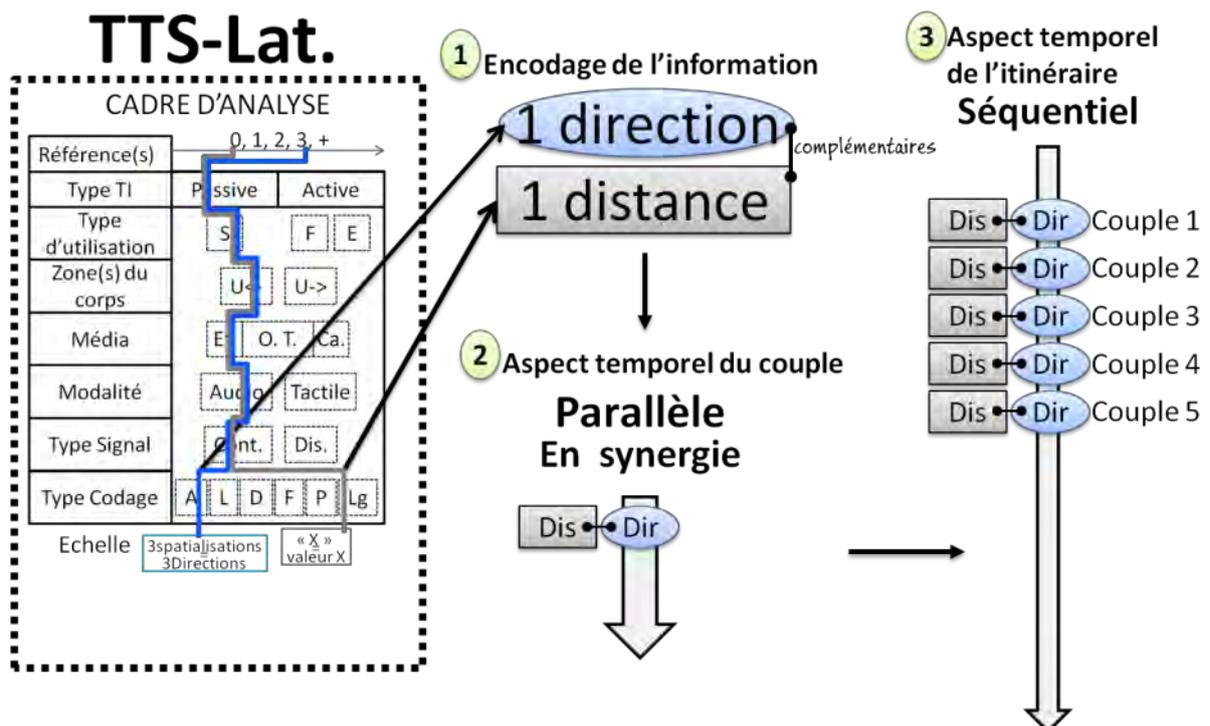


Figure 5-5 : illustration de l'analyse de la technique TextToSpeech Latéralisé (TTS-Lat).

**Matériel** : un casque audio [Plantronics audio 650 USB], la synthèse vocale ["Voix de Virginie" (ScanSoft Virginie\_Dri40\_16kHz), SAPI] et le moteur de son 3D [IrrKlang<sup>43</sup>, API C#].

### 5.4.1.3 Palm Gauge-TextToSpeech (PG-TTS)

La technique Palm Gauge-TextToSpeech (PG-TTS) repose sur la rotation d'un curseur placé sous la main qui se positionne à gauche, en face ou à droite puis, la distance est vocalisée. PG-TTS est une technique multimodale séquentielle.

**Principe** : Un itinéraire est lu segment par segment, et chaque segment est transmis par deux modalités : la première donne la direction via le Palm Gauge, la seconde donne la distance par synthèse vocale (cf. Figure 5-6). Cependant, le PalmGauge est utilisé différemment : au début de l'itinéraire, l'aiguille est orientée en face et va se déplacer soit vers la gauche (direction gauche), soit vers la droite (direction droite). Dans le cas où la direction courante correspond à celle déjà indiquée par l'aiguille alors l'aiguille fait un petit va-et-vient de  $\pm 5^\circ$ . Toutefois, dans le cas où la direction est à l'opposée de la direction indiquée (gauche puis droite ou droite puis gauche), l'aiguille se déplace de  $180^\circ$ . A la fin de l'itinéraire l'aiguille retourne à sa position initiale, c'est-à-dire au centre.

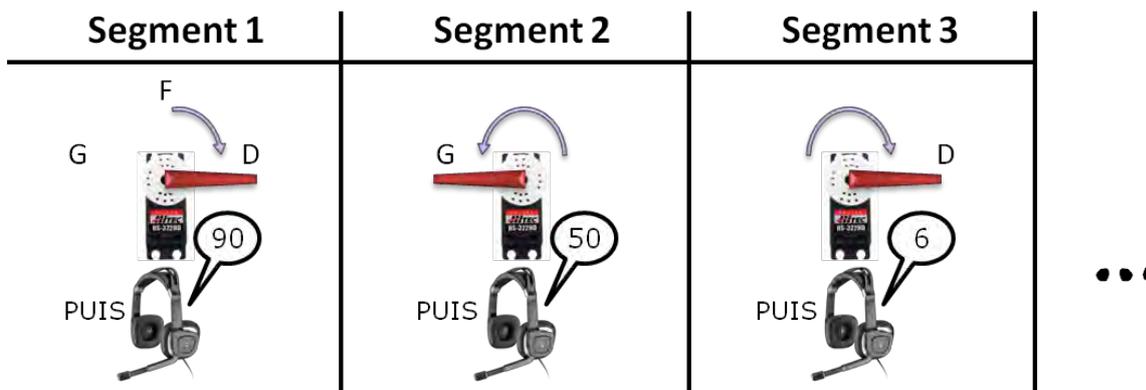


Figure 5-6 : illustration de la technique PG-TTS. Exemple d'utilisation : liste des segments d'un itinéraire de taille trois (Droite-90 ; Gauche-50 ; Droite-6), où les directions sont données par une direction tactile et les distances par synthèse vocale. Les «...» indiquent que la taille de l'itinéraire peut être augmentée.

**Cadre analytique** : technique d'interaction multimodale passive. Le signal en sortie du système vers l'utilisateur s'effectue en alternance avec deux effecteurs. L'un capable de transmettre un stimulus audio de manière continue dont l'information est codée avec le type de codage LANGAGE (parole « X » == valeur de X). L'autre capable de transmettre un

<sup>43</sup> <http://www.ambiera.com/irrklang/>

stimulus tactile continu dont l'information est codée avec le type de codage ANGLE (3 angles == 3 directions). Les informations des couples sont transmises en alternance et les couples sont transmis en séquence (cf. Figure 5-7).

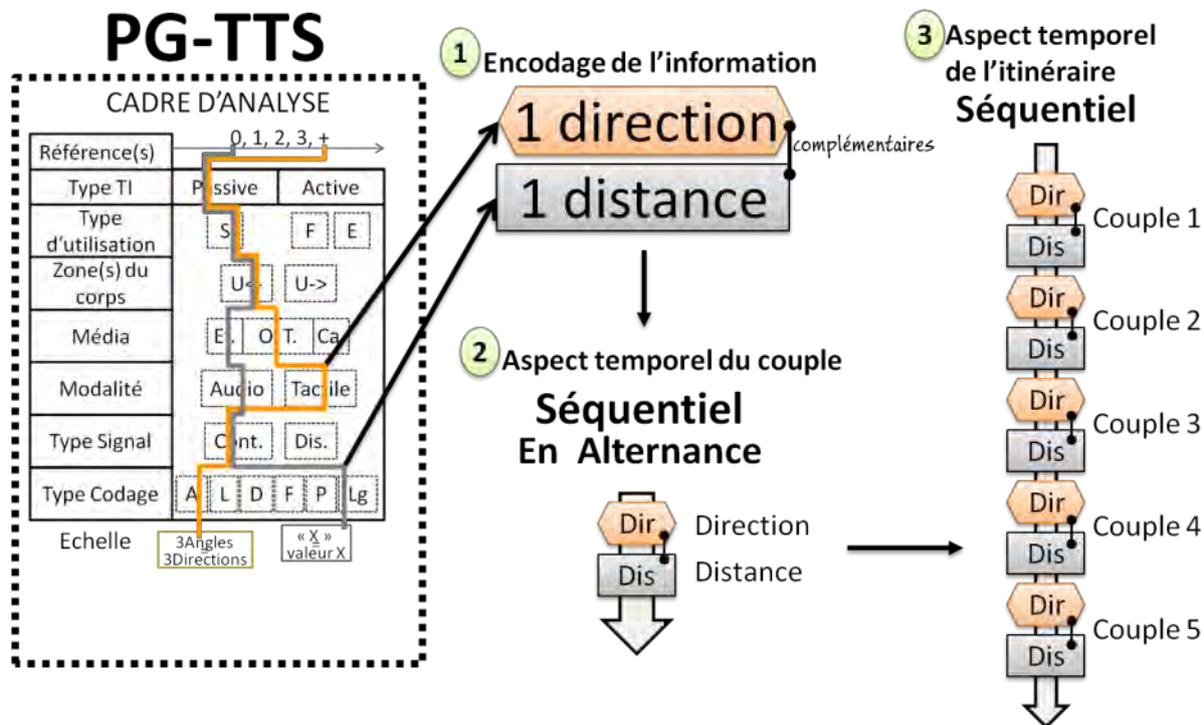


Figure 5-7 : illustration de l'analyse de la technique Palm Gauge-TextToSpeech (PG-TTS).

**Matériel** : un casque audio [Plantronics audio 650 USB], la synthèse vocale ["Voix de Virginie" (ScanSoft Virginie\_Dri40\_16kHz), SAPI]. Le Palm Gauge est constitué d'un servomoteur [HS-322HS Servo] relié à une carte Arduino [1061\_1 PhidgetAdvancedServo 8-motor]. Le support pour la main est un [cylindre cartonné] de 50 mm de hauteur, 90 mm de diamètre et de 5 mm d'épaisseur dans lequel est fixé le servomoteur en son centre. L'aiguille est faite à partir de la peau d'une [balle de tennis] attachée au moyeu du servomoteur avec du fil de fer et du ruban adhésif.

#### 5.4.1.4 Mélodie

La technique Mélodie repose sur une note de musique (timbre et hauteur), sa spatialisation et sa durée. La note et sa spatialisation correspondent de manière redondante à la direction et la durée correspond à la distance (cf. Figure 5-8). L'enchaînement des notes correspond alors à une mélodie qui traduit une séquence de couples direction-distance. Mélodie est une technique multimodale avec fusion de données : la direction et la distance sont deux informations fusionnées en une note spatialisée. Elle a été conçue afin de permettre aux utilisateurs de « chanter la mélodie » pour s'en rappeler. Notons que Mélodie est la seule

technique utilisant le type de codage DURÉE pour la distance, les autres techniques utilisent la parole.

**Principe** : L'utilisateur écoute une mélodie, les notes et leur spatialisation portent l'information des directions. La durée des notes portent l'information de distance (voir Figure 5-8).

**Illustration** :

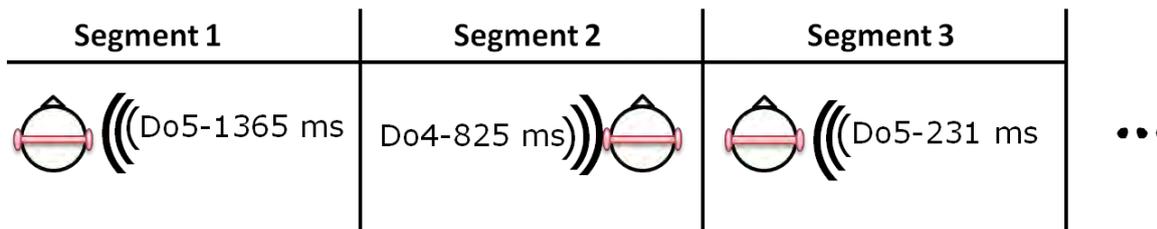


Figure 5-8 : illustration de la technique Mélodie. Exemple d'utilisation : liste des segments d'un itinéraire de taille trois (Droite-90 ; Gauche-50 ; Droite-6) sous la forme d'une mélodie. La note et sa latéralisation indique la direction. Sa durée indique la distance. Les «...» indiquent que la taille de l'itinéraire peut être augmentée.

**Cadre analytique** : technique d'interaction multimodale passive. Le signal en sortie du système vers l'utilisateur s'effectue en synergie avec un effecteur partagé capable de transmettre un stimulus audio de manière continue et dont l'information est composée de trois sous-unités : la première est codée avec le type de codage ANGLE (3 spatialisations == 3 directions), la deuxième est codée le type PULSE (3 notes == 3 directions) et la troisième est codée avec le type DURÉE (1500 ms = 100 valeurs). La spatialisation du son et sa note sont des modalités redondantes donnant trois références explicites :

- Direction EN FACE, note de l'instrument PLANET en do3 majeur spatialisé en face ;
- Direction GAUCHE, note de l'instrument PICOLO en do4 majeur spatialisé à gauche ;
- Direction DROITE, note de l'instrument PICOLO en do5 majeur spatialisé à droite ;

Les informations des couples sont transmises en synergie et les couples sont transmis en séquence (voir Figure 5-9).

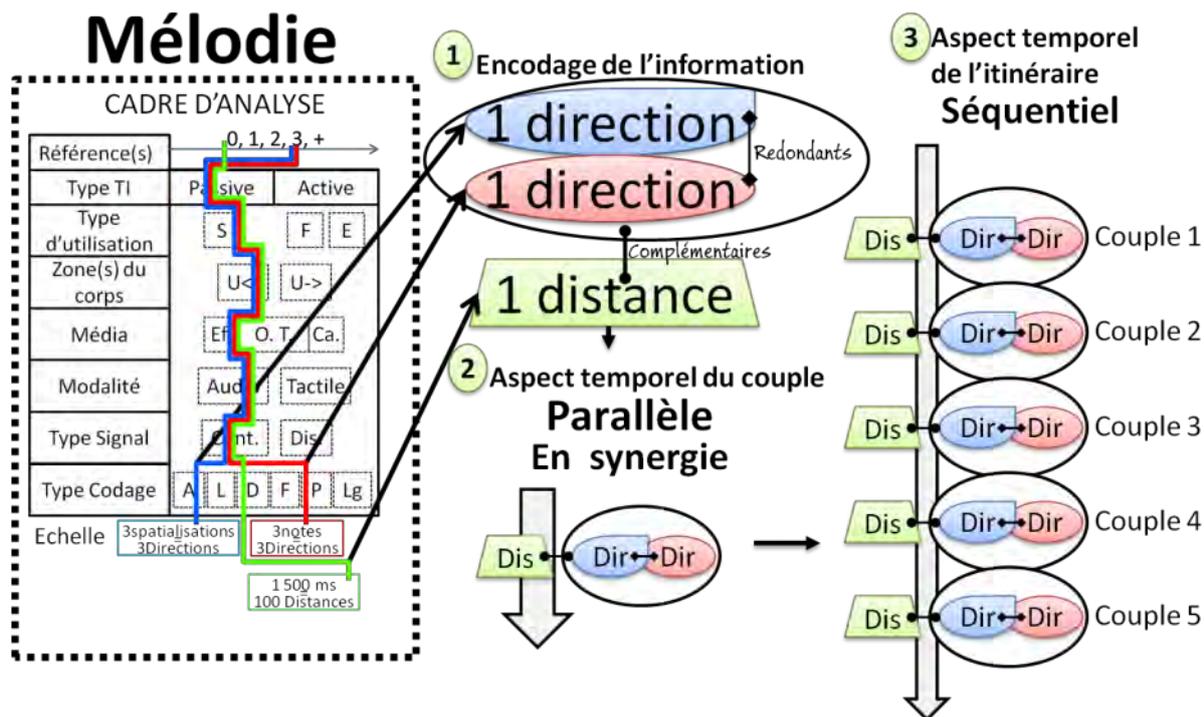


Figure 5-9 : illustration de l'analyse de la technique Mélodie.

**Matériel :** un casque audio [Plantronics audio 650 USB], le moteur de son 3D [IrrKlang<sup>44</sup>, API C#] et le synthétiseur [Piano Virtuel Midi V6.1]<sup>45</sup> pour générer les notes.

### 5.5 Hypothèses de travail

Nous avons émis trois ensembles d'hypothèses basés sur les raisonnements suivants :

1. TTS (condition contrôle) surcharge l'utilisateur d'une succession d'informations verbales et devrait être la moins utilisable ;
2. TTS-Lat, qui est la forme optimisée de TTS, devrait être plus utilisable ;
3. PG-TTS et Mélodie sont des techniques, reposant sur deux principes différents, permettant de mémoriser plus facilement les couples de valeurs. PG-TTS sépare les informations sur les modalités tactiles et auditives. Mélodie est conçue pour faciliter la mémorisation et le rappel en chantant. Ces deux techniques devraient donc être les plus utilisables dans ce contexte.

Afin d'évaluer ces techniques selon ces hypothèses, nous avons conduit une expérimentation.

### 5.6 Protocole expérimental de l'expérience 2.1

<sup>44</sup> <http://www.ambiera.com/irrklang/>

<sup>45</sup> Piano Virtuel Midi V6.1, de Laurent minot, <http://www.madolo.fr/piano-virtuel-midi/>

Les sujets étaient assis sur une chaise, les yeux bandés. Pour chaque sujet, l'expérimentation a été décomposée en quatre blocs, chacun correspondant à une technique d'interaction (TTS, TTS-Lat, Mélodie et PG-TTS) dont l'ordre a été contrebalancé.

Un bloc (cf. Figure 5-10) commençait par quatre essais sur des itinéraires de familiarisation. L'évaluation comportait ensuite deux séries de quatre itinéraires. Les itinéraires de la première série étaient composés de cinq segments (soit 5 couples de directions et distances choisis de façon pseudo-aléatoire) ; les itinéraires de la seconde en contenaient sept. Après chaque itinéraire, le participant devait restituer oralement l'itinéraire perçu (l'expérimentateur le consignait par écrit). Cette tâche de restitution comportait deux rappels (R1 et R2) séparés par une tâche d'interférence (Peterson and Peterson 1959) (dix décréments 3 par 3 à partir d'un nombre aléatoire) visant à établir l'apport de la technique en termes de mémorisation.

A l'issue d'un bloc, les sujets répondaient au questionnaire NASA-TLX qui permet d'évaluer la charge cognitive associée à la technique utilisée. Les deux séries étaient séparées par une courte pause (30 s). Les blocs étaient séparés par une pause plus longue (180 s).

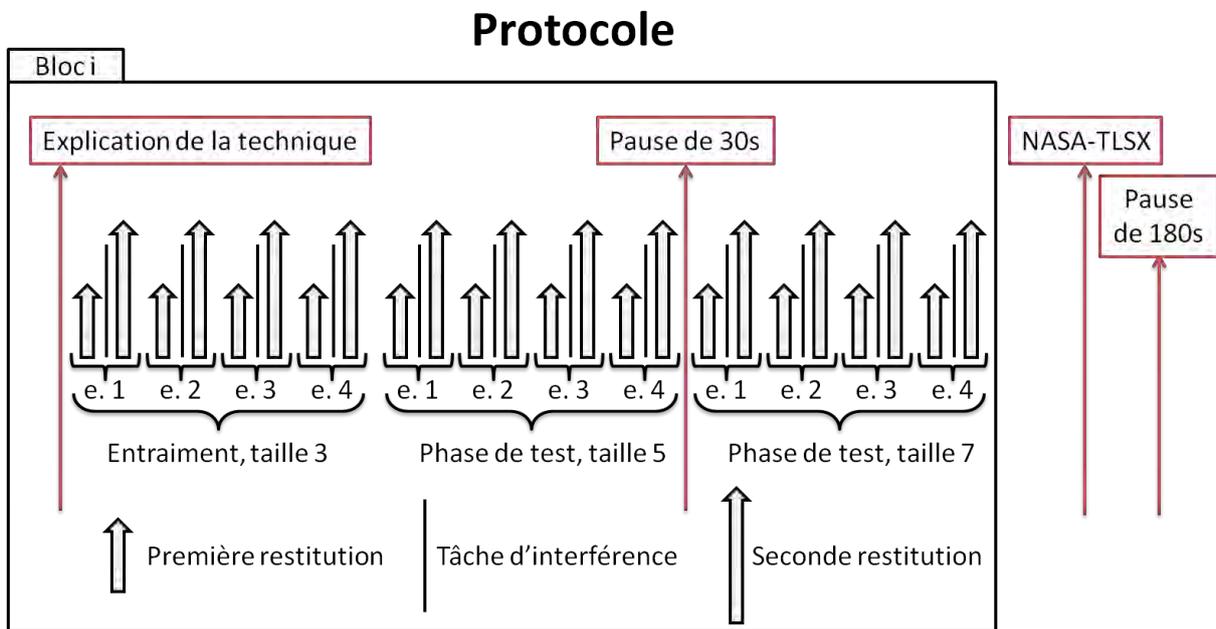


Figure 5-10 : illustration du protocole pour un bloc.

Une fois tous les blocs passés, les sujets devaient ordonner les quatre techniques par ordre de préférence. La durée moyenne de l'expérimentation par sujet était de deux heures.

Cette expérimentation suit donc une conception intra-sujet 4x2x2 avec comme facteurs : les techniques d'interaction, la taille (5 puis 7) et la phase de restitution (1 puis 2). Compte tenu

des ensembles d'hypothèses énoncés, nous avons analysé la satisfaction des utilisateurs vis-à-vis des techniques ainsi que la *robustesse* de la mémorisation et l'*efficacité* de chacune des techniques.

La *robustesse* représente le nombre d'itinéraires restitués de la même façon lors d'un premier rappel puis lors d'un second rappel différé (conformes). La robustesse était évaluée pour les directions seules, les distances seules puis pour les couples (direction ; distance). Nous exprimons cette mesure avec un taux : nombre d'itinéraires conformes sur le nombre d'itinéraires totaux.

L'*efficacité* représente le nombre d'itinéraires restitués correctement lors du premier rappel. L'efficacité a été évaluée pour les directions seules, les distances seules puis pour les couples. Nous exprimons cette mesure avec un taux : nombre d'itinéraires corrects sur le nombre d'itinéraires totaux.

Onze participants ont pris part à l'évaluation : 8 sujets Non Déficients Visuels (Non-DV, âge :  $26,7 \pm 1,7$ ) et 3 sujets Déficients Visuels (DV, âge :  $35,6 \pm 11,5$ ). Nous relatons dans un premier temps les résultats relatifs à la population des Non-DV seuls, puis conduisons une analyse succincte des résultats pour les DV.

## 5.7 Analyse des résultats des sujets Non-DV

Lors des comparaisons, nous avons utilisé les tests non-paramétriques de Friedman ou Wilcoxon avec un seuil de significativité à 0,05 car les données ne suivaient pas une loi normale (Shapiro-Wilk normality test,  $p < 0,01$ ) et aucune transformation n'a permis de les ramener à une forme normale. Les graphiques représentent des moyennes et des écarts-types.

### **ÉTUDE DU NOMBRE D'ÉLÉMENTS CORRECTS LORS DE LA PREMIÈRE PHASE DE RESTITUTION**

Nous rappelons qu'un itinéraire de taille 5 (resp. 7) correspondaient à une séquence de 5 (resp. 7) segments avec une direction et une distance pour chaque, soit 10 éléments en taille 5 et 14 en taille 7.

**Moyenne par taille.** Il ressortait tout d'abord que lors de la première phase de restitution (R1) le nombre moyen d'éléments corrects pour un itinéraire, en fonction de la taille et toutes techniques confondues était en moyenne de 7,6 pour la taille 5 et de 7,8 pour la taille 7.

**Moyenne par technique et par taille.** En affinant cette analyse en prenant en compte les techniques d'interaction et la taille de l'itinéraire, il en ressortait que pour les itinéraires de 5 segments (couple direction-distance), la plus petite moyenne était de 6,9 et la plus grande était de 8,5. Elles étaient de 7,4 et 8,0 pour des itinéraires de taille 7. Les travaux de Miller (Miller 1956) ont montré que l'empan mnésique moyen à court terme est de  $7 \pm 2$  éléments. Les techniques que nous avons conçues permettent évidemment d'atteindre cette moyenne, mais certaines permettaient de le dépasser (cf. Figure 5-11).

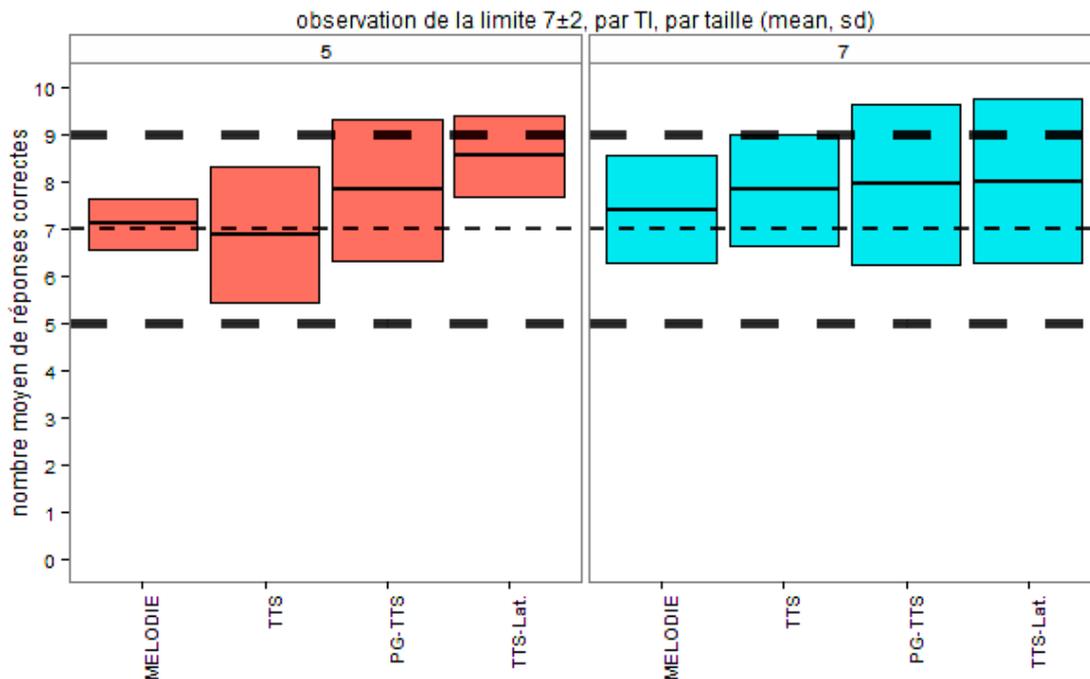


Figure 5-11 : Moyenne du nombre d'éléments correctement restitués (R1) par taille (5 à gauche et 7 à droite) et par technique. Nous pouvons voir que les moyennes de réponses correctes pour MELODIE et TTS étaient proches de 7. Par contre les moyennes pour PG-TTS et TTS-Lat. étaient supérieures, avec un écart-type qui passait la barre des 9 éléments corrects.

**ÉTUDE DES RESULTATS EN TERMES D'EFFICACITE**

L'efficacité représente ici le taux d'itinéraires corrects lors de la première restitution (R1), c'est-à-dire nombre d'itinéraires corrects sur le nombre d'itinéraires totaux. Pour rappel, un itinéraire est considéré comme correct si les 5 (respectivement 7) segments sont restitués correctement.

**Efficacité des itinéraires selon les couples, c'est-à-dire où tous les couples sont corrects, par technique et par taille.** Seuls les itinéraires de taille 5 font apparaitre des différences significatives entre techniques (Friedman :  $X^2 = 12,7$ ,  $df = 3$ ,  $p < 0,01$  ; voir Figure 5-12). TTS-Lat. (0,44) permet d'obtenir une meilleure efficacité que celle obtenue avec les techniques Mélodie (0,03) et TTS (0,06). De même, l'efficacité obtenue pour la technique PG-TTS (0,31)

est significativement supérieure à celle observée pour Mélodie (0,03) et TTS (0,06). Notons, qu'il est quasiment impossible de restituer un itinéraire correctement lorsque celui-ci est constitué de sept segments.

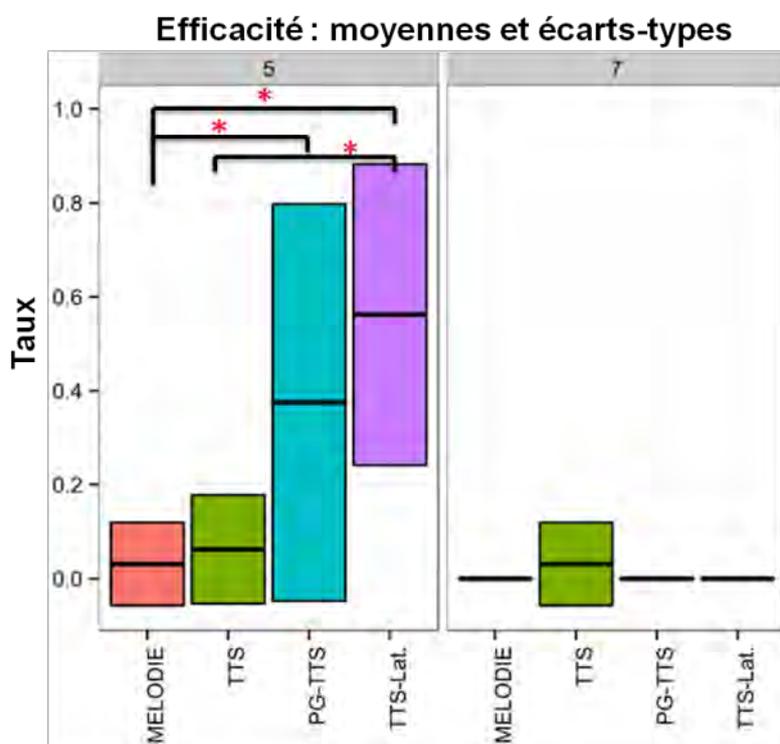


Figure 5-12 : Efficacité des techniques chez les non-DV, selon deux tailles (5 segments à gauche et 7 à droite). (Wilcoxon signed rank test ; \* : p<0,05).

**Efficacité des itinéraires selon les directions seules et les distances seules**, c'est-à-dire où toutes les directions sont correctes et toutes les distances sont correctes. Nous présenterons les résultats ainsi (X% de nombre d'itinéraires corrects selon les directions ; X% du nombre d'itinéraires corrects selon les distances) pour chaque technique et uniquement pour la taille 5 car l'efficacité en taille 7 est quasi-nulle. La Mélodie a (81% ; 3%), la TTS-Lat a (72% ; 56%), le PG-TTS a (63% ; 38%) et la TTS a (69% ; 6%). Concernant la direction seule, pour un itinéraire de taille 5, l'efficacité est excellente avec une moyenne pour toutes les techniques de 71%, la mélodie en tête (81%). Notons en revanche, que l'efficacité observée pour la distance seule est très faible pour Mélodie (3%) et la TTS (6%), et ne dépasse pas 56% avec la techniques TTS-Lat. Ces résultats expliquent en partie la faible efficacité à transmettre des couples direction-distance, et donc à transmettre correctement un itinéraire pour les techniques Mélodie et TTS. Nous nous attendions à avoir ce type de résultat pour la Mélodie, car l'information de la distance est codée sur la durée d'un son et nous avons vu dans une expérience précédente que cela n'était pas très efficace (cf. chapitre 4, Efficacité des huit techniques). Mais il est surprenant de voir que la TTS présente

une faible efficacité à transmettre une distance avec de la parole. Pourtant les techniques PG-TTS et TTS-lat sont basées sur la même modalité pour transmettre la distance, mais présentent des résultats d'efficacité supérieurs (38% pour PG-TTS et 56% pour TTS-Lat).

**ÉTUDE DES RESULTATS EN TERMES DE ROBUSTESSE**

La robustesse représente le taux d'itinéraires restitués à l'identique entre le premier (R1) et le second (R2) rappel (conformes). Ce taux peut faire référence à des itinéraires qui ne sont pas totalement corrects, à condition qu'ils soient restitués de la même façon entre R1 et R2 (ceci correspond à une mémorisation robuste mais basée sur un encodage mnésique faux).

**Robustesse des itinéraires selon les couples, c'est-à-dire où tous les couples sont conformes, par technique par taille.** En considérant les couples direction-distance, la robustesse de la mémorisation avec Mélodie atteint 0,75 en taille 5. Elle est de 0,37 en taille 7. Ces résultats sont significativement supérieurs à ceux mesurés pour les autres techniques (Taille 5 : Friedman rank sum test :  $X^2= 11,7$ ,  $df = 3$ ,  $p<0,01$  ; Taille 7 :  $X^2= 12,9$ ,  $df = 3$ ,  $p<0,01$  ; voir Figure 5-13). En effet, les autres techniques ne dépassaient pas 0,19 en taille 5 et 0,07 en taille 7.

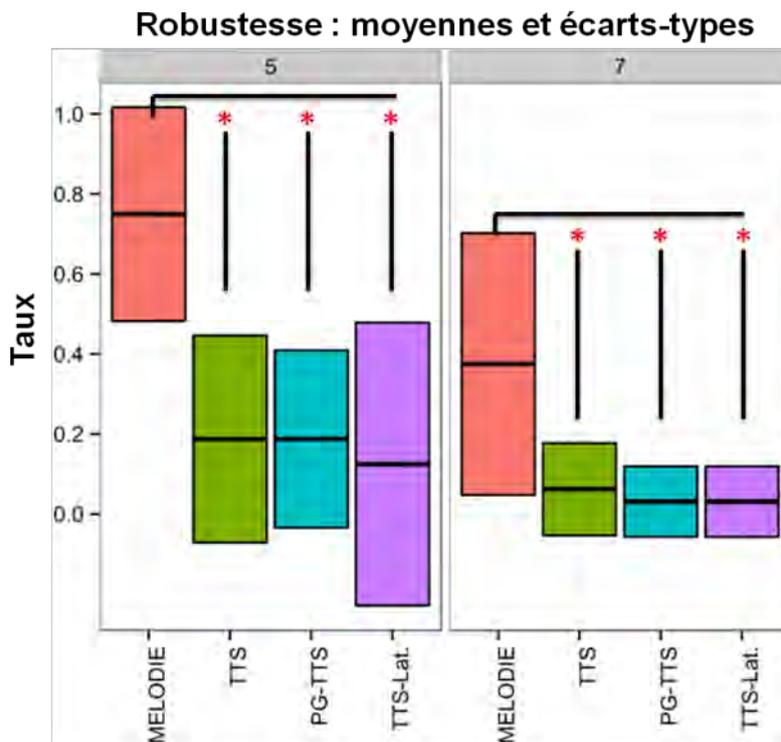


Figure 5-13 : Robustesse de la mémorisation selon la taille de l'itinéraire (5 segments à gauche et 7 à droite). (Wilcoxon signed rank test, \* :  $p<0,05$ ).

**Robustesse des itinéraires selon les directions seules et les distances seules, c'est-à-dire où toutes les directions sont conformes et toutes les distances sont conformes.** Nous

présenterons les résultats ainsi (X% de nombre d'itinéraires conformes selon les directions ; X% du nombre d'itinéraires conformes selon les distances) par technique par taille. La Mélodie présentait une robustesse de (100% ; 75%) pour la taille 5 et de (75% ; 66%) en taille 7. TTS-Lat présentait une robustesse de (78% ; 16%) pour la taille 5 et de (56% ; 31%) en taille 7. PG-TTS présentait une robustesse de (88% ; 34%) pour la taille 5 et de (75% ; 25%) en taille 7. Enfin, TTS présentait une robustesse de (72% ; 53%) pour la taille 5 et de (66% ; 40%) en taille 7. Concernant la direction seule, pour un itinéraire de taille 5 (resp. taille 7), la robustesse est excellente (resp. bonne) avec une moyenne pour toutes les techniques de 84% (resp. 68%), la mélodie en tête pour les deux tailles. Notons en revanche, que la robustesse observée pour la distance seule est différente selon les techniques et les tailles. Alors que la Mélodie présente des résultats que nous qualifierons d'homogènes les autres techniques multimodales présentent de grandes différences entre les directions et les distances. Notre technique témoin, TTS, montre des différences moins importantes mais avec des résultats inférieurs. Le fait que la Mélodie permette de mémoriser en chantant a sans doute favorisé la robustesse des directions et des distances et donc, des couples et des itinéraires. Nous nous attendions à avoir ce type de résultat pour la Mélodie mais, il est surprenant de voir que la modalité de la parole n'a pas favorisé la rétention d'informations.

#### **ÉTUDES QUALITATIVES**

**Satisfaction.** Les techniques préférées des participants Non-DV sont Mélodie et TTS-Lat (voir Figure 5-14 ; Friedman rank sum test  $X^2= 10,35$ ,  $df = 3$ ,  $p<0,02$ ).

**Charge cognitive.** Aucune différence significative n'a été établie concernant la charge cognitive estimée par le score obtenu avec le NASA-TLX (Friedman rank sum test :  $X^2= 4,95$ ,  $df = 3$ ,  $p>0,1$ ).

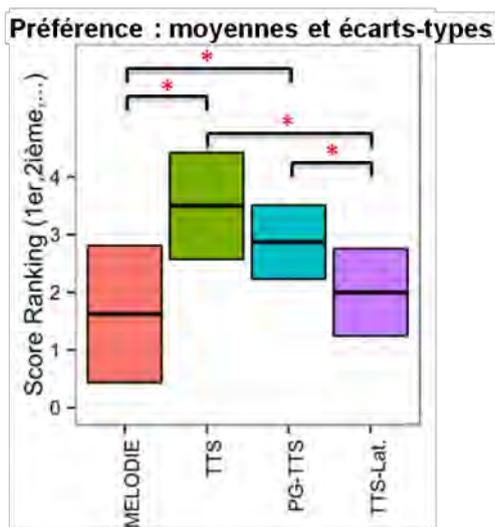


Figure 5-14 : Ordre moyen de préférence des techniques des sujets Non-DV ; Mélodie et TTS-Lat sont préférées à TTS et PG-TTS (Wilcoxon signed rank test ; \* : p<0,05)

### 5.8 Analyse des résultats préliminaires concernant les sujets DV

Nous avons complété cette première analyse par une analyse spécifiquement liée aux participants DV. Compte tenu de l'échantillon très réduit, cette analyse est plus succincte que la précédente et ne peut faire appel à des tests statistiques.

#### ÉTUDE DU NOMBRE D'ÉLÉMENTS CORRECTS LORS DE LA PREMIÈRE PHASE DE RESTITUTION

**Moyenne par taille.** Il ressort tout d'abord que lors de la première phase de restitution (R1) le nombre moyen d'éléments corrects pour un itinéraire, en fonction de la taille et toutes techniques confondues est en moyenne de 7,5 pour la taille 5 et de 7,8 pour la taille 7, ce qui était très similaire pour les non-DV : 7,6 en taille 5 et 7,8 en taille 7.

**Moyenne par technique et par taille d'itinéraire.** Pour les itinéraires de cinq segments, il apparaît que la technique la moins efficace obtient un score de 7,3 et la plus efficace un score de 8,0. Ces scores sont de 7,2 et 8,1, respectivement, pour des itinéraires de taille 7. Comme pour la population des non-DV, ce résultat est cohérent avec la limite de l'empan mnésique à  $7 \pm 2$  éléments (cf. Figure 5-15).

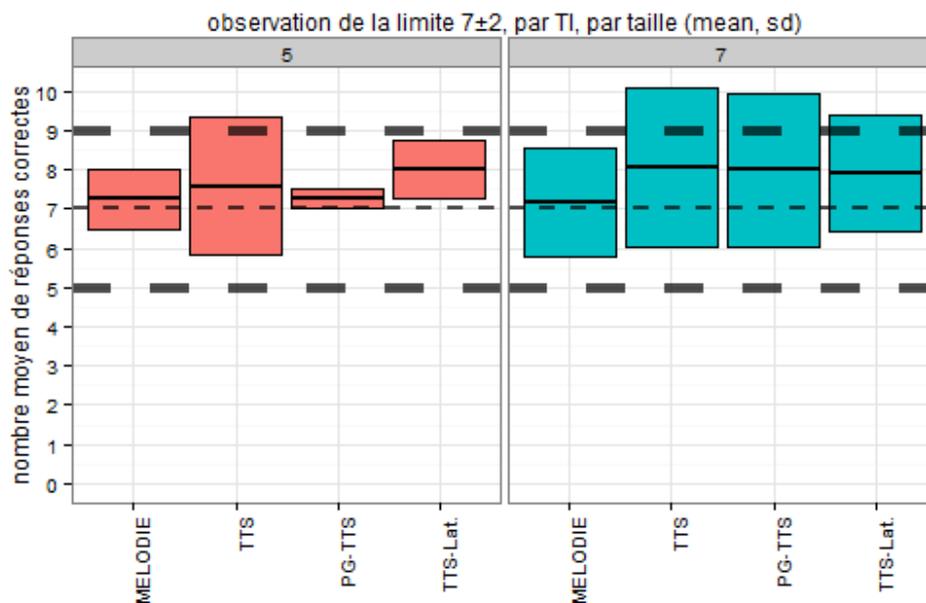


Figure 5-15 : Efficacité des techniques chez les DV mesurée lors du premier rappel, pour des itinéraires de 5 (gauche) ou 7 (droite) segments. Les lignes de tirets noirs, épais, illustrent la limite de l'empan mnésique à  $7 \pm 2$  éléments.

**COMPARAISON DE L'EFFICACITE OBSERVEE CHEZ LES DV ET NON-DV**

Nous avons comparé l'efficacité en ne considérant que les directions seules et que les distances seules selon les deux tailles d'itinéraires et les deux populations. Il apparait que les résultats des DV sont similaires à ceux des Non-DV (points et triangles se chevauchant cf. Figure 5-16).

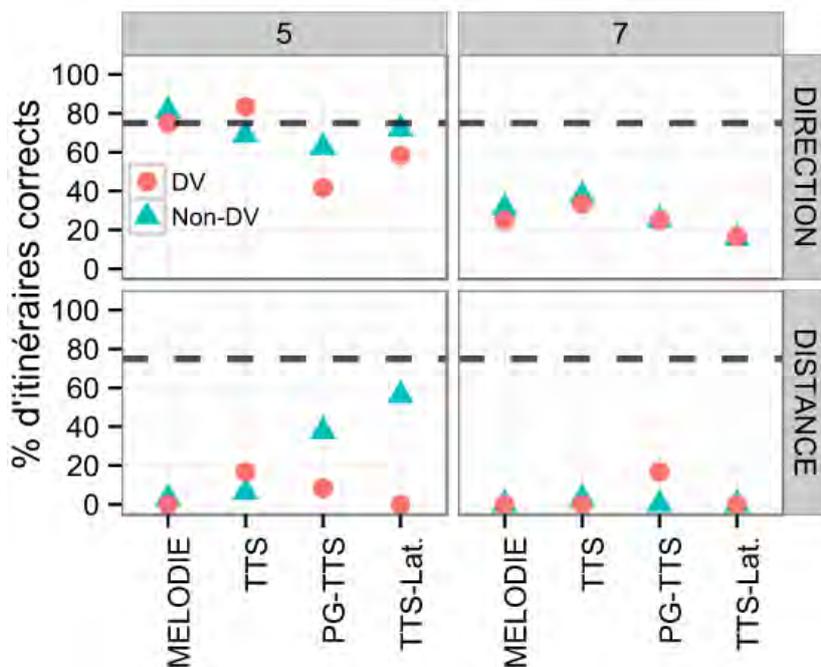


Figure 5-16 : Efficacité des techniques pour les deux populations. Le graphique illustre le pourcentage du nombre d'itinéraires corrects selon les directions (haut) et les distances (bas) observé lors de la première restitution. Les itinéraires de taille 5 sont illustrés à gauche, et ceux de taille 7 à droite. Les lignes de tirets indiquent un seuil d'efficacité à 75%.

**COMPARAISON DE LA ROBUSTESSE OBSERVEE CHEZ LES DV ET NON-DV**

Nous avons également comparé la robustesse en ne considérant que les directions seules et que les distances seules selon les deux tailles d'itinéraires et les deux populations. Il apparaît que la Mélodie facilite la mémorisation en taille 5, cf. Figure 5-17.

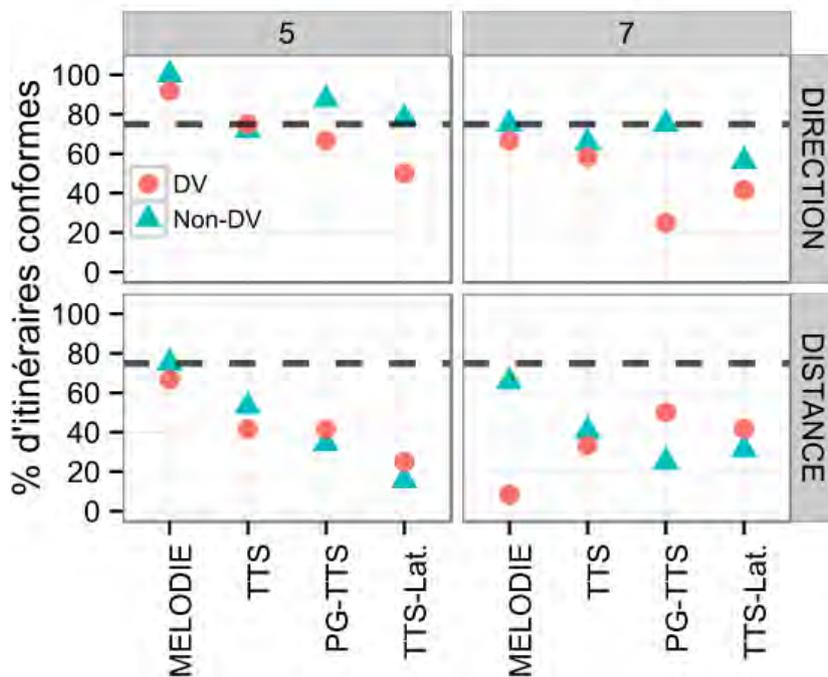


Figure 5-17 : Robustesse des techniques pour les deux populations. Le graphique illustre le pourcentage du nombre d'itinéraires conformes selon les directions (haut) et les distances (bas). Les itinéraires de taille 5 sont illustrés à gauche, et ceux de taille 7 à droite. Les lignes de tirets indiquent un seuil de robustesse à 75%.

**5.9 Analyse spécifique : mémorisation des segments selon leur position dans l'itinéraire**

Dans la Figure 5-18, nous pouvons observer que la technique TTS-Lat permet d'obtenir de bons résultats pour tous les segments en ce qui concerne les directions et les distances en taille 5. Pour la technique PG-TTS, nous pouvons voir que les distances intermédiaires sont moins bien retenues que les distances de début et de fin (pour la taille 5). Pour la technique contrôle monomodale (TTS), nous pouvons constater son inefficacité pour transmettre des couples (direction ; distance) car les utilisateurs ne retiennent que les directions. Enfin, la mélodie ne suit pas exactement le même schéma que les autres, et sa faible efficacité à transmettre correctement une distance la disqualifie comme technique optimisée efficace à transmettre des couples (direction ; distance). Remarquons que pour la taille 7, aucune technique n'est efficace.

Dans la Figure 5-18, nous pouvons aussi observer que certaines courbes dessinent un « V ». Dans la littérature, nous retrouvons le type de courbe en « V » pour des tâches de

mémorisation à court terme. Cette courbe en « V » s'explique en partie par les effets de primauté et de récence (Murdock, Bennet B. 1962), c'est-à-dire que les premières et dernières informations sont mieux retenues que les informations intermédiaires. Ces effets sont connus également sous le nom d' « effet de position sérielle » (Jahnke 1963).

Ces résultats, représentés sous la forme d'un graphique, illustrent l'efficacité de nos techniques pour mémoriser chaque direction et chaque distance, selon sa position dans l'itinéraire. Par exemple, pour la technique PG-TTS, en taille 5, nous observons que le pire ratio est de 0,59 : les utilisateurs restituaient correctement en moyenne 59% des réponses pour les distances du segment 3. L'efficacité de cette technique pour mémoriser un itinéraire complet serait donc plafonnée à 59%, même si les autres éléments étaient parfaitement mémorisés.

Idem pour les autres techniques en taille 5 : TTS-Lat serait plafonnée à 72% de réussite, 31% pour TTS, et 34% pour Mélodie.

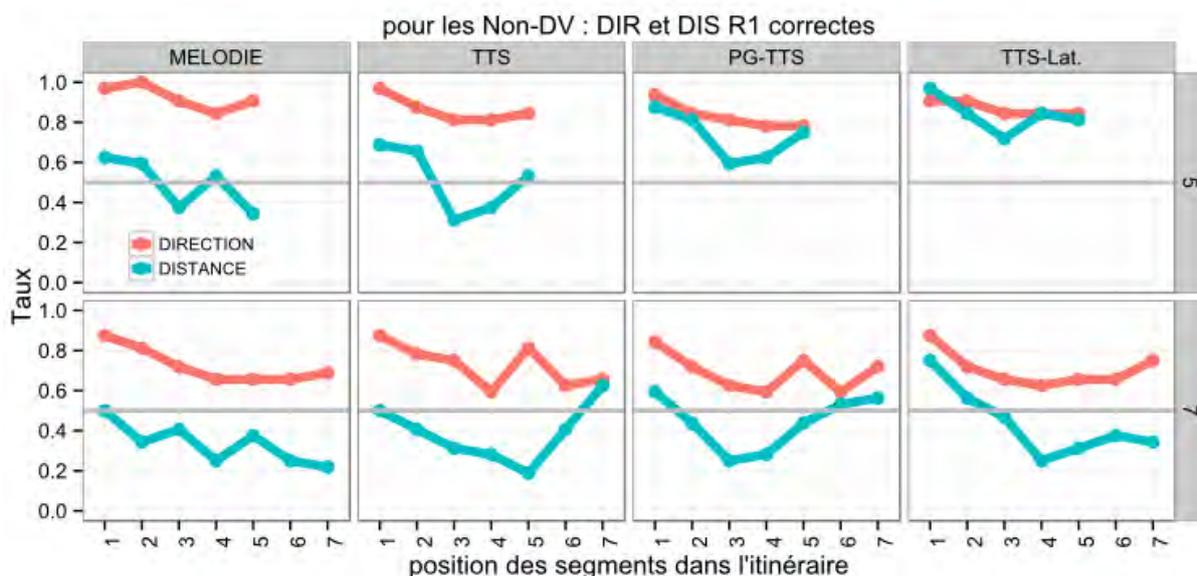


Figure 5-18 : Moyenne des items correctement restitués (R1) par les Non-DV, selon la position du segment dans l'itinéraire. Les graphiques illustrent les directions (rouge) et les distances (bleu) par technique et par taille d'itinéraire (5 en haut et 7 en bas). Seule TTS-Lat, pour la taille 5, résiste à l'effet de position. Rappelons que Mélodie est la seule technique utilisant le codage DURÉE pour la distance, les autres techniques utilisent la parole.

### 5.10 Discussion et conclusion de la deuxième étude

Nous avons conçu et évalué quatre techniques d'interaction non visuelles dont trois sont multimodales. Ces techniques ont été conçues afin d'évaluer leur impact sur la mémorisation d'une séquence d'informations représentative, par exemple, d'un itinéraire pour piétons.

Nous avons étudié les résultats d'un groupe de sujets Non-DV et un groupe de sujets DV. Il est intéressant de noter que la limite théorique de l'empan mnésique ( $7\pm 2$ ) aurait dû rendre difficile la mémorisation de tous les couples de valeurs pour des itinéraires courts (10 éléments), mais surtout pour des itinéraires longs (14 éléments). Néanmoins certaines techniques ont obtenu une efficacité correcte. Les techniques TTS-Lat et PG-TTS semblent notamment permettre d'optimiser la mémorisation d'une séquence de couples de valeurs.

Par ailleurs, Mélodie a montré, en taille 5, la meilleure efficacité pour les directions (0,81) et la pire pour les distances ( $<0,04$ ). Néanmoins sa robustesse pour ces mêmes itinéraires dépasse le taux de 0,75. Compte tenu des retours utilisateurs et du fait qu'elle est la technique préférée des participants, nous pensons toutefois que c'est la technique elle-même qui joue ce rôle en permettant de chanter l'information. Mélodie bénéficiait aussi d'un triple encodage de la direction : la redondance dans la multimodalité pourrait donc jouer un rôle dans la mémorisation des données. Une expérimentation complémentaire devra être conduite pour étudier ces deux facteurs.

Notre technique de référence, la TTS ne s'est pas particulièrement démarquée. Cette technique était toutefois la moins préférée (avec PG-TTS), la moins efficace (avec Mélodie) et n'était pas robuste.

Des améliorations sont requises afin d'accroître la robustesse de TTS-Lat et PG-TTS d'une part, et d'améliorer l'efficacité des trois techniques, notamment celle de Mélodie, d'autre part. Par exemple, en nous appuyant sur les résultats de notre première étude<sup>46</sup>, nous pourrions ajouter des références dans Mélodie, afin que les utilisateurs puissent mieux estimer les distances. Nous pourrions également concevoir et évaluer une version active de ces techniques car nous n'avons que très peu exploité la mémoire kinesthésique. De plus, nous avons vu que les techniques actives présentaient des résultats significativement supérieurs aux techniques passives dans des tâches d'estimation (cf. Facteur 3 : la catégorie Active/Passive (A/P)). Enfin, il se pourrait également que la mémorisation soit directement liée aux profils des utilisateurs (Ministère de l'éducation nationale - DGESCO 2012), donc dans notre contexte de transmission d'informations non visuelles, il serait utile de concevoir et évaluer des techniques basées sur l'audio et d'autres sur le tactile. Cela afin que les utilisateurs puissent avoir le choix de techniques équivalentes afin d'estimer et de mémoriser des itinéraires.

---

<sup>46</sup> Première étude : techniques d'interaction non visuelle pour estimer une valeur

Cette étude a également permis de mettre en évidence deux biais potentiels du protocole expérimental. En effet, il apparaît nécessaire d'adapter le mode de rappel en fonction de la technique d'interaction car Mélodie repose sur l'encodage d'une valeur par la durée de la note. Lors du rappel, nous ne savons pas si l'erreur de distance observée correspond à un problème de décodage (durée de la note vers la distance estimée) ou à un problème de mémorisation de la mélodie dans sa globalité.

Enfin, cette étude se limite à évaluer des techniques dans un cas d'utilisation stricte. Elle ne permet pas de la généraliser sur l'efficacité des techniques *in situ*. Nous avons également envisagé d'autres contextes basés sur l'utilisation de couples propriété-mesure, comme par exemple en géographie (région-superficie ou ville-habitant), pour lesquels de telles techniques pourraient s'avérer pertinentes, et dans lesquels, il conviendrait d'étudier la répliquabilité des résultats observés.

# **Chapitre 6**

## **Troisième étude :**

**assistance tactile à la  
localisation de cibles  
périphériques pour des  
personnes à vision tubulaire**



## 6 Troisième étude : assistance tactile à la localisation de cibles périphériques pour des personnes à vision tubulaire

*Note : cette étude a été réalisée en collaboration avec Damien Camors, doctorant co-encadré par Jean-Baptiste Durand (CerCo, Toulouse) et Christophe Jouffrais (IRIT, Toulouse).*

Dans les chapitres précédents, nous avons présenté des techniques non visuelles efficaces pour transmettre une valeur (étude 1, chapitre 4) avec une première application à la transmission d'un itinéraire (étude 2, chapitre 5). Ce chapitre présente notre troisième étude, dans laquelle nous avons conçu et évalué des techniques d'assistance à la localisation de cibles dans une portion d'espace.

Cette étude s'inscrit dans un contexte d'assistance à la recherche de cibles pour des personnes déficientes visuelles et plus spécifiquement avec une vision tubulaire. Dans ce contexte, nous avons conçu et comparé plusieurs techniques d'interaction, basées sur la modalité vibrotactile pour faciliter la recherche visuelle. En premier lieu, nous avons comparé quatre techniques permettant de transmettre une position à l'aide d'indices tactiles d'anticipation (cf. exp. 3.1), dans une tâche de localisation. Puis, nous avons évalué la technique la plus efficace et la plus satisfaisante dans une tâche de recherche visuelle avec des sujets voyants sur lesquels le champ visuel était artificiellement réduit (cf. exp. 3.2). Enfin, nous avons finalement validé ce résultat avec un patient atteint de glaucome et présentant un déficit visuel important (cf. exp. 3.3).

### **MISE EN CONTEXTE : QU'EST-CE QUE VISION TUBULAIRE ?**

C'est un état de la vision qui résulte de maladie dégénérative de l'œil. La rétinite pigmentaire et le glaucome sont des affections dégénératives de la rétine. Ces deux affections conduisent à une perte progressive du champ visuel et éventuellement à la cécité (Grover, Fishman, and Brown 1998). Ces affections, actuellement incurables, commencent par détériorer la vision périphérique et causent ce que l'on appelle la vision « tubulaire » ou « en tunnel ». Pour comprendre ce qu'est une vision tubulaire, il faut imaginer regarder à travers une paille, le champ visuel est alors réduit à quelques degrés d'angle (cf. Figure 6-1).

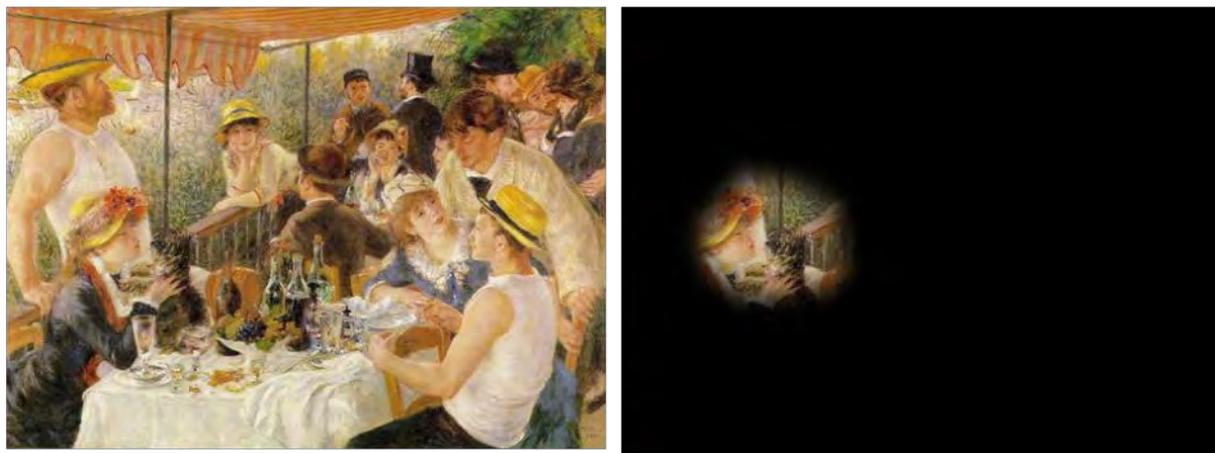


Figure 6-1 : Illustration d'une vue tubulaire : à gauche, « Le déjeuner des canotiers » par Pierre-Auguste Renoir en 1880 ; à droite ce même tableau vu avec une vision tubulaire schématisée.

Il peut y avoir d'autres symptômes associés comme une gêne visuelle importante dans la pénombre (héméralopie) ou une adaptation plus lente aux changements de luminosité. La vision tubulaire a d'importantes conséquences sur la recherche visuelle (Lowe and Drasdo 1992; Smith, Crabb, and Garway-Heath 2011), et par conséquent sur l'autonomie des patients (Herse 2005). Par exemple des tâches usuelles comme trouver un produit sur une étagère ou rechercher un ami dans une foule, deviennent très difficiles et chronophages quand la complexité visuelle de l'environnement s'accroît (Wolfe 1998) et que le champ visuel de l'utilisateur est dégradé (Kuyk, Liu, and Fuhr 2005).

### 6.1 État de l'art sur les technologies d'assistance à la recherche visuelle

De nombreuses technologies ont été conçues pour améliorer la recherche visuelle et la perception de l'espace chez des personnes avec un champ visuel limité. Certaines sont basées sur les capacités visuelles résiduelles et donc sur la modalité sensorielle visuelle, mais d'autres sont basées sur les modalités sensorielles auditives et tactiles.

#### *TECHNIQUES D'ASSISTANCE BASEES SUR LA MODALITE SENSORIELLE VISUELLE*

Des dispositifs spécifiques à la vision tubulaire ont été élaborés afin d'assister la localisation de cibles. Certains dispositifs s'appuient sur le champ visuel résiduel et reposent sur la projection, dans ce champ résiduel, de l'information présente dans la partie aveugle. Une première possibilité consiste à compresser l'information visuelle avec une « minification » optique (Kozlowski and Jalkh 1985; Krefman 1981). D'autres dispositifs utilisent un prisme qui dévie l'information visuelle dans le champ visuel résiduel (Woods *et al.* 2010). Dans plusieurs dispositifs basés sur la réalité augmentée, une information visuelle spécifique (par

exemple la détection de contours) peut être superposée dans la partie intacte du champ visuel (Vargas-Martin and Peli 2002) (cf. Figure 6-2).



**Figure 6-2 : illustration de la technique de superposition de (Vargas-Martin and Peli 2002) : un utilisateur regarde un visage dans une salle et celle-ci est superposée dans le champs de vision résiduel de l'utilisateur avec de la réalité augmentée.**

Dans tous ces cas, le champ visuel résiduel est surchargé avec plus d'informations visuelles, ce qui perturbe l'exploration visuelle, et déforme la relation spatiale entre les objets (Fox 2000). Tous ces inconvénients peuvent avoir un impact important sur la recherche visuelle (Brooks, Rasmussen, and Hollingworth 2010).

Dans le chapitre 2<sup>47</sup> de ce manuscrit, nous avons vu de nombreuses techniques pour transmettre une ou des informations grâce à des techniques d'interaction non visuelle basées, notamment, sur les modalités sensorielles auditives et tactiles. Dans ce contexte d'assistance à une recherche visuelle, nous avons sélectionné des exemples de techniques de transmission d'information.

#### **TECHNIQUES DE TRANSMISSION D'INFORMATIONS ET SUBSTITUTION SENSORIELLE**

Les systèmes de substitution sensorielle<sup>48</sup> convertissent toute une scène visuelle en signaux auditifs (ex. « the VOICE », cf. Figure 6-3). Ils peuvent fournir à l'utilisateur suffisamment d'informations pour une tâche de locomotion, de localisation ou de pointage (Auvray *et al.* 2007; Meijer 1992). Cependant, même une tâche simple de pointage requiert beaucoup de temps et d'attention. De plus, les sons naturels de l'environnement, ont un rôle important dans de nombreuses tâches de la vie quotidienne. Les personnes déficientes visuelles sont donc réticentes à encombrer leur canal auditif avec des informations additionnelles encombrantes et non-spécifiques à la tâche en cours.

<sup>47</sup> Etat de l'art sur la transmission d'information basée sur les modalités tactiles et/ou auditives

<sup>48</sup> Substitution sensorielle : consiste à remplacer une modalité par une autre. Par exemple, remplacer une image (vue) par motif en relief (tactile).

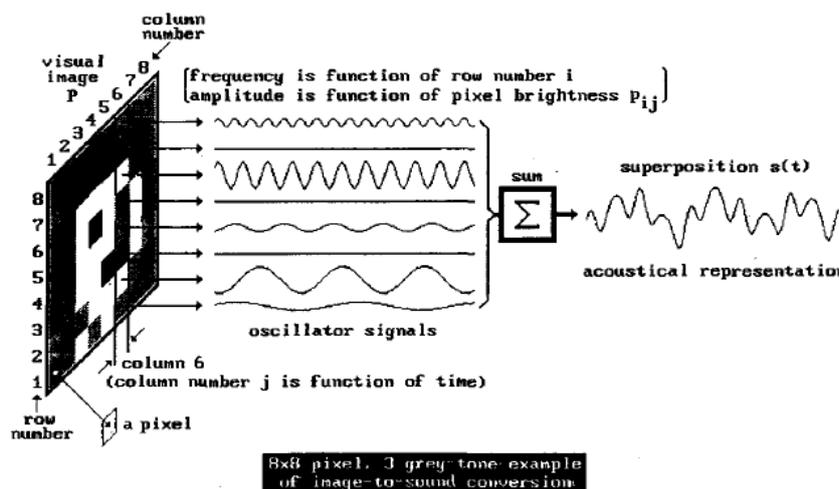


Figure 6-3 : illustration de l'encodage d'une image en signaux auditifs par «The vOICE», issue de (Auvray *et al.* 2007) ; L'encodage est basé sur trois propriétés du son, liées à trois propriétés d'une image : fréquence, durée, intensité (resp. ligne de pixels, colonne de pixel, nuance de gris). Les utilisateurs devaient « estimer » une image avec du son.

La substitution sensorielle a aussi été proposée en utilisant la modalité tactile. Dans ces systèmes, le principe consiste à transformer l'image provenant d'une caméra en un ensemble de signaux tactiles appliqués sur l'abdomen ou la langue (Bach-Y-Rita *et al.* 1969; Paul Bach-y-Rita *et al.* 1998) (voir Figure 6-4). Plusieurs expériences réalisées en condition de laboratoire montrent qu'il est, par exemple, possible de reconnaître des formes géométriques.

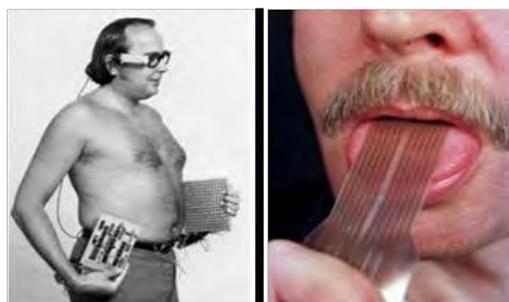


Figure 6-4 : illustration des prototypes issues de (Bach-Y-Rita *et al.* 1969) (gauche), (P Bach-y-Rita *et al.* 1998)(droite).

Cependant, après quarante ans de recherches sur la substitution sensorielle, aucun dispositif n'est utilisable dans des conditions écologiques (Lenay *et al.* 2003). De plus, ils ne sont pas utilisables pour localiser un objet spécifique dans des scènes naturelles pouvant être surchargées.

#### TECHNIQUES DE TRANSMISSION D'INFORMATIONS ET GUIDAGE SENSORIEL

Plutôt que de substituer une scène visuelle dans sa totalité, les dispositifs d'assistance peuvent être utilisés pour localiser précisément une cible ou un point d'intérêt dans la scène

(Brewster and Brown 2004). Dans le cas d'un utilisateur avec une vision tubulaire, il est possible de guider le regard ou la main vers une région spécifique dans l'environnement. Par exemple, il est possible de faciliter le pointage vers une cible grâce à des sons binauraux en fournissant à l'utilisateur l'azimut et l'élévation de la cible (Dramas *et al.* 2008; Katz *et al.* 2012).

Cependant, comme mentionné précédemment, les informations auditives additionnelles interfèrent avec les sons de l'environnement, ce qui réduit leur utilisabilité et peut masquer un danger potentiel (Hoggan *et al.* 2009). Dans cette étude, nous avons conçu un dispositif tactile, privé et peu intrusif, permettant de localiser des cibles spécifiques dans l'espace.

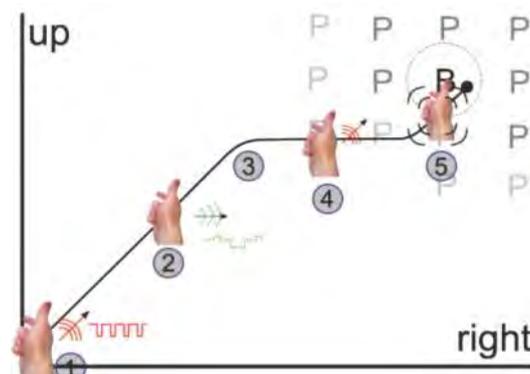
Dans les sections suivantes, nous faisons une synthèse des dispositifs tactiles permettant de localiser des cibles dans l'espace, puis nous décrivons nos choix de conception, ainsi que les différentes expériences que nous avons menées dans le but d'améliorer la localisation de cibles périphériques.

Toutes ces expériences ont été réalisées avec un simulateur de vision tubulaire dans lequel le champ visuel a été restreint artificiellement. Dans la dernière expérience, le dispositif a été testé avec un sujet déficient visuel atteint d'un glaucome.

## 6.2 Dispositifs tactiles pour la perception spatiale

Les stimulations tactiles ne sont pas en compétition avec le traitement de stimuli auditifs ou visuels écologiques. De plus, elles permettent de réaliser des discriminations spatiales et temporelles avec des performances excellentes (Geldard 1957). Les dispositifs tactiles ont été utilisés dans un grand nombre d'applications concernant le guidage, la navigation, la notification, les alertes, etc. Par exemple, Boll et ses collaborateurs ont utilisé un dispositif tactile sur le poignet ou le torse dans le but d'assister la navigation et le guidage (Henze, Heuten, and Boll 2006; Heuten *et al.* 2008). Les dispositifs tactiles peuvent aussi servir comme détecteur d'obstacles afin de prévenir toute collision pendant un déplacement (Cassinelli, Reynolds, and Ishikawa 2006). Des travaux antérieurs ont également porté sur la localisation d'une cible spécifique dans l'environnement. Dans ce domaine, deux types d'indices ont été décrits par Lehtinen *et al.* (Lehtinen *et al.* 2012) : les indices que nous appellerons de *guidage* (« guiding » cues) et d'*anticipation* (« preview » cues). Les indices de guidage fournissent une information continue de l'erreur entre la cible et la partie mobile du corps, généralement la main. Les indices de guidage peuvent inciter l'utilisateur à porter

attention à une position précise dans l'espace, ce qui est particulièrement utile quand la cible est cachée dans une scène très fortement encombrée (Fiannaca, Morelli, and Folmer 2013; Folmer and Morelli 2012; Lehtinen *et al.* 2012; Oron-Gilad *et al.* 2007). Par exemple, (Lehtinen *et al.* 2012) ont placé deux vibreurs dans la paume et deux sur le dos de la main des utilisateurs (cf. Figure 6-5). La main des utilisateurs était alors continuellement guidée vers une lettre B parmi un ensemble de lettres P distractives. Avec le guidage tactile, ils ont observé une réduction significative du temps nécessaire pour trouver la lettre B. Cependant, les sujets admettaient « se concentrer d'abord sur le signal tactile » puis « sauter avec le regard vers la zone indiquée ». Ce faisant, les sujets ont ignoré le guidage tactile continu. Ils ont plutôt utilisé le signal tactile comme un indice d'anticipation avant de commencer la recherche visuelle dans une région déterminée de l'espace. De plus, ils ont mentionné ne pas être convaincus que le guidage soit utile après avoir commencé la recherche. Par conséquent, il semblerait que les indices de guidage soient moins pertinents que les indices d'anticipation. Les indices d'anticipation indiquent la direction ou la partie de l'espace à atteindre avant qu'une quelconque partie du corps ne bouge. Les indices d'anticipation n'accroissent pas la charge cognitive, notamment pendant la phase d'exploration. Ceci est particulièrement important durant une recherche visuelle complexe, et d'autant plus, pour les personnes avec une vision tubulaire.



**Figure 6-5 : illustration issue de (Lehtinen *et al.* 2012). La main des utilisateurs est continuellement guidée vers une lettre B parmi un ensemble de lettres P distractives avec deux messages : l'un signifiant « UP », pour signifier à l'utilisateur qu'il devait monter sa main et, l'autre « RIGHT » pour qu'il décale sa main vers la droite.**

Les indices d'anticipation ont déjà été utilisés dans une tâche de recherche visuelle (C. M. Jones *et al.* 2008; Lindeman *et al.* 2003; Ngo and Spence 2010). Les expériences ont montré une réduction significative du temps de recherche. Malheureusement, ils ont été utilisés pour fournir des informations de direction imprécises (par ex. « dans le coin en haut à gauche »), et de surcroît dans un espace non surchargé. Ainsi, la question demeure : les

indices d'anticipation tactiles peuvent-ils assister des personnes déficientes visuelles, et plus particulièrement, avec une vision tubulaire, afin de trouver une cible spécifique dans l'environnement ? Dans ce cas, les indices d'anticipation doivent indiquer une portion précise de l'espace, suffisamment proche de la cible d'intérêt pour que celle-ci se retrouve directement dans le champ visuel résiduel.

Dans cette étude, nous avons conçu un dispositif d'assistance tactile que nous avons mis en œuvre dans une série de trois expérimentations :

Expérience 3.1 : Comparer quatre encodages différents d'indices tactiles.

Expérience 3.2 : Évaluer le meilleur encodage en conditions simulées.

Expérience 3.3 : Valider l'encodage dans un cas d'étude avec un utilisateur à vision tubulaire.

A la fin de cette étude, nous discuterons des résultats et présenterons les perspectives envisagées pour améliorer la technique et pour l'évaluer *in-situ*.

### **6.3 Expérience 3.1 : Conception et comparaison d'indices tactiles permettant de localiser une cible périphérique**

Dans cette première expérience, nous avons conçu, évalué et comparé quatre techniques d'interaction différentes dont la modalité principale était le vibrotactile. Ainsi, nous avons créé quatre encodages différents d'indices tactiles, selon deux systèmes de coordonnées et deux types de signaux.

#### **6.3.1 Conception des indices tactiles**

Dans cette première expérience, nous avons conçu et comparé quatre codes tactiles permettant de localiser une cible dans l'environnement. La conception du dispositif tactile permettant d'accélérer une tâche de recherche visuelle est détaillée dans la seconde expérience.

#### **REFERENTIELS ET SYSTEMES DE COORDONNEES**

Dans un dispositif d'assistance à la localisation de cibles visuelles, les caméras sont indispensables. Elles peuvent être facilement montées sur des lunettes, et peuvent être utilisées pour localiser des cibles spécifiques dans une scène visuelle. Le système peut aussi intégrer un accéléromètre pour compenser une faible résolution temporelle de la caméra et procurer la position des cibles malgré les déplacements de la tête (voir (Katz *et al.* 2012) par exemple).

Le plus logique des référentiels, pour guider le regard vers la cible, est celui des yeux avec comme origine, la fovéa. En supposant que la cible ait été localisée par la caméra attachée à la tête, et que la position des yeux dans l'orbite soit connue grâce à un oculomètre portable, l'indice tactile peut alors directement indiquer l'erreur entre la fovéa et la cible. Cependant, les oculomètres portables sont chers et doivent être fréquemment calibrés. Comme les yeux et la tête sont synchronisés pendant les mouvements du regard (Nakashima and Shioiri 2014), la tête peut être un référentiel plus pertinent pour assister la recherche visuelle. En effet, cela repose uniquement sur la présence de la caméra montée sur les lunettes. Les indices de guidage peuvent alors directement indiquer l'erreur entre la position de la tête et la position de la cible dans l'image sans aucune transformation de coordonnées.

La localisation d'une cible dans l'espace, dans un repère égo-centré dont l'origine est la tête, repose sur trois coordonnées, à savoir l'azimut (plan horizontal), l'élévation (plan vertical) et la profondeur (ou distance). Cependant la profondeur entre l'utilisateur et la cible n'est pas obligatoire pour la recherche visuelle. En effet, quand la cible est localisée dans le plan de l'image de la caméra, seules deux coordonnées (azimut et élévation) suffisent pour amener le champ visuel en face de la cible. Dans le but de transmettre ces deux coordonnées, nous avons évalué deux systèmes de coordonnées (SC) différents, cartésien et polaire, ancrés sur la tête.

Dans le système cartésien, la position de la cible était décomposée en deux indices tactiles successifs dépendant de l'orientation de la tête. L'azimut (axe horizontal) était présenté en premier, suivi de l'élévation (axe vertical). Dans le système polaire, l'azimut et l'élévation de la cible étaient convertis en deux stimulations successives représentant une inclinaison et une distance radiale (cf. Figure 6-6, ci-après).

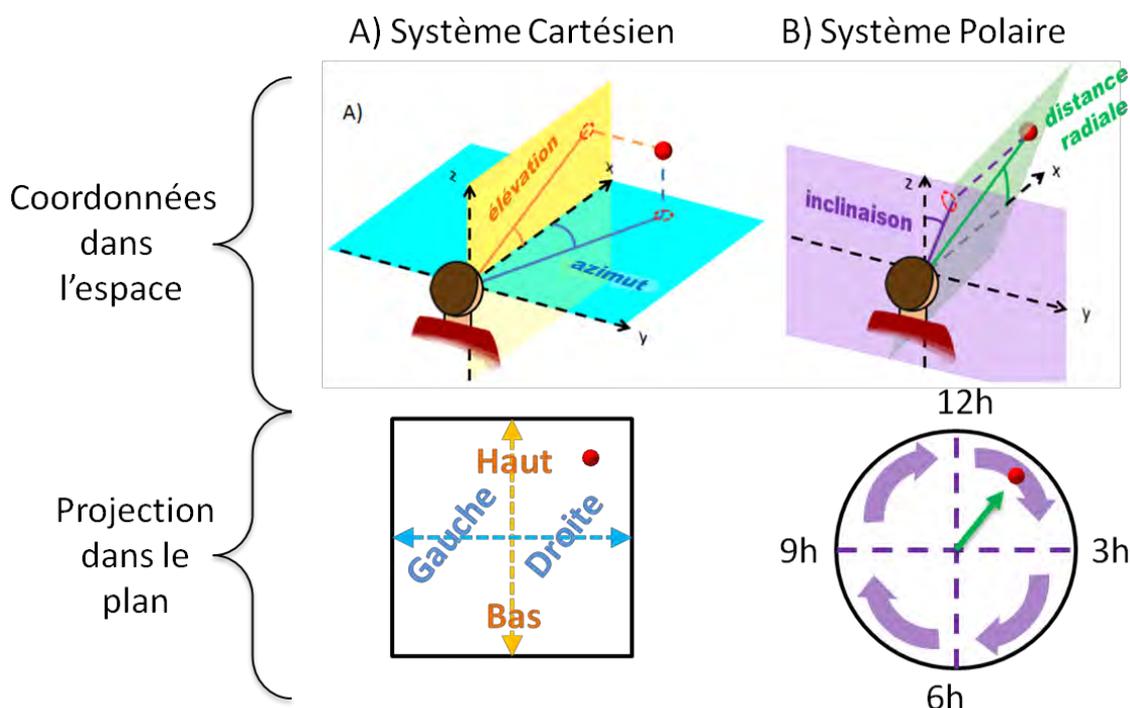


Figure 6-6 : illustration des deux systèmes de coordonnées indiquant une cible rouge : à gauche (A), le système dit cartésien, l'azimut représente la coordonnée 1 (bleu) et l'élévation représente la coordonnée 2 (orange) ; à droite (B), le système dit polaire, l'inclinaison représente la coordonnée (violet) et la distance radiale représente la coordonnée 2 (vert).

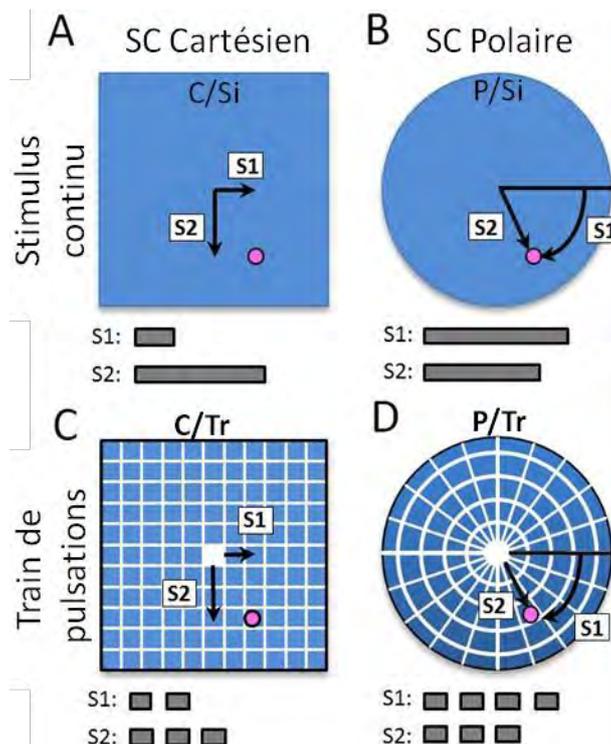
#### ENCODAGE DES COORDONNEES AVEC LA MODALITE VIBROTACTILE

Comme mentionné précédemment, les indices d'anticipation ne guident pas progressivement vers une cible. Après une brève stimulation, l'utilisateur doit être capable d'estimer la position de la cible avec une bonne précision. Si la stimulation n'est pas suffisamment explicite, l'utilisateur devra redemander les indices, ce qui n'est pas utilisable. Par conséquent, l'indice tactile doit être aussi bref que possible et permettre d'accélérer la recherche visuelle. Nous avons évalué deux codes tactiles différents (cf. Figure 6-7, ci-après). Le premier est basé sur une seule vibration de durée variable correspondant à la grandeur de la coordonnée transmise. La durée maximale est de 1,5 secondes. Le choix de cette durée était basé sur des pré-tests, non-détaillés ici, qui indiquaient que la discrimination était possible avec cette durée, et qu'une durée plus longue semblait gênante pour les utilisateurs. La durée maximale de 1,5 secondes correspondait à un azimut ou une élévation maximums de 45°.

Dans le second code, une coordonnée est définie comme un train (une séquence) de courtes pulsations de 200 ms séparées entre elles par 100 ms. Le nombre de pulsations dans le train correspond à la zone cible dans l'espace frontal, avec chaque pulsation indiquant

une zone de 9°. Le train comporte cinq pulsations au maximum, correspondant au même maximum de 45°.

Ces deux codes tactiles ont été évalués dans les deux systèmes de coordonnées décrits précédemment. Par conséquent, nous avons évalué quatre techniques notées : Singlestimulus dans un SC Cartésien (C/Si), TRain de pulsations dans un SC Cartésien (C/Tr), Singlestimulus dans un SC Polaire (P/Si) et TRain de pulsations dans un SC Polaire (P/Tr) (voir Figure 6-7).



**Figure 6-7 : Codage tactile d'une même cible (point rose) avec quatre techniques différentes. Deux codes tactiles ont été utilisés : le Singlestimulus (A, B) et le train de pulsations (C, D). Les coordonnées de la cible sont transmises dans les SC cartésien (A, C) ou polaire (B, D). Les deux coordonnées sont rendues par deux stimulations successives (S1, S2) représentées avec des barres grises. Dans cette expérience, la position de la cible était choisie de façon pseudo-aléatoire dans toute la zone bleue.**

#### **POSITION DU DISPOSITIF TACTILE AUTOUR DU POIGNET**

Les dispositifs tactiles peuvent être placés sur différentes parties du corps humain. Afin d'assister la recherche visuelle, une solution serait de placer un dispositif tactile autour des yeux. La correspondance entre la stimulation et la direction souhaitée regard serait probablement très intuitive. Cependant, la sensibilité tactile sur le visage est très importante. Le dispositif tactile pourrait rapidement provoquer une irritation. De plus, tous les sujets interrogés étaient réticents à porter un dispositif sur le visage.

Les indices directionnels sont le plus souvent transmis par des dispositifs tactiles attachés à une ceinture, sur le dos ou sur les mains. La ceinture est une solution à la fois élégante et

non-intrusive pour transmettre l'azimut (Van Erp *et al.* 2005; Heuten *et al.* 2008), mais pas pour l'élévation. Le dos et le torse présentent une importante surface tactile, ce qui permet d'y placer un nombre important de vibreurs. Les premières études ont utilisé des vibreurs placés sur une chaise dans le but de transmettre des indices d'anticipation sur le dos (C. M. Jones *et al.* 2008; Lindeman *et al.* 2003). Ce dispositif n'était pas mobile ce qui en limitait l'usage. Plus récemment, les vibreurs ont été placés dans des gilets (Van Erp 2005), ce qui peut freiner l'acceptabilité du dispositif d'assistance.

La main est souvent utilisée (Fiannaca *et al.* 2013; Folmer and Morelli 2012; Lehtinen *et al.* 2012; Oron-Gilad *et al.* 2007), ce qui est évidemment un avantage quand un mouvement de celle-ci est requis. En effet, cela évite de transformer les coordonnées entre le capteur et l'effecteur, et par conséquent facilite les tâches de pointage. Cependant, les dispositifs qui sont placés dans la paume de la main posent des problèmes de préhension et de manipulation. Le dos de la main ou le poignet sont de bonnes alternatives. Le poignet peut offrir une discrimination efficace des stimuli (Chen *et al.* 2008; Matscheko *et al.* 2010). Contrairement au dos et à la hanche, le poignet est impliqué dans le processus visuo-moteur permettant la coordination œil-main (Henriques *et al.* 2002). Un autre argument pour fixer le dispositif tactile au poignet est, l'arrivée des « Smartwatch » dans la communauté des déficients visuels. Ces dispositifs possèdent des vibreurs pouvant être activés à distance (par exemple en Bluetooth). Si nécessaire, des vibreurs additionnels peuvent être attachés au bracelet de la montre.

### 6.3.2 Présentation des techniques avec notre cadre analytique

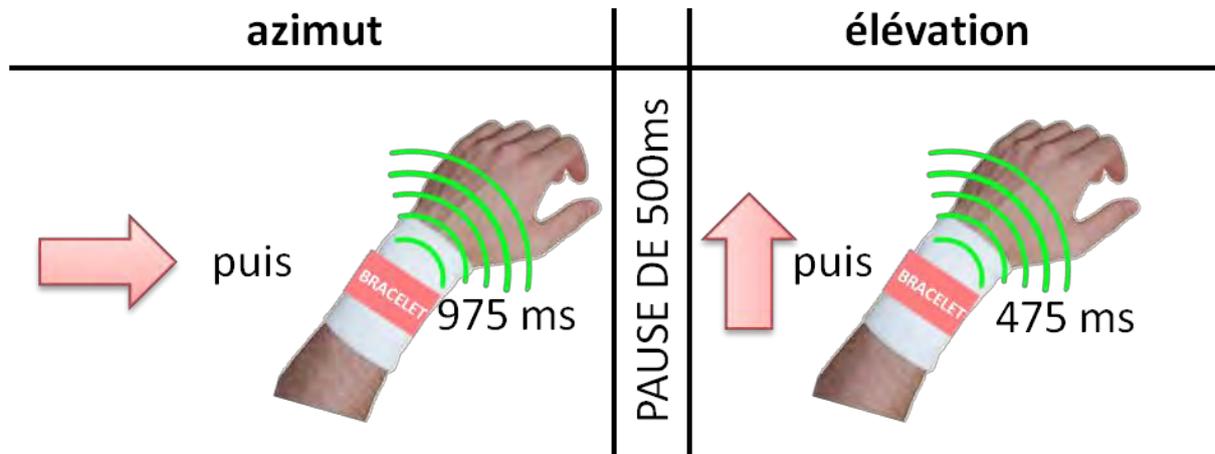
Il est important de noter que la tâche est de localiser une cible. Le signal contenant l'information est toujours envoyé au début de la tâche, une seule et unique fois. Donc d'après nos définitions des types de signal (cf. Différents types de signal), celui-ci est discret. Cela est cohérent avec les définitions d'indice de guidage (continu pendant la tâche) et d'indice d'anticipation (présent uniquement au début de la tâche). Les sections ci-dessous, présentent les quatre techniques d'interaction pour la transmission d'une paire de coordonnées.

#### 6.3.2.1 Singlestimulus dans un SC Cartésien (C/Si)

**Principe** : est similaire au pointage d'une cible avec une tourelle dont la base est posée sur le pont d'un navire. La visée de la tourelle est le champ de vision de l'utilisateur qui avait pour tâche de pointer une position à partir de deux angles : azimut et élévation dont leur

valeur était transmise par des vibrations et leur sens était indiqué par des flèches rouges (cf. Figure 6-8).

**Illustration :**



**Figure 6-8 : illustration de la technique C/Si. Exemple illustré : la cible est 29,25° à droite et 14,25° en haut. La succession des stimuli était : flèche pointant à droite sur l'écran, puis une vibration de 975 ms, puis 500 ms plus tard, flèche pointant en haut, puis une vibration de 475 ms.**

**Cadre analytique :** technique d'interaction multimodale passive. Le signal en sortie du système vers l'utilisateur s'effectue en alternance avec deux effecteurs, l'un capable de transmettre un stimulus visuel de manière discrète dont l'information est codée le type de codage LANGAGE-symbolique (4 images « flèches » == 4 sens) et l'autre capable de transmettre un signal tactile discret dont l'information est codée avec une seule vibration d'une certaine DURÉE (1 500 ms == 45°). Les informations des couples sont transmises en alternance et les couples sont transmis en séquence (cf. Figure 6-9).

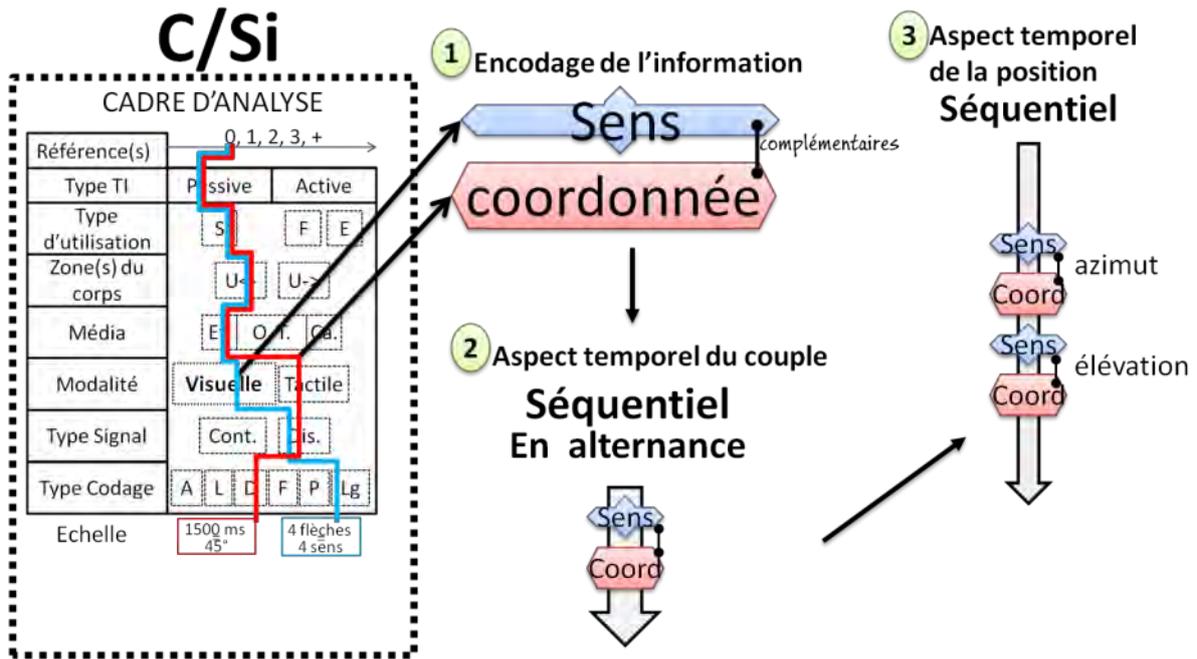


Figure 6-9 : illustration de l'analyse de la technique Singlestimulus dans un SC Cartésien (C/Si).

**Matériel** : casque de réalité virtuelle [NVisor® SX-60 HMD, (NVIS Inc., USA)] pour afficher la scène virtuelle et les flèches rouges. L'affichage est géré par le moteur 3D [Irrlicht v1.8]. Les deux vibrations étaient générées à partir d'un seul moteur vibrant [VPM2] et leur durée était contrôlée par une carte [Arduino® ATmega1280]. Enfin, toute l'expérience était gérée grâce à un programme en C#. (cf. la section Dispositif expérimental et stimuli).

### 6.3.2.2 Train de pulsations dans un SC Cartésien (C/Tr)

**Principe** : est similaire au pointage d'une cible avec une tourelle dont la base est posée sur le pont d'un navire. La visée de la tourelle est le champ de vision de l'utilisateur qui avait pour tâche de pointer une position à partir de deux angles : azimut et élévation dont leur valeur était transmise par des trains de vibrations et leur sens était indiqué par des flèches rouges (cf. Figure 6-10).

Illustration :

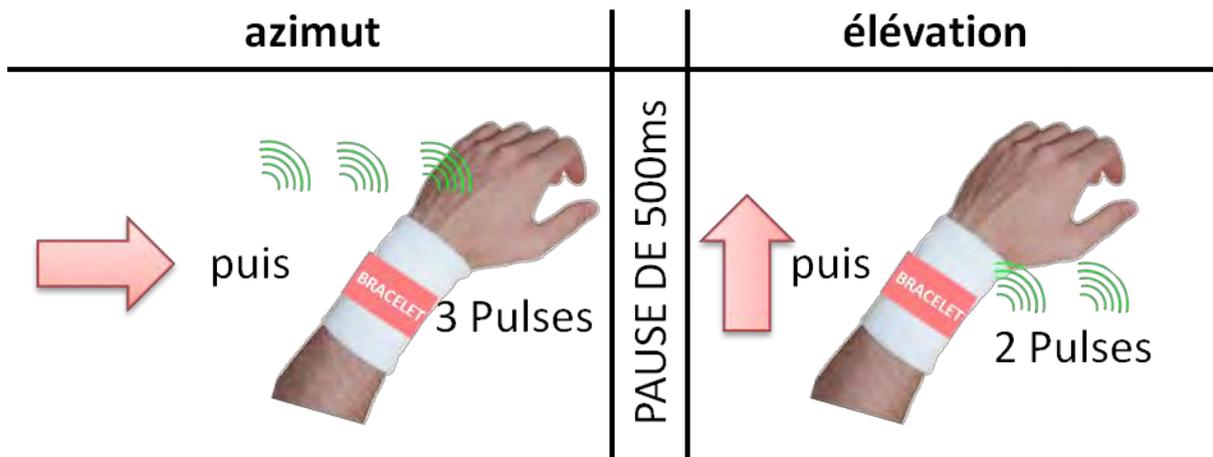


Figure 6-10 : illustration de la technique C/Tr. Exemple illustré : la cible est 27° à droite et 18° en haut. La succession des stimuli était : flèche pointant à droite sur l'écran, puis un train de trois pulsations, puis 500 ms plus tard, flèche pointant en haut, puis un train de deux pulsations.

**Cadre analytique** : technique d'interaction multimodale passive. Le signal en sortie du système vers l'utilisateur s'effectue en alternance avec deux effecteurs, l'un capable de transmettre un stimulus visuel de manière discrète dont l'information est codée avec le type de codage LANGAGE-symbolique (4 images « flèches » == 4 sens) et l'autre capable de transmettre un stimulus tactile discret dont l'information est codée avec le type de codage PULSE (succession de 5 vibrations == 45°). Les informations des couples sont transmises en alternance et les couples sont transmis en séquence (cf. Figure 6-11).

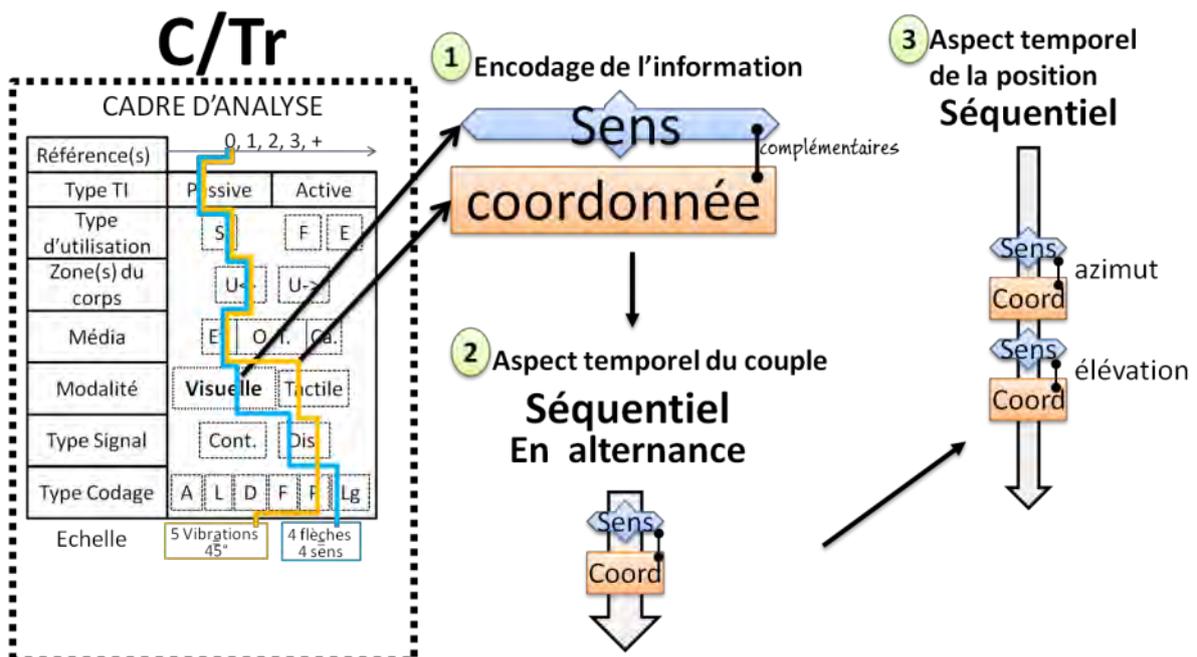


Figure 6-11 : illustration de l'analyse de la technique Train de pulsations dans un SC Cartésien (C/Tr).

**Matériel** : casque de réalité virtuelle [NVisor® SX-60 HMD, (NVIS Inc., USA)] pour afficher la scène virtuelle et les flèches rouges. L’affichage est géré par le moteur 3D [Irrlicht v1.8]. Les trains de pulsations étaient générés à partir d’un seul moteur vibrant [VPM2] et leur nombre de vibrations était contrôlé par une carte [Arduino® ATmega1280]. Enfin, toute l’expérience était gérée grâce à un programme en C#. (cf. la section Dispositif expérimental et stimuli).

### 6.3.2.3 Singlestimulus dans un SC Polaire (P/Si)

**Principe** : est similaire au pointage d’une cible avec une tourelle dont la base est posée sur la face avant d’un cube. La visée de la tourelle est le champ de vision de l’utilisateur qui avait pour tâche de pointer une position à partir de deux angles : inclinaison et distance radiale. La valeur de l’inclinaison était transmise par une vibration d’une certaine durée et le départ de l’inclinaison était transmis par l’une des quatre barres rouges ; puis, tout en gardant l’inclinaison, l’utilisateur pointe la cible avec la distance radiale dont la valeur était transmise par une vibration d’une certaine durée. Notons qu’un indice visuel « cercle rouge » indiquait à l’utilisateur qu’il allait recevoir la seconde coordonnée (cf. Figure 6-12).

**Illustration** :

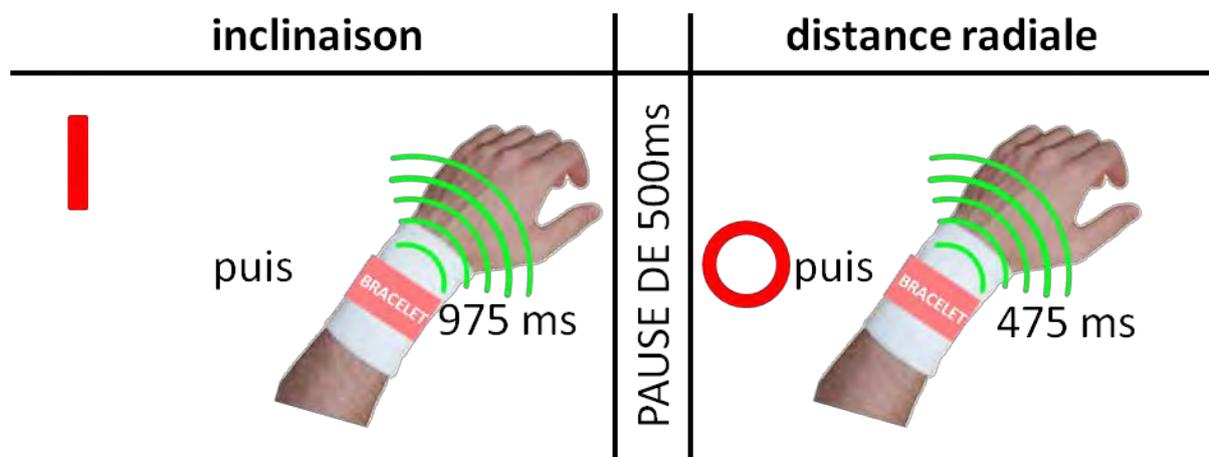


Figure 6-12 : illustration de la technique P/Si. Exemple illustré : la cible est 29,25° à droite et 14,25° en haut. La succession des stimuli était : barre haute sur l’écran, puis une vibration de 975 ms, puis 500 ms plus tard, un cercle rouge, puis une vibration de 475 ms.

**Cadre analytique** : technique d’interaction multimodale passive. Le signal en sortie du système vers l’utilisateur s’effectue en alternance avec deux effecteurs, l’un capable de transmettre un stimulus visuel de manière discrète dont l’information est codée avec le type de codage LANGAGE-symbolique (4 images « barres » == 4 départs + 1 cercle rouge = séparation entre les deux stimuli) et l’autre capable de transmettre un stimulus tactile

discret dont l'information est codée avec une seule vibration d'une certaine DURÉE (1 500 ms == 45°). Les informations des couples sont transmises en alternance et les couples sont transmis en séquence (cf. Figure 6-13).

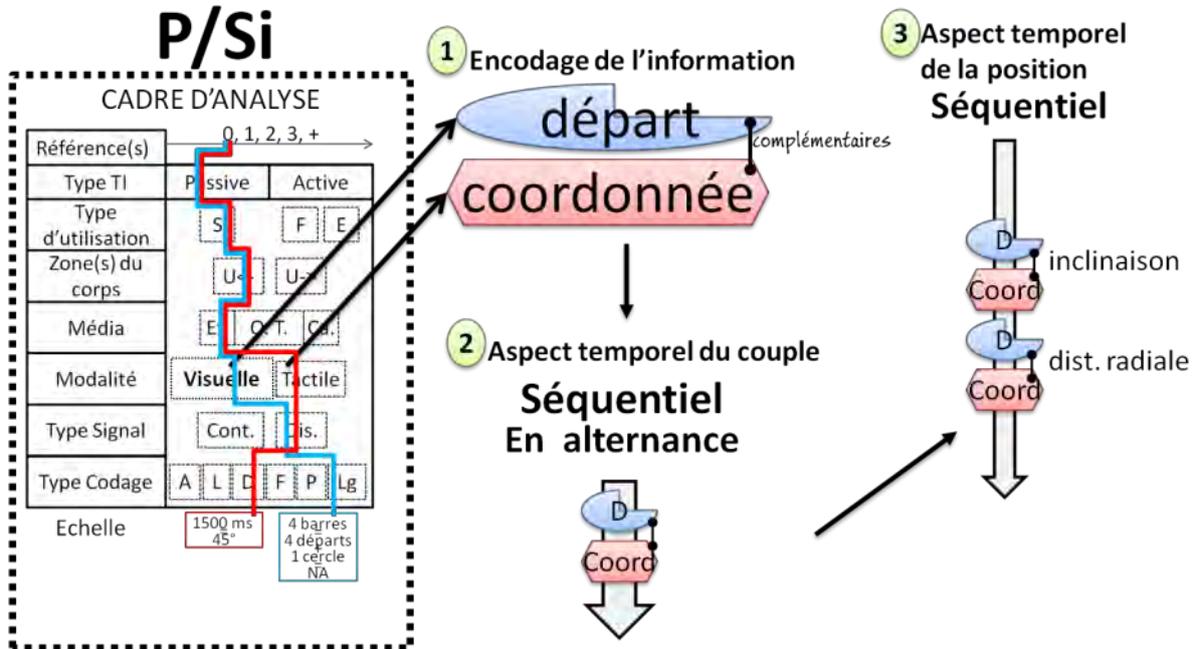


Figure 6-13 : illustration de l'analyse de la technique Singlestimulus dans un SC Polaire (P/Si).

**Matériel** : casque de réalité virtuelle [NVisor® SX-60 HMD, (NVIS Inc., USA)] pour afficher la scène virtuelle et les barres (et cercles) rouges. L'affichage est géré par le moteur 3D [Irrlicht v1.8]. Les vibrations étaient générées à partir d'un seul moteur vibrant [VPM2] et leur durée était contrôlée par une carte [Arduino® ATmega1280]. Enfin, toute l'expérience était gérée grâce à un programme en C#. (cf. la section Dispositif expérimental et stimuli).

#### 6.3.2.4 Train de pulsations dans un SC Polaire (P/Tr)

**Principe** : similaire au pointage d'une cible avec une tourelle dont la base est posée sur la face avant d'un cube. La visée de la tourelle est le champ de vision de l'utilisateur qui avait pour tâche de pointer une position à partir de deux angles : inclinaison et distance radiale. La valeur de l'inclinaison était transmise par un train d'un certain nombre de vibrations et le départ de l'inclinaison était transmis par l'une des quatre barres rouges ; puis, tout en gardant l'inclinaison, l'utilisateur pointe la cible avec la distance radiale dont la valeur était transmise par un train d'un certain nombre de vibrations. Notons qu'un indice visuel « cercle rouge » indiquait à l'utilisateur qu'il allait recevoir la seconde coordonnée (cf. Figure 6-14).

**Illustration** :

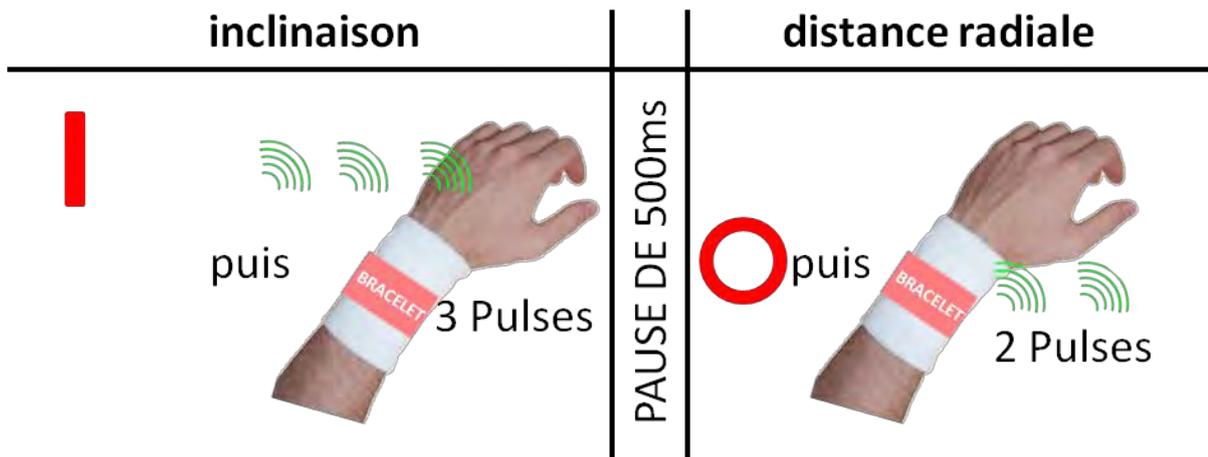


Figure 6-14 : illustration de la technique P/Tr. Exemple illustré : la cible est 27° d'inclinaison et 18° du centre (distance radiale). La succession des stimuli était : barre haute sur l'écran, puis un train de trois pulsations, puis 500 ms plus tard, un cercle rouge, puis un train de deux pulsations.

**Cadre analytique** : technique d'interaction multimodale passive. Le signal en sortie du système vers l'utilisateur s'effectue en alternance avec deux effecteurs, l'un capable de transmettre un stimulus visuel de manière discrète dont l'information est codée avec le type de codage LANGAGE-symbolique (4 images « barres » == 4 départs + 1 cercle rouge = séparation entre les deux stimuli) et l'autre capable de transmettre un stimulus tactile discret dont l'information est avec le type de codage PULSE (succession de 5 vibrations == 45°). Les informations des couples sont transmises en alternance et les couples sont transmis en séquence (voir Figure 6-15).

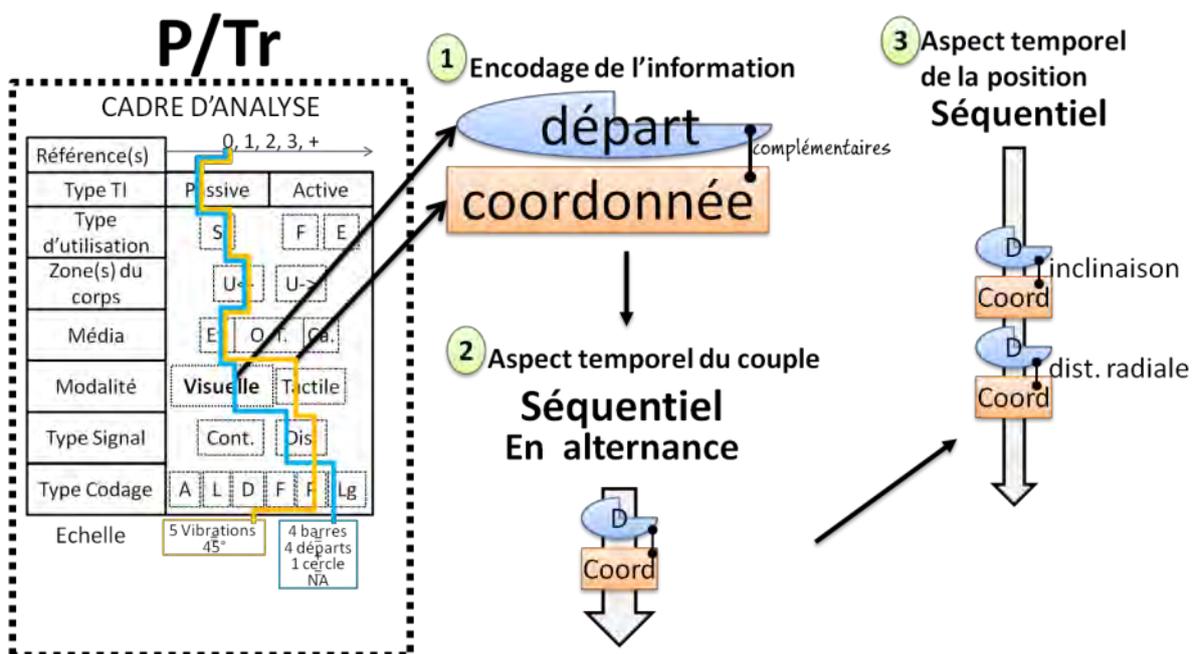


Figure 6-15 : illustration de l'analyse de la technique Train de pulsations dans un SC Polaire (P/Tr)

**Matériel** : casque de réalité virtuelle [NVisor® SX-60 HMD, (NVIS Inc., USA)] pour afficher la scène virtuelle et les barres (et cercles) rouges. L'affichage est géré par le moteur 3D [Irrlicht v1.8]. Les trains de pulsations étaient générés à partir d'un seul moteur vibrant [VPM2] et leur nombre de vibration était contrôlé par une carte [Arduino® ATmega1280]. Enfin, toute l'expérience était gérée grâce à un programme en C#. (cf. la section Dispositif expérimental et stimuli).

### 6.3.3 Évaluation expérimentale (exp. 3.1)

---

Nous avons évalué l'utilisabilité des quatre techniques dans une tâche d'orientation de la tête. Les sujets portaient un casque de réalité virtuelle limitant le champ visuel à dix degrés et un vibreur attaché au poignet (cf. Figure 6-16). Pour chaque condition, ils devaient estimer, le plus précisément possible, la position d'une cible après avoir reçu les deux indices tactiles correspondant à sa coordonnée. Nous avons utilisé un plan expérimental à 2x2 facteurs, avec le système de coordonnées (Cartésien vs. Polaire) et le code tactile (single vs. train) comme facteurs principaux. Chaque sujet a passé les quatre conditions dans un ordre pseudo-aléatoire. Pour chaque condition, les vingt premiers essais étaient utilisés comme une phase de familiarisation et n'étaient pas inclus dans les résultats. Ils étaient suivis par quarante essais où les mouvements de la tête étaient enregistrés.

Dans cette première expérience, un quadrant du référentiel était sélectionné aléatoirement pour chaque essai et était indiqué par un indice visuel. Dans le repère Cartésien, deux flèches rouges étaient affichées sur les deux demi-axes du quadrant choisi. Dans le repère polaire, une barre rouge indiquait un demi-axe parmi les quatre possibles. Les deux coordonnées, correspondant à la position de la cible dans ce quadrant, étaient alors transmises par un seul vibreur, séparées par un silence de 0,5s. Ce couple de modalités contraignait la stimulation sur un seul vibreur et permettait aux sujets de se concentrer uniquement sur le décodage de la stimulation tactile des deux coordonnées.

À la fin de chaque condition, les sujets devaient remplir un questionnaire d'utilisabilité (SUS) sur la technique qu'ils venaient d'utiliser. Pour chaque sujet, l'expérience a duré environ une heure et demie.

#### **PARTICIPANTS**

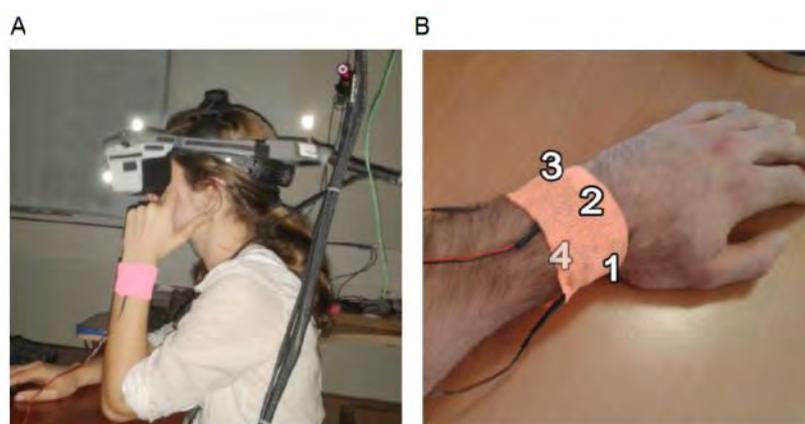
Six sujets (quatre hommes et deux femmes, âgés de  $27,2 \pm 2,6$  ans) ont réalisé l'expérience. Conformément à la Déclaration d'Helsinki, ils ont donné leur consentement éclairé par écrit,

avant de participer à l'étude. L'étude a été approuvée par un comité local d'éthique (CLERIT).

#### **DISPOSITIF EXPERIMENTAL ET STIMULI**

Les sujets étaient placés dans une scène visuelle artificielle contenant seulement un fond bruité (ressemblant à un ciel étoilé). Le fond était utilisé pour fournir un feedback (retour d'information) visuel pendant les mouvements de la tête, mais ne fournissait aucune information sémantique ou symbolique pouvant servir de repère. Pendant l'expérience, le champ visuel du sujet dans le casque de réalité virtuelle était réduit à dix degrés par un masque noir couvrant toute la périphérie du champ visuel.

Les stimuli visuels étaient présentés dans un casque de réalité virtuelle [**NVisor® SX-60 HMD, NVIS Inc., USA**] avec une résolution de 1280 x 1024 pixels, équivalant à 44 x 33 degrés d'angle visuel. L'orientation de la caméra (virtuelle) était directement liée à celle de la tête grâce à un système de capture de mouvement cadencé à 60 Hz [**OptiTrack®, Natural Point, USA**]. Les sujets étaient donc libres de regarder dans l'espace virtuel en bougeant la tête. Les sujets étaient assis sur une chaise fixe qui limitait les mouvements de leur tronc. Ils étaient libres de bouger les yeux, mais une croix rouge présentée, au centre du champ visuel, servait de point d'ancrage. L'environnement virtuel (EV) était généré et animé par le moteur 3D open source **Irrlicht** (v1.8). Le dispositif tactile, spécialement conçu pour l'expérience, utilisait une carte [**Arduino® ATmega1280**] et des vibreurs [**VPM2**]. Les stimuli tactiles étaient transmis via un seul vibreur placé sur le dessus du poignet (cf. Figure 6-16-B). Le vibreur était calibré pour une fréquence de 250 Hz, optimale pour une stimulation tactile sur la peau (Makous *et al.* 1995).



**Figure 6-16 : Photos d'un sujet voyant équipé avec l'interface tactile au poignet et le casque de réalité virtuelle permettant de simuler la vision tubulaire. Les mouvements de la tête sont enregistrés par un système infrarouge, visible en A. Dans l'expérience 3.1, les sujets ne disposent que d'un seul vibreur, placé sur le poignet par une bande de fixation élastique et cohésive (A). Dans les expériences 3.2 et 3.3, quatre**

**vibreurs sont placés sur le poignet gauche (B). Le quatrième vibreur, placé sur la partie antérieure du poignet n'est pas visible sur la photo.**

#### **TACHE COMPORTEMENTALE**

Les sujets ont été testés dans une tâche d'orientation de la tête en présence de l'expérimentateur. Le sujet devait estimer le plus précisément possible, la position d'une cible dans un environnement virtuel. A chaque nouvel essai, une cible était générée au hasard dans un espace de  $90^\circ \times 90^\circ$  d'angle visuel, centré sur la position de la tête. En début d'expérience, chaque sujet devait définir une position de repos, avec la tête bien droite. Cette position était symbolisée par une sphère verte ( $2^\circ$  de diamètre) et allumée quand nécessaire. A chaque début d'essai, le sujet devait aligner le centre de son champ visuel (indiqué par une croix rouge) et la sphère verte. Après une durée aléatoire de 1,4 s à 2 s, les coordonnées de la cible étaient transmises avec le dispositif tactile. A la fin de la transmission, la sphère verte s'effaçait. Le sujet devait alors orienter sa tête vers la position de la cible le plus précisément possible, puis valider cette orientation en appuyant sur la barre d'espace. La validation était toujours suivie par l'affichage de la position de la cible (sphère rouge) et de la position estimée (sphère blanche). Cette phase permettait au sujet d'avoir un feedback (retour d'information) de sa performance à chaque essai. Un nouvel essai commençait lorsque le sujet était revenu à sa position de repos.

#### **6.3.4 Résultats de l'expérience 3.1**

L'erreur de pointage était calculée pour chaque essai, puis moyennée par sujet et par technique. Elle correspond à la distance entre la position de la cible et celle pointée par le sujet. Une analyse de variance à mesures répétées (ANOVA-RM) a montré un effet significatif du type de référentiel (Cartésien vs. Polaire ;  $F=4,6$ ,  $df=1$  ;  $p=0,043$ ) mais pas d'effet du code tactile (Single vs. Train), ni aucune interaction entre ces deux facteurs. La moyenne de l'erreur de pointage liée au référentiel Cartésien ( $9,0^\circ \pm 2,1^\circ$ ) était plus réduite que celle liée au référentiel polaire ( $10,6^\circ \pm 1,6^\circ$ ).

Afin de révéler l'existence d'un biais cognitif ou moteur systématique, comme une sous-estimation ou une surestimation des deux coordonnées, nous avons calculé la moyenne de l'erreur de pointage dans chaque référentiel. Nous n'avons pas observé de biais pour le référentiel Cartésien. Toutefois dans la condition P/Tr, les sujets sous-estimaient systématiquement la distance angulaire (test-t,  $p<0,05$ , après correction de Bonferroni). De même, dans la condition P/Si, les sujets surestimaient la distance angulaire (test-t,  $p<0,05$ , après correction de Bonferroni). De plus, indépendamment des techniques, les résultats

montrent un effet significatif des positions spatiales des cibles (proches/éloignées) sur la précision de pointage. Les sujets sont moins bons à pointer une cible éloignée, avec un ratio moyen de 1,7.

Nous avons finalement comparé les scores d'utilisabilité pour chaque technique par une ANOVA à deux facteurs. Nous avons observé un effet significatif du code tactile sur l'utilisabilité ( $p < 0,05$ ). Le train de pulsations a été considéré comme plus satisfaisant que le Singlestimulus (test-t,  $p < 0,05$ ) (cf. Figure 6-17-B). Pendant les interviews finales, quatre sujets ont considéré la technique C/Tr comme étant leur préférée. Les deux autres sujets ont considéré la technique C/Tr comme leur deuxième choix.

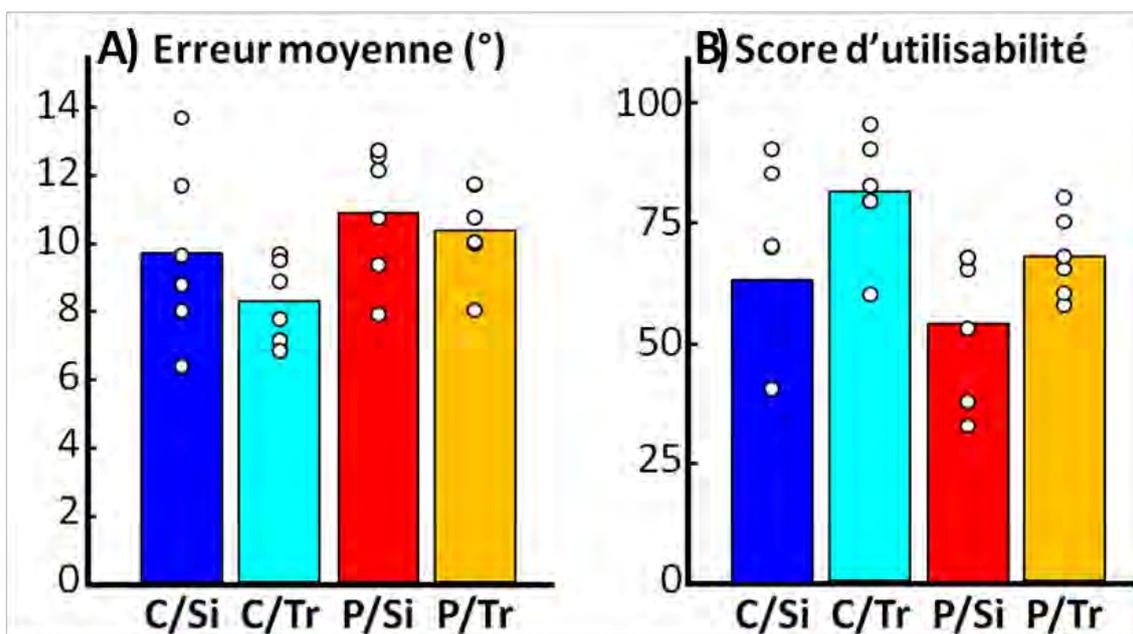


Figure 6-17 : efficacité (A) et satisfaction (B) pour les quatre techniques conçues. A) L'erreur moyenne de pointage pour chacune des techniques est représentée par des barres de couleur. La moyenne de l'erreur de pointage par sujet est représentée avec des points blancs. La technique C/Tr montre l'erreur moyenne la plus faible (8,2°) avec l'écart-type le plus faible (1,2°). B) Le score moyen de satisfaction (SUS) pour chacune de ces techniques est représenté par des barres de couleur. Le score de satisfaction par sujet est représenté avec des points blancs. La technique C/Tr reçoit le score le plus élevé avec 81,3±12,0.

### 6.3.5 Conclusions de l'expérience 3.1

Les réponses de pointage sont plus précises dans le référentiel cartésien. Nous n'avons pas observé de biais systématique et la moyenne de l'erreur est de 9,0 degrés d'angle visuel seulement. Au contraire, les stimuli tactiles dans le référentiel polaire ont conduit à des résultats significativement moins précis, avec des biais systématiques. Par conséquent, il apparaît que le référentiel cartésien est plus précis pour communiquer la position d'une cible à l'aide du dispositif tactile.

La satisfaction liée à ces deux systèmes n'est pas significativement différente. Cependant, le score est meilleur lorsque les coordonnées sont communiquées avec le train de pulsations plutôt qu'avec le Singlestimulus. En combinant précision et satisfaction, la technique reposant sur un train de pulsations dans un référentiel cartésien (C/Tr) est la meilleure parmi les quatre que nous avons conçues. De plus, cette technique a montré la plus faible variabilité inter sujets et la meilleure moyenne au SUS. Le score SUS de cette technique a atteint  $81 \pm 12$ , ce qui correspond à un « bon » score d'utilisabilité (Bangor *et al.* 2008).

## **6.4 Expérience 3.2 : recherche visuelle d'une petite cible dans un salon virtuel encombré**

Dans cette deuxième expérience, nous avons évalué la technique la plus efficace et préférée (C/Tr) dans des conditions proches d'une situation écologique. Nous avons simulé un champ visuel sévèrement réduit ( $10^\circ$  d'angle visuel) avec un casque de réalité virtuelle. Les sujets étaient immergés dans un salon et avaient pour tâche de localiser le plus rapidement possible une cible dans cette salle.

### **6.4.1 Participants**

Dix sujets (six hommes et quatre femmes,  $25,4 \pm 1,6$  ans) ont réalisé l'expérience. Ils ont donné par écrit leur consentement éclairé avant de participer à l'étude, conformément à la Déclaration d'Helsinki. L'étude a été approuvée par un comité d'éthique local (CLERIT).

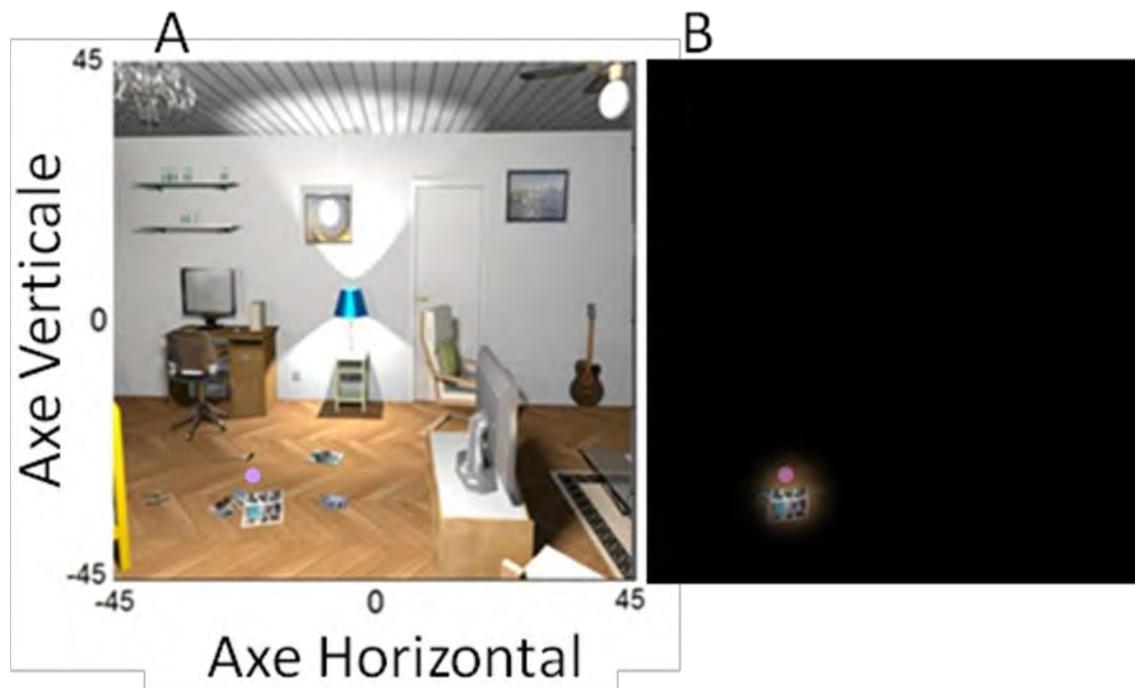
### **6.4.2 Environnement virtuel, dispositif tactile et stimuli**

La configuration expérimentale utilisée dans cette expérience était identique à celle utilisée dans l'expérience précédente.

Une scène virtuelle représentant un salon a été créée avec le logiciel Sweet Home 3D<sup>49</sup>. Une grande quantité d'objets (lampe, chaise, livre, etc.) a été placée de façon homogène dans l'ensemble de la scène visuelle (cf. Figure 6-18-A). La cible (sphère rose) était placée au hasard dans la scène virtuelle, dans une partie de  $90^\circ \times 90^\circ$  d'angle visuel. La scène virtuelle était centrée sur la position de la tête avant chaque essai. Nous nous sommes assurés que la cible et les éléments de la scène visuelle ne pouvaient pas être confondus. Par conséquent, les teintes roses n'étaient pas utilisées dans la scène visuelle. Toutefois, afin d'éviter un effet « pop-out » de la cible, celle-ci était légèrement transparente. Contrairement à une situation écologique, la cible était disposée de façon aléatoire dans la scène et non sur des

<sup>49</sup> Sweet Home 3D, Copyright (c) 2005-2015 Emmanuel PUYBARET / eTeks <[info@eteks.com](mailto:info@eteks.com)> ; <http://www.sweethome3d.com/fr/>

objets. Ce choix permet de mesurer des temps de recherche non biaisés par le contexte, dépendant uniquement de l'assistance tactile. Pendant toute l'expérience, le champ visuel à l'intérieur du casque virtuel était limité à 10°. Un masque noir était appliqué sur le pourtour de la scène (virtuelle), ce qui simulait une vision tubulaire sévère (cf. Figure 6-18-B).



**Figure 6-18 : A – Portion de l'environnement visuel. La cible rose apparaît en bas à gauche. B – Exemple de la simulation artificielle de la vision tubulaire de 10°.**

Suite aux résultats de la première expérience, les deux coordonnées de la cible étaient transmises avec deux trains de pulsations tactiles encodés dans un référentiel Cartésien. Cependant, afin d'être dans des conditions plus réalistes, les indices tactiles étaient transmis, dans cette expérience, par quatre vibreurs disposés autour du poignet gauche, comme s'ils étaient répartis sur un bracelet (cf. Figure 6-16-B). Les vibreurs étaient attachés avec un bandage. Nous avons choisi les axes et leurs directions de façon arbitraire, en fonction d'une position fréquente de la main gauche au repos, la paume vers le bas. Les coordonnées horizontales étaient transmises par les vibreurs placés vers le pouce (direction droite) et le petit doigt (dir. gauche). Les coordonnées verticales étaient transmises par les vibreurs de la partie dorsale (dir. haut) et ventrale (dir. bas). Comme dans la première expérience, les pulsations tactiles avaient une durée fixe de 200 ms séparées par 100 ms. Le nombre de pulsations dans le train correspondait à la zone cible dans l'espace, avec chaque pulsation indiquant une zone de 9°. Le maximum de stimuli était de cinq et correspondait à l'angle maximal de 45°.

**PRESENTATION DE LA TECHNIQUE D'ASSISTANCE A LA RECHERCHE VISUELLE AVEC NOTRE CADRE D'ANALYSE**

**Principe** : similaire au pointage d'une cible avec une tourelle dont la base est posée sur le pont d'un navire. La visée de la tourelle est le champ de vision de l'utilisateur qui avait pour tâche de pointer une position à partir de deux angles : azimut et élévation dont leur valeur était transmise par des trains de vibrations et leur sens était indiqué par la position du vibreur sur le poignet (cf. Figure 6-19).

**Illustration** :

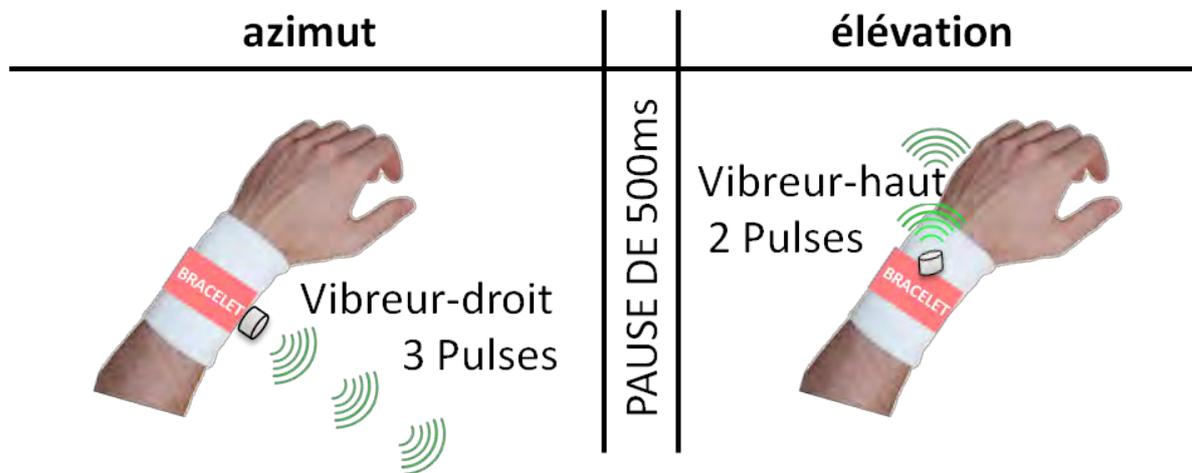


Figure 6-19 : illustration de la technique d'assistance à la recherche visuelle. Exemple pour une cible située à 27° à droite et à 18° en haut : 3 vibrations sur le vibreur droit, puis 500 ms plus tard, 2 vibrations sur le vibreur haut.

**Cadre analytique** : technique d'interaction multimodale passive. Le signal, en sortie du système vers l'utilisateur, s'effectue en synergie avec un effecteur partagé capable de transmettre un stimulus tactile discret dont l'information est composée de deux sous-unités : l'une codée avec le type de codage ANGLE (4 positions = 4 sens) et l'autre avec le type de codage PULSE, c'est-à-dire ici, une succession de vibrations (5 vibrations == 45°). La position des vibreurs autour du poignet donne quatre références explicites : gauche, droite, haut et bas (voir Figure 6-20).

# Tech. d'assistance

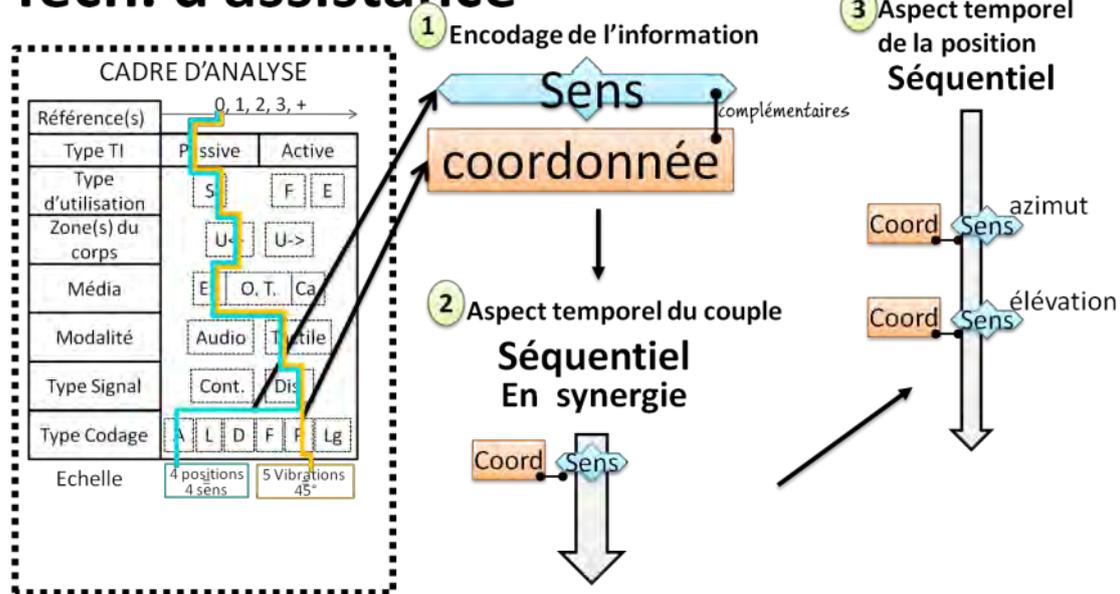


Figure 6-20 : illustration de l'analyse de la technique d'assistance à la recherche visuelle, basée sur la technique de l'exp. 3.1 C/Tr.

**Matériel :** casque de réalité virtuelle [NVisor® SX-60 HMD, (NVIS Inc., USA)] pour afficher la scène virtuelle. L'affichage est géré par le moteur 3D [Irrlicht v1.8]. Les trains de vibrations étaient générés à partir de quatre moteurs vibrants [VPM2] disposés en croix autour du poignet et leur nombre de vibrations était contrôlé par une carte [Arduino® ATmega1280]. Enfin, toute l'expérience était gérée grâce à un programme en C#. (cf. la section Dispositif expérimental et stimuli).

## 6.4.3 Procédure expérimentale (exp. 3.2)

Les sujets étaient évalués individuellement dans une tâche de recherche visuelle, en présence de l'expérimentateur. Le sujet devait estimer, le plus rapidement possible, la position d'une cible dans un environnement virtuel. Comme dans la tâche précédente, la position de repos de la tête a été définie au début de l'expérience, et était marquée dans l'environnement virtuel par une sphère verte (2° de diamètre) quand nécessaire. Le sujet devait maintenir sa position de repos pendant une durée aléatoire entre 700 ms et 1400 ms. Puis, les coordonnées de la cible étaient transmises au sujet. À la fin de la stimulation, la sphère verte disparaissait, et les lumières de la scène s'allumaient. Le sujet devait alors rechercher et trouver la cible le plus rapidement possible, puis valider sa réponse en appuyant sur la barre d'espace.

Pour chaque technique, le sujet commençait par une phase de familiarisation incluant quinze essais qui n'étaient pas enregistrés. Ensuite, la phase de test était répartie en deux

blocs de vingt essais où le temps de recherche et l'ensemble des mouvements de tête étaient enregistrés.

Dans cette expérience, la technique d'assistance était comparée à une condition contrôle absolument identique, sauf qu'elle n'incluait aucun indice tactile concernant l'emplacement de la cible. Chaque sujet a passé quarante essais répartis en deux blocs correspondant aux deux techniques. Cinquante pour cent des sujets ont commencé par la condition contrôle. Les autres ont commencé par la technique d'assistance.

L'ensemble de l'expérience a duré environ soixante minutes par sujet. À la fin de l'expérience, nous avons demandé aux sujets de remplir un questionnaire de satisfaction (SUS) concernant la technique d'assistance.

#### **6.4.4 L'analyse des données**

---

Nous avons mesuré le temps de recherche pour chaque essai. Il correspond à la durée écoulée entre l'envoi de la première stimulation et la validation du sujet. Cela correspond au temps nécessaire pour effectuer la recherche visuelle. La précision temporelle de 60 Hz nous a permis d'analyser finement, pour chaque essai, la trajectoire de la tête au cours de la recherche visuelle. La distance totale parcourue a été calculée pour chaque essai. De plus, nous avons calculé un ratio exprimant la distance totale sur la distance optimale. La distance optimale est la plus courte distance entre l'origine du référentiel et la cible. Ce ratio est complémentaire au temps de recherche.

#### **6.4.5 Résultats de l'expérience 3.2**

---

Dans la condition contrôle, nous avons observé différentes stratégies de recherche visuelle en fonction des sujets (cf. Figure 6-21). Certains sujets ont effectué un motif de balayage sur l'axe vertical ou sur l'axe horizontal (cf. Figure 6-21-A et B resp.). D'autres sujets ont effectué des cercles concentriques autour du centre de la scène visuelle (cf. Figure 6-21-C). Ces motifs particuliers n'ont pas été observés en présence d'indices tactiles. Dans ces conditions, les sujets déplaçaient tout d'abord leur tête sur l'axe horizontal, puis sur l'axe vertical (cf. Figure 6-21-D et E). Nous avons également observé chez certains sujets, des mouvements en diagonale. Cela suggère une bonne intégration des deux coordonnées (horizontales et verticales) avant le mouvement de la tête (cf. Figure 6-21-F).

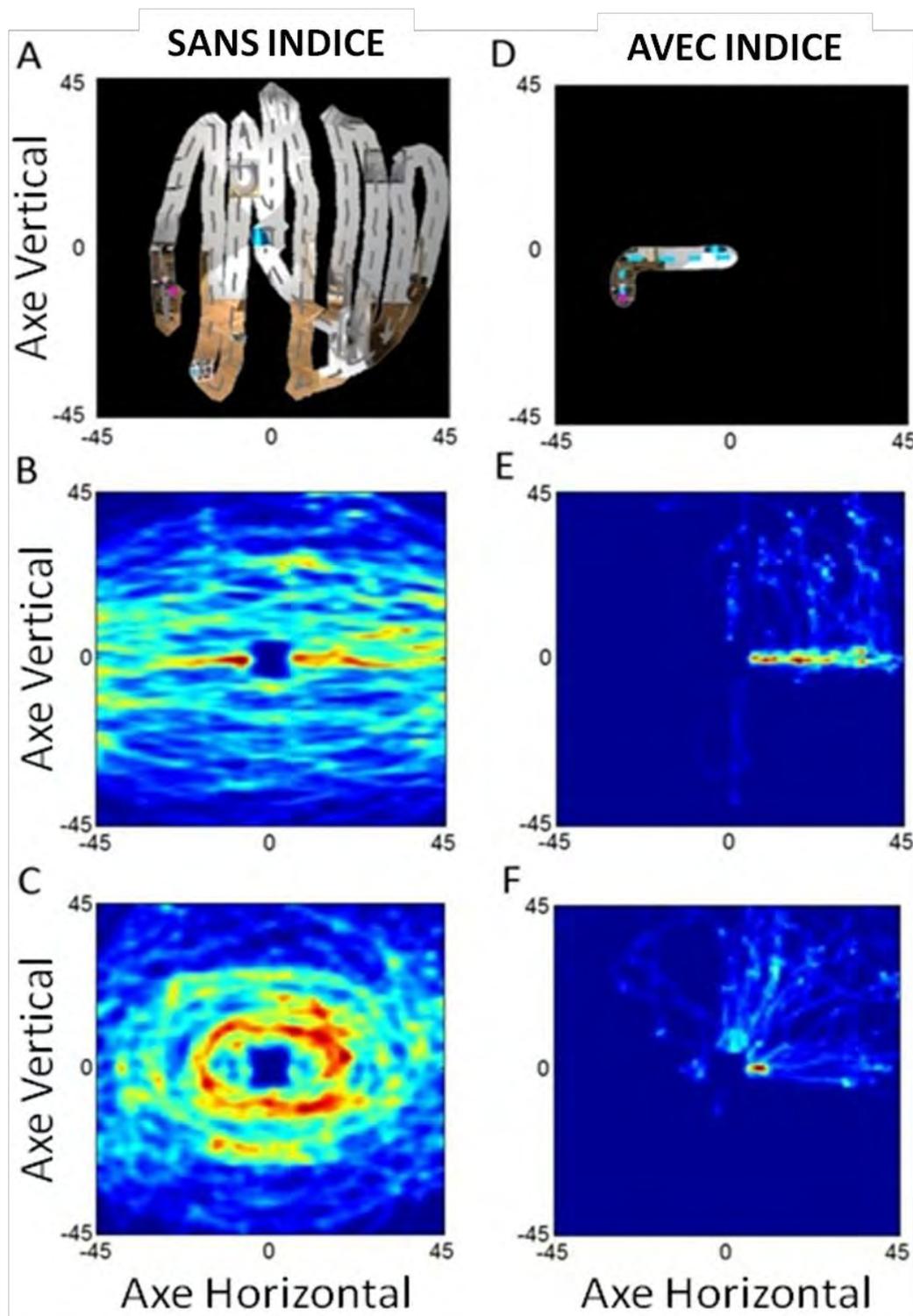


Figure 6-21 : Exemples de mouvements de tête observés pendant une recherche visuelle dans les conditions contrôle, sans indice (A, B, C) et avec la technique d'assistance, avec indice, (D, E, F). A - Balayage sur l'axe vertical. B – Carte de densité d'une recherche avec un balayage principalement sur l'axe horizontal. C – Carte de densité d'une exploration en cercles concentriques. D – Mouvements ordonnés le long de l'axe horizontal puis vertical suivant l'ordre de transmission des deux indices tactiles. E – Carte de densité d'un mouvement commençant essentiellement par l'axe horizontal. F- Carte de densité d'un mouvement intégré. Ici, le sujet a été capable d'intégrer les deux coordonnées avant de bouger, puis de déplacer la tête en faisant des mouvements obliques vers les cibles. Chaque carte de densité représente les mouvements de tête d'un sujet. La position de la cible et les mouvements de tête ont été normalisés avec une transformation symétrique sur l'axe horizontal puis vertical, en plaçant les cibles dans le coin supérieur droit. Le rouge indique des densités élevées et le bleu des densités nulles.

Nous avons calculé les temps de recherche médians par sujet et par technique. Les moyennes de ces temps de recherche médians sont présentées dans la Figure 6-22-A. Les sujets passaient en moyenne  $16,8 \pm 2,5$  secondes pour trouver une cible dans la condition contrôle, sans indice ; et  $5,4 \pm 0,8$  secondes avec la technique d'assistance, avec indices tactiles. Nous avons observé une différence similaire pour chaque sujet. La différence est très significative (test-t,  $p < 0,001$ ). Nous avons également calculé la distance médiane parcourue par sujet et par condition. Chaque distance a été comparée à la distance optimale (distance la plus courte entre la position de la cible et la position initiale du sujet). En condition contrôle, la distance médiane était en moyenne dix fois plus importante que la distance optimale (cf. Figure 6-22-B). Avec des indices tactiles, ce ratio est en moyenne de  $1,2 \pm 0,2$ . Nous avons observé un ratio similaire pour tous les sujets. Par conséquent, l'analyse a montré une différence très significative (test-t  $p < 0,001$ ). Nous n'avons observé aucune amélioration de leur temps de recherche, ni de la distance parcourue, que ce soit dans les conditions contrôle ou avec la technique d'assistance.

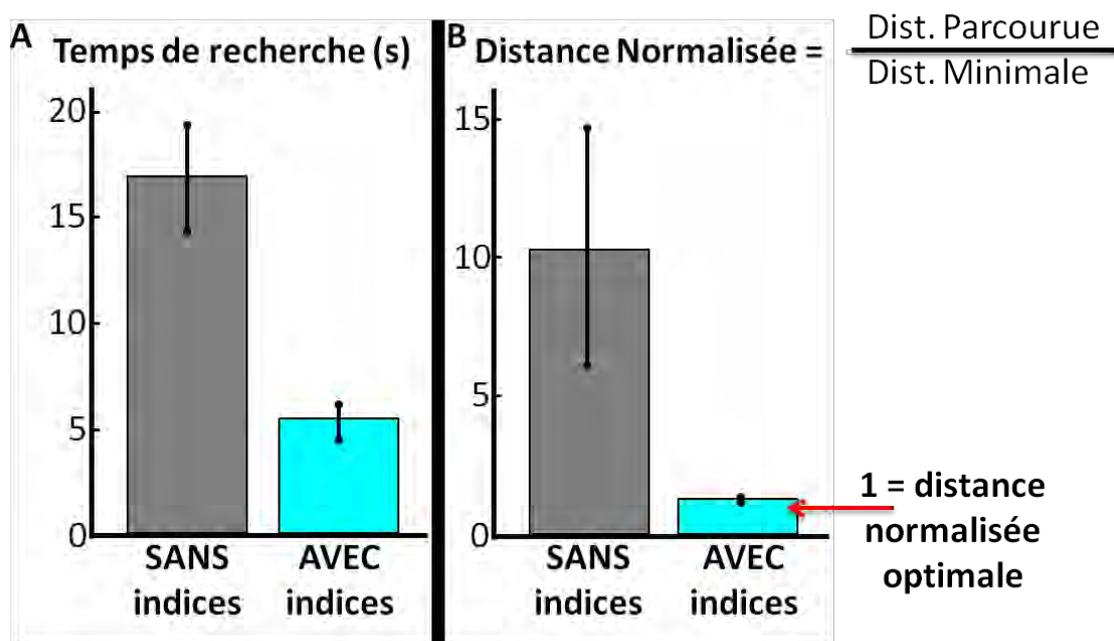


Figure 6-22 : Performance de la recherche visuelle (A – temps de recherche et B- distance parcourue) avec et sans indices tactiles. En gris, la condition contrôle sans indice ; en cyan, la technique d'assistance, avec indices.

Une des erreurs observées avec la technique d'assistance était liée à l'inversion de la direction sur les axes. Nous avons calculé les erreurs d'inversion selon une règle simple. Une erreur est détectée lorsque la tête se déplace dans la direction opposée, et atteint le seuil de cinq degrés d'angle visuel. En moyenne, les sujets ont fait  $4,5 \pm 3,3$  % (resp.  $8,8 \pm 4,0$  %) d'erreurs d'inversion sur l'axe horizontal (resp. vertical). Le score de satisfaction (SUS) pour

la technique d'assistance était en moyenne de  $82 \pm 16$ , ce qui est considéré comme « bon » (Bangor *et al.* 2008).

#### **6.4.6 Conclusions de l'expérience 3.2**

L'expérience a montré que la recherche visuelle avec une assistance tactile était en moyenne trois fois plus courte. De plus, l'assistance tactile était fiable et satisfaisante. L'expérience a donc clairement démontré l'utilisabilité de la technique pour assister une recherche visuelle dans un environnement visuellement encombré avec la vue tubulaire simulée.

### **6.5 Expérience 3.3 : étude de cas sur un sujet avec une vision tubulaire**

Dans cette dernière expérience, nous avons souhaité vérifier que les résultats observés avec une simulation sont toujours valides dans le cas d'un utilisateur déficient visuel. Nous avons donc répété l'expérience avec un sujet ayant un champ visuel restreint dû à un glaucome.

#### **6.5.1 Participant**

Un homme de 47 ans a réalisé cette expérience complémentaire. Il s'est porté volontaire suite à une annonce passée dans un centre d'éducation spécialisée pour déficients visuels (CESDV-IJA). Il a donné son consentement éclairé par écrit avant de participer à l'étude conformément à la Déclaration d'Helsinki. Le sujet présentait un champ visuel bilatéral de  $10^\circ \times 15^\circ$ . Ce sujet avait une perception altérée des couleurs et une adaptation lente à la luminosité, ce qui est souvent observé chez ces patients (Nelson, Aspinall, and O'Brien 1999).

#### **6.5.2 Environnement virtuel, dispositif tactile et cibles**

Le sujet était placé dans les mêmes conditions que précédemment, excepté les points suivants. La couleur de la cible a été changée en un dégradé de couleur (du jaune au blanc), et la transparence a été supprimée. Dans le but de réduire l'inconfort dû aux changements de luminosité, nous l'avons gardée constante pendant toute l'expérience. Évidemment, le champ visuel n'était pas artificiellement réduit par un masque noir.

#### **6.5.3 Procédure expérimentale (exp. 3.3)**

La procédure était la même que dans l'expérience précédente, à l'exception que la lumière n'était jamais éteinte pendant les essais. Le sujet a tout d'abord réalisé une série de trente essais (dont dix pour se familiariser) sans assistance tactile. Puis, il a réalisé une série de trente-cinq essais (dont quinze pour se familiariser) avec assistance tactile.

### 6.5.4 Résultats de l'expérience 3.3

Nous avons calculé la durée médiane pour trouver la cible avec et sans indices tactiles. Elle était significativement réduite avec les indices tactiles (9,4 s) comparée à la condition contrôle (27,3 s) (Wilcoxon  $p < 0,001$ ), cf. Figure 6-23. Le score de satisfaction (SUS) pour la technique d'assistance était de 72,5, ce qui est considéré comme « bon » (Bangor *et al.* 2008). Enfin, la discussion ouverte avec l'utilisateur nous a permis de recueillir des résultats qualitatifs. Celui-ci a notamment remarqué que « c'est vraiment une technique prometteuse » qui va « l'aider à localiser un objet ». Si elle était disponible, il s'en « servirait régulièrement ». Il a aussi précisé qu'il a « eu parfois quelques difficultés à détecter les vibrations », et que, par conséquent, « il faudrait ralentir la fréquence des trains de stimulations ».

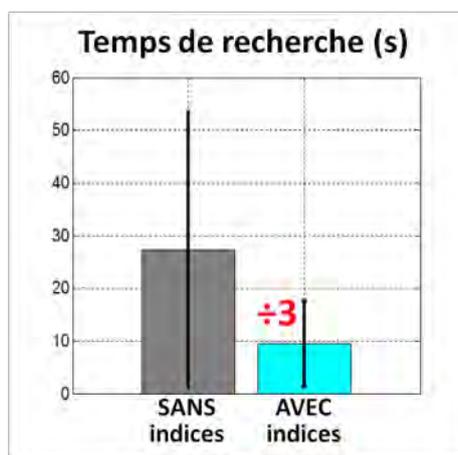


Figure 6-23 : Performance de la recherche visuelle en temps de recherche, avec et sans indices tactiles. En gris, la condition contrôle sans indices ; en cyan, la technique d'assistance, avec indices.

## 6.6 Discussion et conclusion de la troisième étude

Dans la première expérience, nous avons conçu quatre codes tactiles différents transmis par un seul vibreur fixé au poignet. Les indices d'anticipation étaient transmis avant tout mouvement de tête, et visaient à fournir l'assistance pour localiser une cible dans un espace frontal. La première observation est que tous ces codes permettent à l'utilisateur d'estimer correctement la position d'une cible dans un repère égocentrique lié à l'orientation de la tête, ce qui est similaire à la façon dont la mémoire peut stocker des informations spatiales (Medendorp *et al.* 2008). Toutefois, la comparaison directe de ces quatre techniques dans une tâche d'orientation de la tête avec une vision tubulaire artificielle a montré qu'un train de pulsations transmis dans un système de coordonnées Cartésien était plus efficace et plus satisfaisant. Dans cette technique, deux trains de pulsations correspondent à l'abscisse et

l'ordonnée de la cible relatives à l'orientation courante de la tête. Et, un nombre croissant de pulsations (vibrations) dans le train indique que la portion recherchée de l'espace est plus éloignée que la position actuelle.

Dans l'expérience suivante, nous avons évalué cette technique, considérée comme la plus efficace et la plus satisfaisante, dans une situation plus proche d'une tâche de recherche visuelle réelle. Ainsi, les sujets avaient un champ visuel équivalent à une vision tubulaire de 10°. Ils devaient trouver une petite cible dans un salon encombré. Les sujets portaient un dispositif tactile conçu par nos soins, composé de quatre vibreurs répartis autour du poignet. Nous avons choisi d'associer une direction par vibreur (Haut, Bas, Gauche, Droite) et un axe par paire de vibreurs (vertical vs. horizontal). L'évaluation de la technique a montré l'efficacité de ce dispositif puisqu'il a permis une estimation très efficace de la localisation de la cible dans un environnement surchargé. Encore une fois, il est important de noter que ce n'est pas une technique de guidage mais d'anticipation. L'information est fournie de façon brève, avant tout mouvement, et permet d'estimer la position de la cible dans un repère centré sur la tête. La recherche visuelle, réalisée dans un environnement encombré, était trois fois plus rapide avec les indices d'anticipation qu'en leur absence. La distance parcourue pour trouver la cible était aussi dix fois plus courte avec les indices. De plus, tous les sujets testés ont montré une réduction significative du temps de recherche et de la distance parcourue. Une autre observation importante est que les directions étaient dans la grande majorité des cas (>90%) correctement interprétées, ce qui suggère d'une part que le bracelet permet une bonne discrimination des vibrations autour du poignet, mais aussi que le code que nous avons utilisé est facile à interpréter.

Comme nous l'avons mentionné, le choix de la durée des stimulations tactiles a été basé sur des pré-tests succincts permettant de se situer dans une zone de discriminabilité confortable. Au niveau physiologique, la discrimination temporelle entre deux stimulations d'une même partie du corps est en moyenne entre 25 ms et 40 ms (Hoshiyama, Kakigi, and Tamura 2004). Il serait donc possible d'envisager de raccourcir le temps de stimulation (i.e. la durée des trains de pulsations). Cependant, nous avons noté que le ressenti des utilisateurs était très variable. Les sujets voyants (exp. 3.1 et 3.2) n'ont fait aucun commentaire sur une éventuelle optimisation des stimulations, alors que le sujet déficient visuel (exp. 3.3) aimerait les ralentir. Il est probable que de nombreux facteurs influencent la perception des stimuli tactiles (déficience visuelle éventuelle mais aussi fatigue, âge etc.). Il

serait donc nécessaire de réaliser une étude systématique permettant d'optimiser la durée et la fréquence des stimuli. Nous pourrions aussi imaginer un système dans lequel l'utilisateur peut aisément choisir ces deux paramètres.

Au cours de l'expérience, nous n'avons observé ni augmentation ni diminution des performances de la recherche visuelle entre les essais. Cette observation signifie que l'utilisateur avait besoin de moins de quinze essais, correspondant à la phase de familiarisation, pour atteindre un bon niveau de performance. Cela suggère un apprentissage rapide du code tactile. Enfin, bien que notre prototype soit un prototype de laboratoire bon marché (quatre vibreurs attachés avec un bandage), il a été, de façon surprenante, très apprécié. Avec des scores de satisfaction et d'efficacité très bons, il apparaît que ce dispositif, destiné aux personnes ayant un champ visuel restreint, pourrait être très utilisable pour transmettre une connaissance a priori de cibles environnantes. Une première amélioration de notre approche serait d'augmenter sa portée, i.e. d'indiquer la position d'objets dans une zone plus large, voire de 360° autour de l'utilisateur. Afin d'accroître les performances de l'approche, nous nous focaliserions sur la réduction du temps de transmission des coordonnées, tout en conservant la bonne discriminabilité de chacune des vibrations. La diminution de la durée des vibrations et des intervalles serait une bonne piste, mais il y en a d'autres, comme par exemple, transmettre en simultanément les deux coordonnées ce qui permettrait également de réduire le temps de la transmission d'information. Une autre piste envisagée serait d'accroître la résolution de la grille de recherche qui est actuellement de 121 positions de cible (11 x 11) afin de rechercher des cibles en extérieur, donc plus éloignées de l'utilisateur, et donc plus difficile à trouver avec la résolution de la grille actuelle.

Il était important de valider notre approche sur un sujet déficient visuel. Différentes études comportementales ont déjà montré que cette même approche est fiable dans des tâches de cognition spatiale (Fortenbaugh *et al.* 2007; Fortenbaugh, Hicks, and Turano 2008; Turano *et al.* 2005). Néanmoins, les personnes avec une vision tubulaire peuvent présenter d'autres déficits tels que la faible détection des contrastes, la confusion des couleurs, ou la perte de l'acuité visuelle (Gawande *et al.* 1989; Young and Fishman 1982). Or ces déficits n'étaient pas simulés dans notre étude. Ils peuvent également développer des stratégies particulières pour la recherche visuelle lors de la perte progressive de leur champ visuel. Ces stratégies peuvent entraver les résultats que nous avons observés dans cette expérience. Par

conséquent, nous avons évalué le dispositif sur un sujet ayant un champ visuel restreint en raison d'un glaucome. Son champ visuel était cohérent avec la simulation, mais sa capacité à détecter les couleurs était aussi altérée. Nos résultats montrent qu'avec des indices d'anticipation, il était trois fois plus rapide pour trouver une cible. Le ratio est similaire à celui observé avec la simulation, bien que le sujet déficient visuel soit légèrement moins rapide pour trouver les cibles dans les deux conditions. Une capacité plus faible de détection des contrastes et des couleurs pourrait expliquer cette différence. Au cours du débriefing, le sujet déficient visuel mentionnait qu'une partie de son attention était portée sur la discrimination des vibreurs actifs. Toutefois, le sujet a montré un grand intérêt pour le dispositif d'assistance et le score (72,5) qu'il a donné confirme sa satisfaction. Néanmoins, la suppression de la transparence a probablement augmenté l'effet pop-out de la cible et facilité sa détection. Cependant dans l'exp. 3.3, l'effet pop-out était identique dans les deux conditions. Et après une analyse supplémentaire, le pourcentage de réduction en termes de durée, entre la condition contrôle et la technique d'assistance, observé dans les exp. 3.2 et 3.3 était similaire (67% pour l'exp. 3.2 et 65% pour l'exp. 3.3). Cela confirme l'hypothèse que la diminution du temps de recherche dans l'exp. 3.3 est bien due aux informations tactiles.

La prochaine étape consistera à tester l'utilisabilité du dispositif en situation réelle. Il peut facilement être intégré dans des lunettes intelligentes actuellement en développement (i.e. Google Glass) ou même commercialisées (OrCam). Tous ces dispositifs sont capables de reconnaître et localiser des cibles d'intérêts, et de calculer leur position relative à celle de la tête. Les algorithmes de vision par ordinateur sont en constante progression, et reconnaissent déjà des visages ou des objets avec précision. Il est possible de distinguer des milliers d'objets sur la base d'images en temps réel (Dean *et al.* 2013). Ces cibles d'intérêt pourraient être détectées, soit sur demande grâce à des commandes vocales, soit être suggérées par le dispositif de façon automatique (par ex. des obstacles potentiels ou des cibles récurrentes), ce qui améliorerait la compréhension spatiale de l'environnement autour de l'utilisateur. Une version commerciale de notre bracelet tactile pourrait facilement être intégrée dans le bracelet d'une « Smartwatch » et communiquer par Bluetooth. Une étude récente a notamment mis en évidence l'intérêt croissant que les personnes déficientes visuelles portent aux Smartphones et autres appareils portables (Ye *et al.* 2014). Il est également intéressant de noter qu'en plus de la recherche visuelle, un dispositif tactile sur le poignet pourrait être utile dans d'autres situations quotidiennes

problématiques pour les déficients visuels, comme la navigation piétonne (Brock *et al.* 2014) ou le transport (Brunet *et al.* 2013).

# **Chapitre 7**

## **Résumé des contributions, conclusion générale et perspectives**



## 7 Résumé des contributions, conclusion générale et perspectives

Dans le cadre de cette thèse, nous avons travaillé sur la conception et l'évaluation de techniques d'interaction non visuelle optimisées pour de la transmission d'information(s). Dans le chapitre 2, nous avons proposé un état de l'art sur les modalités auditives et tactiles, mais aussi sur les propriétés de la multimodalité. Cet espace de connaissance a été nécessaire pour la conception de nos techniques. C'est pourquoi, dans le chapitre 3, nous avons synthétisé ces connaissances sous la forme de propriétés générales de conception puis nous les avons regroupées dans des facteurs de conception et enfin, nous les avons organisées et intégrées dans notre cadre d'analyse sous la forme de paramètres. Afin d'introduire les trois chapitres suivants, nous avons pris le temps de décrire huit techniques d'interaction dans le chapitre 3, et d'expliquer trois cas d'études réalisées dans ces travaux de thèse.

### 7.1 Résumé des résultats de la première étude

Le chapitre 4, présente la première étude et porte sur l'estimation d'une seule valeur entre 1 et 100. Cette étude a mis en œuvre quatre facteurs de conception, sur lesquels nous nous sommes basés pour concevoir dix-neuf techniques d'interaction non visuelle. Ces techniques ont été évaluées et comparées entre elles par trois variables observées : l'erreur d'estimation, le score SUS des techniques et leur ordre de préférence.

Les résultats de la première expérience (exp. 1.1) ont montré que l'utilisation de la modalité auditive (ici, un son) était équivalente à la modalité vibrotactile. Ce qui signifie que l'utilisation d'un stimulus auditif pour estimer une valeur entre 1 et 100 avec sa durée est équivalente à l'utilisation d'un stimulus vibrotactile pour estimer une valeur entre 1 et 100 avec sa durée. De même, l'utilisation d'un feedback (retour d'information) auditif, pour marquer la fin de la tâche d'estimation, est équivalente à l'utilisation d'un feedback vibrotactile pour marquer la fin de la tâche d'estimation. Nous pouvons en déduire que le fait que ces deux modalités soient équivalentes dans cette tâche, offre une certaine flexibilité de conception, mais aussi d'adaptation suivant le contexte d'utilisation. En effet, il serait plus pertinent de choisir la modalité vibrotactile pour utiliser une des ces techniques dans un environnement bruyant par exemple. Mais aussi, cela laisse à l'utilisateur le choix par préférence, c'est-à-dire qu'il peut choisir la modalité avec laquelle il a le plus d'affinité. Enfin, dans les travaux de (Coutaz *et al.* 1995; Nigay and Coutaz 1997) sur les « CARE-like

properties of the user » (ou « USER-CARE properties »), les auteurs définiraient ces deux modalités comme U-Equivalent.

Les résultats de la seconde expérience (exp. 1.2) ont validé nos trois hypothèses. La première est basée, notamment, sur les observations de (Gibson 1962) qui a démontré qu'une participation active d'un utilisateur dans une tâche de reconnaissance de forme était deux fois plus efficace qu'une participation passive. Nous avons voulu reproduire cette observation dans notre contexte de transmission d'information, et plus précisément dans une tâche d'estimation d'une valeur entre 1 et 100. Nos résultats ont montré que les techniques dites « Active » (voir Le tableau de synthèse : huit paramètres de base (Tableau 3-4)) étaient meilleures que les techniques dites « Passive », ce qui est cohérent avec les observations de (Gibson 1962). Notre seconde hypothèse est basée sur les travaux de (Fraisse 1984), où P. Fraisse suggère que la présence d'un son avec durée fixe, juste avant ou après un son avec une durée variable améliorerait la précision de l'estimation de la durée (variable) du son. Nous avons voulu tester sa recommandation, pour d'une part des stimuli auditifs, et d'autre part des stimuli tactiles. Nos résultats ont montré que la présence d'une référence juste avant le stimulus auditif variable améliore l'efficacité de 50% (cf. Techniques de bases versus optimisées : Audio), ce qui est cohérent avec les conclusions de (Fraisse 1984). Cependant, nous avons poussé l'hypothèse plus loin et avons montré que la présence de références pendant le stimulus améliore l'efficacité de 81%. Toutefois, le nombre de références n'était pas le même d'une technique à l'autre, ce qui peut expliquer la grande différence entre le fait de mettre la référence juste avant ou pendant le stimulus, nous en discuterons ci-après. Notre troisième hypothèse est basée sur les observations de l'expérience 1.1 des facteurs de conception « Actif/Passif » (**A/P**) et « AvecRéférences / SansRéférence » (**AR/SR**) dans l'expérience 1.2. Nous avons souhaité ordonner ces facteurs et leurs quatre combinaisons de la manière suivante : **A+AR** > **P+AR** > **A+SR** > **P+SR**. Nos résultats n'ont pas été homogènes quant au respect de cette règle, c'est-à-dire que l'analyse de nos variables observées était contradictoire. Par exemple, en termes d'efficacité l'hypothèse est violée tandis qu'en termes de dispersion (écart-type) l'hypothèse est respectée. Afin de prendre en compte les résultats d'un point de vue général, nous avons construit une nouvelle mesure, que nous appelons « Puissance d'impact », à partir de cinq

autres<sup>50</sup>. Selon cette nouvelle mesure, l'ordre **A+AR > P+AR > A+SR > P+SR** est respecté, et notre troisième hypothèse est validée par cette mesure. Cependant, lorsque nous appliquons cette hypothèse à toutes nos techniques d'interaction de cette première étude, certaines techniques ne sont pas à leur place dans le classement, c'est-à-dire qu'elles violent l'hypothèse. Le Tableau 7-1 rappelle ces techniques.

Classement	Techniques	Puissance d'impact	Hypothèse violée
1	« Règle-Graduée », A+AR, 6 Réf.	0.93	NON
2	« PalmGauge » ; P+AR, X Réf.	0.84	OUI
3	« Slider » ; A+AR, 2 Réf.	0.80	NON
4	« BipMultiple », P+AR, 6 Réf.	0.80	OUI
5	« Pédale » ; A+AR, 2 Réf.	0.76	NON
6	« Volant » ; A+AR, 3 Réf.	0.75	NON
7	« DualStimuli » ; P+AR, 1 Réf.	0.58	NON
8	« Règle-Lisse » ; A+SR	0.31	NON
9	« BipSolo » ; P+SR	0.20	NON
10	« SingleStimulus » ; P+SR	0.18	NON
11	« ClicOnMe » ; A+SR	0.14	OUI
12	« Pipe » ; A+SR	0.09	OUI

**Tableau 7-1 : Classement des 12 techniques de la première étude, en fonction de leur « puissance d'impact » ; avec en rouge celles qui violent la règle. A+AR > P+AR > A+SR > P+SR.**

Les sections suivantes discutent de ce problème à l'aide de notre cadre d'analyse et des facteurs de conception.

#### **P+AR VERSUS A+AR : PALMGAUGE ET BIPMULTIPLE**

Dans notre première expérience, nous avons montré que les techniques avec références étaient plus efficaces que celles sans référence. Cependant, nous n'avons pas entrepris de mesurer l'impact du nombre et/ou de la nature des références dans les résultats. Il ressort toutefois que le classement par puissance d'impact semble suivre deux règles :

- Règle 1. Plus le nombre de références est élevé, plus la technique aura une puissance d'impact élevée ;

<sup>50</sup> Voir chapitre 4 : la puissance d'impact d'une technique résulte d'un calcul dans lequel nous faisons la moyenne des rangs normalisés des classements ordonnés pour cinq mesures : précision, dispersion, score SUS, préférence, coefficient de dérive.

Règle 2. Si l'utilisateur, pendant l'utilisation d'une technique, peut explorer de manière active les références avant, après ou pendant le stimulus, alors la technique aura une puissance d'impact élevée.

La première règle semble évidente, la seconde mérite quelques explications. Nous avons vu que fournir des références, avant, après ou pendant le stimulus, permettait d'améliorer l'estimation (cf. chapitre 4<sup>51</sup>). Nous avons vu également qu'une exploration active permettait d'améliorer la reconnaissance de formes (voir (Gibson 1962)), donc cela devrait aussi fonctionner pour estimer une valeur située physiquement sur un objet (ici, en forme de « segment » ou de « demi-cercle »). Vérifions donc, pour chacune des techniques de la première étude (exp. 1.1 et 1.2), si les règles R1 et R2 peuvent s'appliquer. Le Tableau 7-2 reprend le classement des douze techniques de la première étude.

Classement	Techniques	Puissance d'impact	Règles suivies
1	« Règle-Graduée », A+AR, 6 Réf.	0.93	R1 & R2
2	« PalmGauge » ; P+AR, X Réf.	0.84	R1 & R2
3	« Slider » ; A+AR, 2 Réf.	0.80	R2
4	« BipMultiple », P+AR, 6 Réf.	0.80	R1
5	« Pédale » ; A+AR, 2 Réf.	0.76	R2
6	« Volant » ; A+AR, 3 Réf.	0.75	R2
7	« DualStimuli » ; P+AR, 1 Réf.	0.58	-
8	« Règle-Lisse » ; A+SR	0.31	-
9	« BipSolo » ; P+SR	0.20	-
10	« SingleStimulus » ; P+SR	0.18	-
11	« ClicOnMe » ; A+SR	0.14	-
12	« Pipe » ; A+SR	0.09	-

**Tableau 7-2 : Classement des 12 techniques de la première étude, en fonction de leur « puissance d'impact » ; avec les règles R1 et R2 qu'elles suivent.**

La technique « PalmGauge » n'a pas de référence physique, toutefois, cette technique permet à l'utilisateur de choisir et donc, d'explorer ses propres références en associant des zones de sa main à des valeurs. Ainsi, dans cette technique, l'utilisateur a la possibilité, non seulement, d'explorer activement ses références, mais aussi d'en choisir autant qu'il peut en

<sup>51</sup> Première étude : techniques d'interaction non visuelle pour estimer une valeur

discriminer sur sa main. Or, d'après (Johnson and Phillips 1981), la main est une zone du corps avec une grande capacité à discriminer des stimuli proches les uns des autres.

Cependant, avec ces deux nouvelles règles, la technique « Volant » n'est pas à sa place dans le classement. En effet, la technique « Volant » a plus de références explicites que la « Pédale » mais a une puissance d'impact inférieure. Dans la conception de la technique « Volant » qui compte trois références, nous avons compté comme références : la référence marquant le début de la rotation (milieu), et celles des deux butées physiques de fin de rotation (vers la gauche et vers la droite). Il y a là, sans doute, une erreur de jugement. En effet, les utilisateurs allaient soit vers la gauche soit vers la droite, donc d'un point de vue fonctionnel, l'utilisateur ne pouvait utiliser que la butée physique de gauche ou que la butée physique de droite. Il serait donc judicieux de réduire son nombre de références explicites à deux. Ainsi, les techniques « Pédale » et « Volant » sont quasiment équivalentes en termes de puissance d'impact. Cela peut s'expliquer par leur conception similaire à partir de notre cadre d'analyse : seule la zone du corps sollicitée, l'échelle et le dispositif diffèrent. En effet, la « Pédale » qui sollicite le poignet, et a une échelle de 22° pour 100 valeurs et, le « Volant » qui sollicite les deux bras, a lui une échelle de 90° pour 100 valeur (si l'on considère une seule butée). Parmi ces trois facteurs, il serait intéressant de déterminer le ou lesquels impactent les résultats et expliquer, par exemple, le fait que la « Pédale » a de moins bons résultats quantitatifs, mais a de meilleurs résultats qualitatifs que le « Volant ».

Ainsi, les règles R1 et R2 justifient le classement des six premières techniques en termes de puissance d'impact et nous avons expliqué pourquoi les techniques « PalmGauge » et « BipMultiple » violent l'hypothèse générale :  $A+AR > P+AR > A+SR > P+SR$ , par la présence de nombreuses références dans « BipMultiple » (R1) et parce que le « PalmGauge » permet à l'utilisateur d'explorer de manière active de nombreuses références avant le stimulus (R1&R2).

#### ***P+SR VERSUS A+SR : CLICONME ET PIPE***

Pour les techniques sans référence, le fait qu'elles permettent de rechercher activement la valeur améliore la puissance d'impact. Cependant, les techniques « ClicOnMe » et « Pipe » ne suivent pas cette règle.

Pour la technique « ClicOnMe », l'utilisateur déclenche explicitement la recherche de la valeur et dispose d'un feedback (retour d'information) immédiat de son action par la perception kinesthésique de son mouvement. Mais il s'agit d'un contrôle isométrique : on

peut dire que l'utilisateur recherchait « statiquement » la valeur. Il semble donc apparaitre, ici, que la notion d'actif doit prendre en compte le caractère isométrique (par opposition à isotonique) de la technique. Une perspective intéressante, de cette étude de l'impact du caractère actif de l'implication de l'utilisateur, consisterait donc à distinguer les techniques actives isométriques et isotoniques.

La technique « Pipe » présente la plus faible puissance d'impact parmi toutes les techniques évaluées. Reprenons l'analyse de cette technique, nous pouvons voir qu'elle est sans référence, donc sa puissance d'impact est inférieure à toute technique en ayant au moins une et c'est le cas. Cette technique est active, du type d'utilisation IN/FEEDBACK et utilise la modalité tactile dans un signal continu avec le type de codage ANGLE, comme les techniques « PalmGauge », « Pédale » et « Volant », qui elles, sont loin devant en termes de classement par puissance d'impact. Ainsi, il semblerait que la combinaison de facteurs « technique active »  $\cap$  « type IN/FEEDBACK »  $\cap$  « modalité tactile »  $\cap$  « signal continu »  $\cap$  « type de codage ANGLE » ne soit pas la cause de la faible puissance d'impact de la technique « Pipe ». Toujours est-il, que parmi les facteurs de conception de cette technique n'ayant pas été traités, il reste la « zone du corps », l'« échelle » et le « dispositif ». Les zones du corps sollicitées sont les poignets et l'échelle vaut 90 ANGLES == 100 Valeurs, c'est-à-dire que l'utilisateur devait incliner le « pipe » de 9° pour atteindre la valeur 10. Or dans la technique « Pédale » le poignet est également sollicité et son échelle est de 22 ANGLES == 100 valeurs, c'est-à-dire que l'utilisateur devait incliner la « Pédale » de 2,2° pour atteindre 10. Cependant, cette technique est très bien classée, ce qui nous amène à penser que ces deux facteurs ne semblent pas pouvoir être, non plus, la cause. Il reste le « dispositif » lui-même, comme cause probable de sa faible puissance d'impact. Ainsi, selon notre cadre analytique, le « Pipe » affiche une mauvaise puissance d'impact car cette technique n'a aucune référence ou que le dispositif ne semble pas être adapté à une tâche d'estimation de valeur entre 1 et 100.

## 7.2 Résumé des résultats de la deuxième étude

Le chapitre 5, présente la deuxième étude et porte sur l'estimation et la mémorisation d'un itinéraire, c'est-à-dire une séquence de couples direction-distance. Cette étude a mis en œuvre trois facteurs spécifiques de conception (l'échelle, le type de transmission (l'aspect temporel) et le dispositif partagé) et l'utilisation de la parole, sur lesquels nous nous sommes basés pour concevoir quatre techniques d'interaction non visuelle. Ces techniques

ont été évaluées et comparées entre elles par quatre variables observées : le nombre d'itinéraires correctement restitués, le nombre d'itinéraires restitués de manière identique entre les deux restitutions d'un itinéraire, le score NASA-TLX et leur ordre de préférence.

Les résultats de cette expérience (exp. 2.1) ont confirmé que l'utilisation de la parole seule n'est pas très efficace, et ont montré que son usage, en complément de d'autres modalités, améliore l'efficacité et la satisfaction. Cependant, les résultats ont montré que les techniques utilisant la parole, exclusivement ou partiellement, ne sont pas robustes à une tâche d'interférence et ne facilite pas la rétention d'information. Ces deux résultats sont discutés dans les sections suivantes.

#### ***L'EFFICACITE DE LA PAROLE SEULE (TTS)***

Dans la littérature, nous avons vu que la modalité auditive « parole » utilisée dans des techniques d'interaction était efficace pour transmettre des informations de navigation, par exemple pour guider un utilisateur vers un point d'intérêt (Kane *et al.* 2011) ou énumérer une liste de points d'intérêt dans une rue (Panëels *et al.* 2013). Cependant, nous avons montré dans notre deuxième étude (exp. 2.1) que la parole seule n'est ni efficace ni satisfaisante pour transmettre un itinéraire, ou pour permettre sa mémorisation. Notons que les utilisateurs n'avaient droit qu'à un seul essai pour percevoir et retenir les informations. Il est évident qu'un nombre infini d'essais permettrait de mémoriser un itinéraire composé de cinq couples direction-distance. En effet, dès le premier essai, d'autres techniques, basées, notamment, sur les propriétés de la multimodalité, se sont montrées plus efficaces que la parole seule.

#### ***L'EFFICACITE DE LA PAROLE AVEC D'AUTRES MODALITES***

La technique « TTS-Lat. », qui transmet une direction et une distance en parallèle, affiche une efficacité de près de 44%, c'est-à-dire que les utilisateurs réussissaient à parfaitement restituer dix informations dans 44% des essais alors que la parole ne permettait que 6% d'efficacité. Ainsi, en spatialisant la parole, nous avons amélioré les performances d'un facteur sept. Nous avons également évalué la technique « PG-TTS », qui transmet une direction avec la modalité tactile puis la distance avec la parole : dans ce cas, les performances sont améliorées d'un facteur cinq.

#### ***LA ROBUSTESSE DE LA MELODIE, TECHNIQUE INSPIREE DU JEU SIMON***

Nous avons conçu et évalué la technique « Mélodie » permettant de faciliter la mémorisation. Nous sommes partis d'un jeu de mémorisation des années 80, basé sur le

son et la lumière, le *Simon*<sup>52</sup>, pour concevoir un système basé uniquement sur du son et capable de transmettre une séquence de couple d'informations. Cependant, dans le jeu *Simon* il s'agit d'une séquence de note, de couleur et de position, et dans notre technique « Mélodie » c'est une séquence de notes spatialisées d'une certaine durée transmettant un itinéraire (séquence de couples direction-distance). Le résultat de cette technique améliore les performances de la robustesse de la mémorisation d'un facteur 4 comparativement à notre technique témoin, qu'est la parole seule mais aussi, comparativement aux techniques utilisant la parole avec une autre modalité (« TTS-Lat » et « PG-TTS »). Notons toutefois, que les utilisateurs mémorisaient des informations de distance trop erratiques pour que cette technique soit efficace. Néanmoins, cette technique sans parole démontre que l'utilisation de la parole (seule ou non) limite fortement la robustesse de la mémorisation d'un grand nombre d'informations. De plus, il est possible que la modalité de restitution (DUREE d'un son) induise un biais dans les réponses relatives aux « distances » pour cette technique car c'est la seule qui n'utilisait pas la modalité parole pour transmettre les distances. Enfin, en ne considérant que les directions, les excellentes performances de Mélodie montrent que l'utilisation synergique et redondante de plusieurs modalités pour encoder une information, favorise la mémorisation de celles-ci.

### 7.3 Résumé des résultats de la troisième étude

Le chapitre 6, présente la troisième étude et porte sur la localisation d'une cible dans une portion d'espace à partir de la transmission d'un couple de coordonnées avec des stimuli vibrotactiles. Cette étude a mis en œuvre deux types de codage (DUREE vs. PULSE) dans deux systèmes de coordonnées différents (Cartésien vs. Polaire). Après avoir évalué et comparé ces quatre techniques (exp. 3.1), nous avons adapté la meilleure des techniques et l'avons évaluée et comparée à une technique de contrôle ; c'est-à-dire que nous avons comparé une technique d'assistance avec indice d'anticipation à une technique sans indice avec des sujets voyants (exp. 3.2) et un sujet déficient visuel (exp. 3.3). Ces techniques ont été évaluées et comparées entre elles par le biais de quatre variables observées : l'erreur de pointage, le score SUS, la préférence et le temps de réaction (exp. 3.2 et 3.3 uniquement).

Les résultats de cette étude (exp. 3.1, 3.2 et 3.3) ont montré que notre technique « train de pulsations transmis dans un système de coordonnées Cartésien » est plus efficace et plus satisfaisante que les autres. Dans les exp. 3.2 et 3.3, la technique d'assistance a permis de

<sup>52</sup> <http://www.1980-games.com/jeux-reflexion/jeux-flash/simon.php>

réduire le temps de recherche et de pointage de la cible, d'un facteur trois comparativement à la technique sans indice. Toutefois, les résultats de l'exp. 3.1 montrent que les techniques basées sur le type de codage DUREE sont moins efficaces et moins préférées que celles basées sur le type PULSE. Nous approfondirons la discussion de ce résultat dans la section suivante.

#### ***DIFFERENCE ENTRE LES TYPES DE CODAGE DUREE ET PULSE***

Dans notre troisième étude (exp. 3.1), nous avons comparé deux types de codage différents : DUREE et PULSE. Les résultats ont montré que le type PULSE permettait une meilleure efficacité et satisfaisait plus les utilisateurs que le type DUREE. Pourtant, le type DUREE permet une précision théorique immensément plus grande que le type PULSE. En effet, le type PULSE, ne permet qu'une précision théorique de 9° tandis que le type DUREE offre une précision théorique de 0.03°. En d'autres termes, il y a cent-vingt-une positions possibles avec le type PULSE et plus de deux millions avec le type DUREE, soit une résolution presque 18 000 fois plus de fine avec le type DUREE. Les utilisateurs ont reconnu le potentiel de précision du type DUREE, mais étaient incapables d'être surs de leurs réponses (contrairement avec le type de codage PULSE). Néanmoins, ce résultat est cohérent avec ceux de l'exp. 1.1. En effet, la technique C/Si (Cartésien-SingleStimulus) est basée sur la technique SingleStimulus, qui est classée 10<sup>ième</sup> en termes de puissance d'impact pour estimer une valeur entre 1 et 100 (cf. Tableau 4-12). Pourtant, nous avons montré qu'il est possible d'estimer précisément une valeur entre 1 et 100 avec le type de codage DUREE pour la technique « BipMultiple». Afin de vérifier si le type de codage PULSE est vraiment supérieur au type de codage DUREE, nous devrions concevoir une technique vibrotactile équivalente (cf. Conclusion de la première étude) à la technique « BipMultiple» pour transmettre une paire de coordonnées d'une cible, l'évaluer dans un tâche similaire à celle de l'exp. 3.1 et comparer les résultats avec ceux obtenus pour C/Tr (cartésien-Train de pulsation) dans l'exp. 3.1.

## **7.4 Conclusion générale**

Ces travaux de thèse, nous ont amenés à construire un cadre d'analyse regroupant les facteurs de conception communs des techniques d'interaction présentées dans notre état de l'art. A partir de ce cadre nous avons conçu, évalué et comparé près d'une trentaine de techniques dans trois études distinctes.

#### ***TROIS ETUDES ...***

Dans la première étude (exp. 1.1 et exp. 1.2), dix-neuf techniques ont été évaluées et comparées dans une tâche d'estimation d'une seule valeur entre 1 et 100. Toutefois, les résultats de l'exp. 1.1 ont montré que les techniques alternatives (version vibrotactile) étaient équivalentes à leur technique de base (version auditive) : les utilisateurs ont donc le choix entre les modalités auditive ou vibrotactile pour réaliser une tâche d'estimation de valeur entre 1 et 100 et peuvent ainsi choisir la modalité sensorielle plus adéquate en fonction du contexte. Cette première étude a permis de valider l'hypothèse qu'être actif pendant une transmission d'informations (pendant la tâche), était plus efficace et plus satisfaisant qu'être passif. Elle a aussi permis de valider l'hypothèse selon laquelle percevoir des références explicites, juste avant, après ou pendant la transmission d'une information, améliore l'estimation de cette information : par exemple, « BipMultiple » est 81% plus efficace que « SingleStimulus ». Nous avons également observé que chacune des techniques sans référence, générait une dérive (significative) de l'erreur au cours du temps. En effet, à partir du moment où la phase de familiarisation était finie, les estimations des participants étaient de plus en plus fausses. Ce phénomène n'a pas été observé pour les techniques avec au moins une référence. Enfin, les résultats de cette première étude ont permis d'établir les relations suivantes entre deux facteurs de conception **[Actif > Passif (A>P)]** et **[AvecRéférences > SansRéférence (AR>SR)]** et leurs quatre combinaisons **[A+AR > P+AR > A+SR > P+SR]** (en termes de puissance d'impact).

Dans la deuxième étude (exp. 2.1), quatre techniques ont été évaluées et comparées dans une tâche d'estimation et de mémorisation d'un itinéraire composé d'une séquence de couples direction-distance. Les résultats de cette étude ont montré les limites de l'utilisation de la parole, mais aussi les avantages liés à des interactions multimodales synergiques (parallèles et complémentaires). En effet, la parole seule s'est montrée ni efficace ni satisfaisante pour transmettre un itinéraire, ni pour permettre sa mémorisation. Toutefois, deux techniques multimodales utilisant en partie la parole se sont montrées efficaces pour transmettre un itinéraire. TTS-Lat (parole spatialisée) et PG-TTS (angle tactile puis parole), ont permis d'optimiser la transmission d'un itinéraire et ont facilité la mémorisation à court terme de 10 informations dans 44% des cas pour TTS-Lat et 31% des cas pour PG-TTS. Pour rappel, notre technique témoin (TTS, parole seule) n'a permis de mémoriser 10 informations que dans 6% des cas seulement. Ainsi, en spatialisant la parole (TTS-Lat), nous avons amélioré les performances d'un facteur sept et en répartissant les informations sur deux canaux sensoriels, tactile et auditif (PG-TTS), nous avons amélioré les performances d'un

facteur cinq. La « Mélodie » est notre technique basée uniquement sur du son sans parole et plus précisément, sur une séquence de notes spatialisées. Cette technique s'est montrée plus robuste d'un facteur quatre, comparativement à toutes les autres techniques utilisant, exclusivement ou en partie, la modalité de la parole. Ce résultat démontre que l'utilisation de la parole (seule ou non) limite fortement la robustesse de la mémorisation d'un grand nombre d'informations avec seulement un seul essai.

Dans notre troisième étude, quatre techniques ont été évaluées et comparées dans une tâche de localisation d'une cible à partir d'une paire de coordonnées (exp. 3.1). Les résultats de l'exp. 3.1 ont montré que la meilleure technique parmi les quatre, était celle basée sur le système de coordonnées cartésien et le type de codage PULSE, c'est-à-dire que les participants localisaient les cibles selon deux axes (un vertical et un horizontal), selon une valeur de -5 à 5 pour chaque axe, soit une grille d'une résolution de 11 x 11. Dans l'exp. 3.2 et 3.3, nous avons repris cette technique d'assistance, puis nous l'avons évaluée et comparée à une technique témoin dans une tâche de recherche visuelle d'une cible dans un salon surchargé visuellement. Ceci par un groupe de participants voyants (exp. 3.2) puis par un participant présentant une vision tubulaire (exp. 3.3). Les résultats de ces deux expériences (3.2 et 3.3) ont montré que la technique d'assistance réduisait le temps de recherche d'un facteur trois comparativement à la technique témoin. Ce dispositif, un bracelet composé de quatre moteurs vibrants, pourrait être facilement intégré dans des Smartwatch et détecter des cibles d'intérêt, soit sur demande soit de façon automatique (par exemple les obstacles potentiels), facilitant ainsi la recherche visuelle et la perception spatiale de l'environnement.

#### *... ET UN CADRE D'ANALYSE ...*

Le cadre d'analyse dont nous nous sommes servi pour concevoir nos techniques, ne concerne que les techniques permettant de transmettre une information. Toutefois, nous avons étudié l'impact de certains facteurs et paramètres de conception sur l'efficacité des techniques et sur la satisfaction des utilisateurs.

Dans les expériences 1.1 et 1.2, nous avons démontré que le facteur AvecRéférences améliore la confiance des utilisateurs et les estimations des informations et que le facteur techniques Actives améliore légèrement les estimations des techniques. Nous avons également observé que le type de codage LONGUEUR favorise une meilleure estimation d'une valeur entre 1 et 100 comparativement à des techniques similaires (nombre de

références similaires, actives, mais type de codage différent) et améliore la confiance des utilisateurs dans leurs réponses. Nous avons constaté que dans des conditions bien précises (exp. 1.1), la modalité vibrotactile est équivalente à la modalité sonification (auditive sans parole).

Dans l'expérience 2.1, nous avons démontré que la parole (type de codage LANGAGE) seule était inefficace et insatisfaisante dans une tâche de mémorisation d'une séquence de couple direction-distance à un seul essai (l'information est transmise une seule et unique fois). Cependant, l'utilisation des différents types d'aspect temporel d'une transmission d'information (types de transmission), basés sur les propriétés et types de la multimodalité, a permis d'optimiser des techniques d'interaction utilisant en partie la parole. L'utilisation synergique et complémentaire de la parole et de sa spatialisation ainsi que, l'utilisation alternée et complémentaire d'un angle tactile et de la parole, ont permis d'améliorer l'efficacité à mémoriser un itinéraire, soit 10 éléments, d'un facteur sept (TTS-Lat) et cinq (PG-TTS), comparé à une utilisation exclusive de la parole. De plus, notre technique conçue pour mémoriser de manière robuste (Mélodie), s'est montrée très robuste pour mémoriser des itinéraires de taille cinq, trois itinéraires sur quatre sont mémorisés de manière robuste, et excellentement robuste pour les directions (100% en taille 5 et 75% en taille 7). Ces résultats démontrent que l'utilisation de la parole limite la robustesse de la mémorisation d'un grand nombre d'informations et que la redondance des modalités améliore la robustesse.

Dans les expériences 3.1, 3.2 et 3.3, nous avons observé que le type de codage PULSE améliore la précision et facilite l'intégration d'une paire de coordonnées d'une cible comparativement au type de codage DUREE. Cependant, la résolution de la grille encodée par notre technique C/Tr (Cartésien-Train de pulsations) est 18 000 fois inférieure à celle de la grille encodée par notre technique C/Si (Cartésien-SingleStimulus). Nous avons vu que ce résultat est cohérent avec ceux de l'exp. 1.1 (cf. Différence entre les types de codage DUREE et PULSE). Cependant, nous pensons que la technique C/Si pourrait-être améliorée en se basant sur la technique « BipMultiple ».

**... POUR CONCEVOIR ET EVALUER DE TECHNIQUES D'INTERACTION NON VISUELLE OPTIMISEES POUR DE LA TRANSMISSION D'INFORMATION(S).**

Nous avons conçu, évalué et comparé près d'une trentaine de techniques permettant de transmettre trois types d'information : <valeur>, <position> et <itinéraire>. Nous avons vu quels facteurs de conception améliorent ou facilitent, la perception d'une information et

peuvent satisfaire les utilisateurs. Nous avons vu également que la redondance de plusieurs modalités pour une même information améliore la robustesse de la mémorisation d'une information. Enfin, l'équivalence des modalités auditives (sonifiées) et vibrotactiles ajoute de la flexibilité dans la conception. Nous avons vu que certains aspects temporels, <synergique et complémentaire> et <alternée et complémentaire>, issus des propriétés <CASE et CARE> de la multimodalité améliorent l'efficacité de la mémorisation d'un grand nombre d'informations. Enfin, nous avons vu que transmettre une information en se basant sur le type de codage PULSE, donne confiance aux utilisateurs et assure une bonne efficacité des techniques. Nous pouvons également faire cette remarque pour les techniques avec un grand nombre de références (Règle 1, cf. Résumé des résultats de la première étude) ou pour les techniques permettant aux utilisateurs d'explorer activement les références (avant ou pendant la tâche) (Règle 2, cf. Résumé des résultats de la première étude).

*In fine*, afin de concevoir et d'évaluer des techniques d'interaction non visuelle optimisées pour la transmission d'information, dans des situations où la perception visuelle est fortement contrainte ou déficiente, il est nécessaire de rendre perceptible l'information sous une forme non visuelle. Ceci en prenant en compte les capacités sensorielles et mnésiques humaines et en considérant certains facteurs de conception, tels que le nombre de références, l'implication de l'utilisateur, le type de codage des informations, le type de transmission, la zone du corps sollicitée et le dispositif.

## 7.5 Perspectives

Le travail présenté dans cette thèse ouvre plusieurs perspectives.

Une première serait de compléter l'étude de l'impact des facteurs non traités ici. Prenons l'exemple des zones du corps : peut-on estimer une valeur entre 1 et 100 avec l'inclinaison d'un doigt ou d'une pression sur ce doigt ? Nous pourrions notamment nous appuyer sur les travaux de (Hajnal *et al.* 2011) pour concevoir un dispositif utilisant un PHANTOM capable de transmettre une valeur entre 1 et 100 avec une inclinaison (ANGLE) ou une pression (FORCE) avec la modalité d'interaction haptique RETOUR DE FORCE. Nous pourrions le faire, d'une manière similaire, pour les pieds ou les jambes, cependant, pour le coude nous nous inspirerions des travaux de (Enokibori and Mase 2014) et de leur dispositif de tissu « intelligent ». Notons que ceux de (Wagner *et al.* 2013) portent sur l'efficacité de

techniques d'interaction en fonction de la latéralisation des participants. Ainsi, nous devrions également faire une étude sur l'impact de la latéralisation des participants sur les résultats.

Un second exemple de facteur de conception à étudier, serait de comparer plus assidument les différents types de codage : par exemple, est-il plus efficace d'estimer une valeur entre 1 et 100 avec un doigt et avec le type de codage ANGLE (par inclinaison, par rotation, etc.), LONGUEUR (ex. par translation sur écran tactile), DUREE (ex. par vibration), FORCE (ex. par pression), PULSE (ex. nombre de tape sur un bouton) ou un langage (ex. dessiner un nombre) ? Nous pourrions faire de même pour chacun des doigts de la main gauche puis ceux de la main droite mais aussi pour toutes les zones du corps.

Une deuxième perspective serait de décliner les types d'informations à transmettre à partir des trois étudiés dans le cadre de cette thèse : <valeur>, <position>, et <itinéraire>. Nous avons vu qu'une valeur pouvait être une direction, une distance ou une coordonnée mais, elle pourrait être également : un numéro de département ou de salle de cours, une heure, un poids, une taille, etc. Toutefois, de nombreux d'exemples existent déjà pour transmettre une valeur avec une modalité non visuelle et sans utiliser la parole : par exemple, les indications d'une faible charge de batterie d'un Smartphone par des vibrations ou du son, l'heure par un affichage en relief (aiguille à nues, montres brailles<sup>53</sup>, etc.) sans parler de toutes les techniques présentées dans notre état de l'art. Cependant, en ce qui concerne les informations de type <couple de valeurs>, nous avons vu comment transmettre une position d'une cible (cf. Troisième étude : assistance tactile à la localisation de cibles périphériques pour des personnes à vision tubulaire), mais nous pourrions explorer de nouvelles interactions pour transmettre des informations par couple. Par exemple, en situation d'évacuation, suite à un incendie de forêt, la transmission d'un couple <Direction ; Niveau-de-Danger> pourrait faciliter le choix d'un utilisateur à prendre la bonne direction pour échapper aux flammes ou rejoindre l'abri le plus sûr. Un autre exemple, serait de transmettre le couple <Propriété ; Mesure> pour analyser un fruit chez le primeur (ex. <Pomme ; 300g>) ou pour analyser une zone en contexte de navigation (ex. <Trottoir gauche ; 120 cm de large>). Enfin, pour les informations de type <itinéraire>, c'est-à-dire une séquence de couple <direction-distance>, nous pourrions imaginer de décliner ce

---

<sup>53</sup> Voir <http://www.dotincorp.com/>

type d'informations pour transmettre un graphique sous la forme d'une séquence de couples <propriété ; mesure>. Par exemple, pour un histogramme illustrant le nombre d'inscrits dans une catégorie : <Sport ; 45> ; <Musique ; 6> ; <Anglais ; 25> ; <Dessin ; 12> ; <Jeux de société ; 12>. Un autre exemple serait de transmettre une séquence de couples <mesure ; mesure> pour transmettre la représentation graphique d'une fonction mathématique. Pour ce faire, nous concevrons une technique d'interaction inspirée des travaux de (Marzo *et al.* 2015) pour créer un « holographique acoustique » afin que l'utilisateur puisse percevoir les valeurs sur une échelle continue ( $f(x)=ax+b$ , où  $a,b,x \in \mathbb{R}$ ).

Enfin, une troisième perspective serait de développer concrètement une plateforme opérationnelle permettant le prototypage rapide de nouvelles techniques d'interaction non visuelle optimisées pour de la transmission d'information. Il conviendrait en effet de réaliser le développement concret, complet et robuste d'une application permettant d'interconnecter des dispositifs nouveaux ou déjà existants dans une base de données, à des systèmes interactifs conçus pour transmettre ou recevoir des informations par la génération de différents types de signaux pendant une tâche utilisateur. Le verrou technologique de ce projet se situe probablement au niveau des interconnexions car cette plateforme devra permettre de réaliser facilement le couplage entre les dispositifs et l'information à transmettre mais aussi, entre les dispositifs eux-mêmes. L'apport indéniable d'une telle plateforme sera double : faciliter le prototypage de ces techniques d'interaction non visuelle optimisées et constituer un support à une exploration systématique des différents attributs d'une technique d'interaction non visuelle identifiés dans notre cadre d'analyse. Pour y parvenir, nous pourrions nous appuyer sur les travaux de cette thèse bien sûr, mais aussi sur le projet WComp<sup>54</sup> (Ferry *et al.* 2011; Weibel *et al.* 2006) qui permet de faire du prototypage rapide avec une approche par composant et sur les travaux de (Dubois *et al.* 2014) qui ont développé un modèle de conception pour les systèmes interactifs mixtes (SIM).

---

<sup>54</sup> <http://www.wcomp.fr/wikiwcomp>



# **Partie 8 :** **Bibliographie**



## 8 Bibliographie

1. Agrawal, Pyush et al. 2004. "Multimodal Interface Platform for Geographical Information Systems (GeoMIP) in Crisis Management." Pp. 339–40 in *Proceedings of the 6th International Conference on Multimodal Interfaces, ICMI '04*. New York, NY, USA: ACM. Retrieved (<http://doi.acm.org/10.1145/1027933.1027997>).
2. Altenhoff, Bliss M. et al. 2012. "Effects of Calibration to Visual and Haptic Feedback on near-Field Depth Perception in an Immersive Virtual Environment." Pp. 71–78 in *Proceedings of the ACM Symposium on Applied Perception - SAP '12*, vol. 1. New York, USA: ACM Press. Retrieved (<http://dl.acm.org/citation.cfm?doid=2338676.2338691>).
3. Appelle, Stuart. 1991. "Haptic Perception of Form: Activity and Stimulus Attributes." in *The Psychology of Touch*, ch7. Morton A. Heller, William Schiff.
4. Auvray, Malika, Sylvain Hanneton, and J. Kevin O'Regan. 2007. "Learning to Perceive with a Visuo – Auditory Substitution System: Localisation and Object Recognition with 'The vOICe.'" *Perception* 36(3):416–30. Retrieved (<http://www.perceptionweb.com/abstract.cgi?id=p5631>).
5. Bach-y-Rita, P., K. A. Kaczmarek, M. E. Tyler, and J. Garcia-Lara. 1998. "Form Perception with a 49-Point Electrotactile Stimulus Array on the Tongue: A Technical Note." *Journal of rehabilitation research and development* 35(4):427–30.
6. Bach-Y-Rita, Paul, Carter C. Collins, Frank A. Saunders, Benjamin White, and Lawrence Scadden. 1969. "Vision Substitution by Tactile Image Projection." *Nature* 221(5184):963–64.
7. Bach-y-Rita, Paul, Kurt A. Kaczmarek, M. E. Tyler, and J. Garcia-Lara. 1998. "Form Perception with a 49-Point Electrotactile Stimulus Array on the Tongue: A Technical Note." *Journal of rehabilitation research and development* 35(4):427–30. Retrieved (<http://www.ncbi.nlm.nih.gov/pubmed/10220221>).
8. Bangor, Aaron, Philip T. Kortum, and James T. Miller. 2008. "An Empirical Evaluation of the System Usability Scale." Pp. 574–94 in *International Journal of Human-Computer Interaction*, vol. 24. Retrieved (<http://www.tandfonline.com/doi/abs/10.1080/10447310802205776#.VTXGZPvzj8>).
9. Barrass, Stephen and Gregory Kramer. 1999. "Using Sonification." *Multimedia Systems - Special issue on audio and multimedia* 7(1):23–31. Retrieved (<http://link.springer.com/10.1007/s005300050108>).
10. Barrass, Stephen and Paul Vickers. 2011. "Chapiter 7 : Sonification Design and Aesthetics." Pp. 145–71 ; p.149 in *The Sonification Handbook*, edited by J. G. Hermann, T., Hunt, A., and Neuhoff.
11. Bastien, J. M. Christian and Dominique L. Scapin. 1992. "A Validation of Ergonomic Criteria for the Evaluation of Human-computer Interfaces." *International Journal of Human-Computer Interaction* 4(2):183–96.
12. Bellik, Y. and D. Teil. 1992a. "Définitions Terminologiques Pour La Communication Multimodale." in *IHM'92, Quatrièmes Journées sur l'Ingénierie des Interfaces Homme-Machine*.
13. Bellik, Y. and D. Teil. 1992b. "Les Types de Multimodalités." in *IHM'92, Quatrièmes Journées sur l'Ingénierie des Interfaces Homme-Machine*.
14. Blattner, Meera, Denise Sumikawa, and Robert Greenberg. 1989. "Earcons and Icons: Their Structure and Common Design Principles." *Human-Computer Interaction* 4(1):11–44. Retrieved ([http://www.tandfonline.com/doi/abs/10.1207/s15327051hci0401\\_1](http://www.tandfonline.com/doi/abs/10.1207/s15327051hci0401_1)).
15. Bolelli, Levent et al. 2004. "Multimodal Interaction for Distributed Collaboration." Pp. 327–28 in *Proceedings of the 6th International Conference on Multimodal Interfaces, ICMI '04*. New York, NY, USA: ACM. Retrieved (<http://doi.acm.org/10.1145/1027933.1027990>).
16. Bolt, Richard a. 1980. "'Put-That-There': Voice and Gesture at the Graphics Interface." *Proceedings of the 7th annual conference on Computer graphics and interactive techniques - SIGGRAPH '80* 262–70.
17. Bourdon, B. 1914. "Localisation En Avant, En Arrière, Entre Avant et Arrière." *L'année psychologique* 21(1):86–89. Retrieved ([http://www.persee.fr/web/revues/home/prescript/article/psy\\_0003-5033\\_1914\\_num\\_21\\_1\\_8081](http://www.persee.fr/web/revues/home/prescript/article/psy_0003-5033_1914_num_21_1_8081)).
18. Brewster, Stephen S. a. and L. M. Lorna M. Brown. 2004. "Tactons: Structured Tactile Messages for Non-Visual Information Display." Pp. 15–23 in *Proceedings of the fifth conference on Australasian user interface (AUIC '04)*, vol. 28, edited by A. Cockburn. Darlinghurst, Australia, Australia: Australian Computer Society, Inc. Retrieved (<http://crpit.com/Vol28.html>).

19. Brock, Anke. 2013. "Touch the Map!: Designing Interactive Maps for Visually Impaired People." *SIGACCESS Access. Comput.* 105, January, 9–14. Retrieved (<http://dl.acm.org/citation.cfm?id=2444800.2444802>).
20. Brock, Anke, Slim Kammoun, Marc Macé, and Christophe Jouffrais. 2014. "Using Wrist Vibrations to Guide Hand Movement and Whole Body Navigation." *I-Com* 13(3):19–28.
21. Brock, Anke M., Philippe Truillet, Bernard Oriola, Delphine Picard, and Christophe Jouffrais. 2015. "Interactivity Improves Usability of Geographic Maps for Visually Impaired People." *Human-Computer Interaction* 30(2):156–94. Retrieved (<http://www.tandfonline.com/doi/abs/10.1080/07370024.2014.924412>).
22. Brock, Derek, James a. Ballas, and Daniel C. McFarlane. 2005. "Encoding Urgency in Legacy Audio Alerting Systems." *Proceedings of the 11th international Conference on Auditory Display* 190–97.
23. Brooks, Daniel I., Ian P. Rasmussen, and Andrew Hollingworth. 2010. "The Nesting of Search Contexts within Natural Scenes: Evidence from Contextual Cuing." *Journal of experimental psychology. Human perception and performance* 36(6):1406–18.
24. Brunet, Lucie et al. 2013. "'Invitation to the Voyage': The Design of Tactile Metaphors to Fulfill Occasional Travelers' Needs in Transportation Networks." Pp. 259–64 in *2013 World Haptics Conference, WHC 2013*. IEEE. Retrieved (<ftp://158.130.73.148/stuff/WHC13-usb/papers/O3/brunet.pdf>).
25. Buchs, Galit, Shachar Maidenbaum, and Amir Amedi. 2015. "Augmented Non-Visual Distance Sensing with the EyeCane." Pp. 209–10 in *Proceedings of the 6th Augmented Human International Conference on - AH '15*. New York, New York, USA: ACM Press. Retrieved (<http://dl.acm.org/citation.cfm?doid=2735711.2735780>).
26. Capelle, Christian, Charles Trullemans, Patricia Arno, and Claude Veraart. 1998. "A Real-Time Experimental Prototype for Enhancement of Vision Rehabilitation Using Auditory Substitution." *IEEE Transactions on Biomedical Engineering* 45(10):1279–93.
27. Cassell, J. et al. 1999. "Embodiment in Conversational Interfaces: Rea." Pp. 520–27 in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '99*. New York, NY, USA: ACM. Retrieved (<http://doi.acm.org/10.1145/302979.303150>).
28. Cassinelli, Alvaro, Carson Reynolds, and Masatoshi Ishikawa. 2006. "Augmenting Spatial Awareness with Haptic Radar." Pp. 61–64 in *2006 10th IEEE International Symposium on Wearable Computers*. IEEE.
29. Chatillon, J. 2006. "Limites d'Exposition Aux Infrasons et Aux Ultrasons. Étude Bibliographique." *Revue Hygiène et Sécurité Du Travail -- Référence INRS ND 2250*, 67–77.
30. Chen, HY, Joseph Santos, Matthew Graves, Kwangtaek Kim, and Hong Z. Tan. 2008. "Tactor Localization at the Wrist." *Proceedings of the EuroHaptics 2008 Conference* 209–18. Retrieved February 10, 2014 ([http://link.springer.com/chapter/10.1007/978-3-540-69057-3\\_25](http://link.springer.com/chapter/10.1007/978-3-540-69057-3_25)).
31. Chevrin, Vincent. 2004. "Les Interactions Multimodales et Multicanal Dans UMR." P. 4 in *RJC IHM*.
32. Choi, Sam-ha, Hee-Dong Chang, and Kyung-Sik Kim. 2004. "Development of Force-Feedback Device for PC-Game Using Vibration." *ACM SIGCHI International Conference on Advances in computer entertainment technology - ACE '04* 325–30. Retrieved (<http://portal.acm.org/citation.cfm?doid=1067343.1067391>).
33. Coleman, Paul D. 1963. "An Analysis of Cues to Auditory Depth Perception in Free Space." *Psychological Bulletin* 60(3):302–15. Retrieved (<http://content.apa.org/journals/bul/60/3/302>).
34. Coutaz, J., L. Nigay, Daniel Salber, and Ann Blandford. 1995. "Four Easy Pieces for Assessing the Usability of Multimodal Interaction: The CARE Properties." *InterAct* (June):1–7. Retrieved (<http://danielsalber.com/publications/interact1995.pdf>).
35. Coutaz, Joëlle and Laurence Nigay. 1994. "Les Propriétés 'CARE' Dans Les Interfaces Multimodales." Pp. 7–14 in *IHM'94*.
36. Cremetz, Olivier. 2007. "Rapport Soutenance Du 23 Février 2007 : MODÉLISATION DES MÉCANISMES D'ADAPTATION DYNAMIQUE ET CONTEXTUELLE D'UNE INTERACTION MULTIMODALE." *Polytech'Lille* 15.

37. Csapó, Ádám and György Wersényi. 2013. "Overview of Auditory Representations in Human-Machine Interfaces." *ACM Comput. Surv.* 46(2):19:1–19:23. Retrieved (<http://doi.acm.org/10.1145/2543581.2543586>).
38. Dean, Thomas et al. 2013. "Fast, Accurate Detection of 100,000 Object Classes on a Single Machine." Pp. 1814–21 in *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*. Retrieved (<http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6619081\npapers3://publication/doi/10.1109/CVPR.2013.237>).
39. Dingler, Tilman, Jeffrey Lindsay, and Bruce N. Walker. 2008. "Learnability of Sound Cues for Environmental Features: Auditory Icons, Earcons, Spearcons, and Speech." Pp. 1–6 in *14th International Conference on Auditory Display*. Paris, France.
40. Dramas, Florian, Bernard Oriola, Brian F. G. Katz, Simon J. Thorpe, and Christophe Jouffrais. 2008. "Designing an Assistive Device for the Blind Based on Object Localization and Augmented Auditory Reality." Pp. 263–64 in *Proceedings of the 10th international ACM SIGACCESS conference on Computers and accessibility - Assets '08*. Retrieved ([http://portal.acm.org/citation.cfm?id=1414529\nhttp://portal.acm.org/ft\\_gateway.cfm?id=1414529&type=pdf&coll=GUIDE&dl=GUIDE&CFID=55075414&CFTOKEN=82534184\nhttp://doi.acm.org/10.1145/1414471.1414529](http://portal.acm.org/citation.cfm?id=1414529\nhttp://portal.acm.org/ft_gateway.cfm?id=1414529&type=pdf&coll=GUIDE&dl=GUIDE&CFID=55075414&CFTOKEN=82534184\nhttp://doi.acm.org/10.1145/1414471.1414529)).
41. Dubois, Emmanuel, Christophe Bortolaso, Damien Appert, and Guillaume Gauffre. 2014. "An MDE-Based Framework to Support the Development of Mixed Interactive Systems." *Science of Computer Programming* 89(PB):199–221. Retrieved (<http://dx.doi.org/10.1016/j.scico.2013.03.007>).
42. Dubois, Véronique, Valérie Gyselinck, and Hugues Choplin. 2003. "Multimodalité et Mémoire de Travail." Pp. 187–98 in *Environnements informatiques pour l'apprentissage humain, Strasbourg*.
43. Edworthy, J., S. Loxley, and I. Dennis. 1991. "Improving Auditory Warning Design: Relationship between Warning Sound Parameters and Perceived Urgency." *Human factors* 33(2):205–31.
44. Ellard, Colin G. and Sarah C. Shaughnessy. 2003. "A Comparison of Visual and Nonvisual Sensory Inputs to Walked Distance in a Blind-Walking Task." *Perception* 32(5):567–78.
45. Enokibori, Yu and Kenji Mase. 2014. "Human Joint Angle Estimation with an E-Textile Sensor." Pp. 129–30 in *Proceedings of the 2014 ACM International Symposium on Wearable Computers - ISWC '14, ISWC '14*. New York, USA: ACM Press. Retrieved (<http://dl.acm.org/citation.cfm?doid=2634317.2634331>).
46. Van Erp, Jan B. F., Hendrik A. H. C. Van Veen, Chris Jansen, and Trevor Dobbins. 2005. "Waypoint Navigation with a Vibrotactile Waist Belt." *ACM Transactions on Applied Perception* 2(2):106–17.
47. Van Erp, Jan. 2005. "Presenting Directions with a Vibrotactile Torso Display." *Ergonomics* 48(3):302–13.
48. Etter, R. and M. Specht. 2005. "Melodious Walkabout: Implicit Navigation with Contextualized Personal Audio Contents." *Ext Proc. Pervasive 2005* 43–49.
49. Farcy, René et al. 2006. "Electronic Travel Aids and Electronic Orientation Aids for Blind People: Technical, Rehabilitation and Everyday Life Points of View." P. 12 in *Conference & Workshop on Assistive Technologies for People with Vision & Hearing Impairments (CVHI 2006)*, edited by M.A. Hersh.
50. Ferry, Nicolas et al. 2011. "WComp, a Middleware for Ubiquitous Computing." in *Ubiquitous Computing*. InTech. Retrieved (<http://www.intechopen.com/books/ubiquitous-computing/wcomp-a-middleware-for-ubiquitous-computing>).
51. Fiannaca, Alexander, Tony Morelli, and Eelke Folmer. 2013. "Haptic Target Acquisition to Enable Spatial Gestures in Nonvisual Displays." *Proceedings of Graphics Interface 2013* 213–19. Retrieved December 1, 2014 (<http://dl.acm.org/citation.cfm?id=2532165>).
52. Folmer, Eelke and Tony Morelli. 2012. "Spatial Gestures Using a Tactile-Proprioceptive Display." *Proceedings of the Sixth International Conference on Tangible, Embedded and Embodied Interaction* 1:139–42. Retrieved (<http://dl.acm.org/citation.cfm?id=2148161>).
53. Fontana, Federico and Davide Rocchesso. 2008. "Auditory Distance Perception in an Acoustic Pipe." *ACM Transactions on Applied Perception (TAP)* 5(3):16:1–16:15. Retrieved (<http://portal.acm.org/citation.cfm?doid=1402236.1402240>).

54. Fortenbaugh, Francesca C., John C. Hicks, Lei Hao, and Kathleen a Turano. 2007. "Losing Sight of the Bigger Picture: Peripheral Field Loss Compresses Representations of Space." *Vision research* 47(19):2506–20.
55. Fortenbaugh, Francesca C., John C. Hicks, and Kathleen a Turano. 2008. "The Effect of Peripheral Visual Field Loss on Representations of Space: Evidence for Distortion and Adaptation." *Investigative ophthalmology & visual science* 49(6):2765–72.
56. Fox, Charles R. 2000. "Low Vision Rehabilitation for Patients with Severe Glaucoma." Pp. 3031–43 in *Principles and Practice of Ophthalmology*.
57. Fraisse, P. 1984. "Perception and Estimation of Time." *Annual review of psychology* 35:1–36.
58. Froehlich, P. and F. Hammer. 2004. "Expressive Text-to-Speech: A User-Centred Approach to Sound Design in Voiceenabled Mobile Applications." P. 4 in *Proc. Second Symposium on Sound Design*. Paris, France.
59. Fujimoto, Emily and Matthew Turk. 2014. "Non-Visual Navigation Using Combined Audio Music and Haptic Cues." Pp. 411–18 in *Proceedings of the 16th International Conference on Multimodal Interaction - ICMI '14*. New York, USA: ACM Press. Retrieved (<http://dl.acm.org/citation.cfm?id=2663243>).
60. Garzonis, Stavros, Christopher Bevan, and Eamonn O'Neill. 2008. "Mobile Service Audio Notifications: Intuitive Semantics and Noises." Pp. 156–63 in *Proceedings of the 20th Australasian Conference on Computer-Human Interaction: Designing for Habitus and Habitat, OZCHI '08*. New York, NY, USA: ACM. Retrieved (<http://opus.bath.ac.uk/30341/>).
61. Gaver, William. 1986. "Auditory Icons: Using Sound in Computer Interfaces." *Human-Computer Interaction* 2(2):167–77. Retrieved ([http://www.tandfonline.com/doi/abs/10.1207/s15327051hci0202\\_3](http://www.tandfonline.com/doi/abs/10.1207/s15327051hci0202_3)).
62. Gaver, William W. 1988. "Everyday Listening and Auditory Icons." University of California, San Diego.
63. Gawande, a a, W. J. Donovan, a P. Ginsburg, and M. F. Marmor. 1989. "Photoaversion in Retinitis Pigmentosa." *The British journal of ophthalmology* 73(2):115–20.
64. Geldard, Frank a. 1957. "Adventures in Tactile Literacy." *American Psychologist* 12(3):115–24.
65. Gibson, J. J. 1962. "Observations on Active Touch." *Psychological review* 69(6):477–91.
66. Gonot, A., N. Chateau, and M. Emerit. 2006. "Usability of 3D-Sound for Navigation in a Constrained Virtual Environment." Pp. 1–16 in *the 120th Audio Engineering Society Convention (Paris)*. Pairs.
67. Gordon, G. D. M. 1978. *And International Congress of Physiological Sciences. Active Touch: The Mechanism of Recognition of Objects by Manipulation : A Multi-Disciplinary Approach*. Pergamon Press Oxford ; New York.
68. Grover, Sandeep, G. A. Fishman, and Jeremiah Brown. 1998. "Patterns of Visual Field Progression in Patients with Retinitis Pigmentosa." *Ophthalmology* 105:1069–75.
69. Gygi, Brian and Valeriy Shafiro. 2009. "From Signal to Substance and Back: Insights from Environmental Sound Research to Auditory Display Design." Pp. 306–29 in *Proceedings of the 6th International Conference on Auditory Display, CMMR/ICAD'09*, edited by Sølvi Ystad, Mitsuko Aramaki, Richard Kronland-Martinet, and Kristoffer Jensen. Copenhagen, Denmark: Springer-Verlag. Retrieved ([http://link.springer.com/10.1007/978-3-642-12439-6\\_16](http://link.springer.com/10.1007/978-3-642-12439-6_16)).
70. Hajnal, Alen, Daniel T. Abdul-Malak, and Frank H. Durgin. 2011. "The Perceptual Experience of Slope by Foot and by Finger." *Journal of Experimental Psychology: Human Perception and Performance* 37(3):709–19. Retrieved (<http://doi.apa.org/getdoi.cfm?doi=10.1037/a0019950>).
71. Hartmann, W. M. and B. Raked. 1989. "On the Minimum Audible Angle--a Decision Theory Approach." *The Journal of the Acoustical Society of America* 85(5):2031–41.
72. Hellier, E. J., J. Edworthy, and I. Dennis. 1993. "Improving Auditory Warning Design: Quantifying and Predicting the Effects of Different Warning Parameters on Perceived Urgency." *Human factors* 35(4):693–706.
73. Hemmert, Fabian. 2010. "Interfaces beyond the Surface: A Structural Approach to Embodiment." Pp. 2903–6 in *CHI '10 Extended Abstracts on Human Factors in Computing Systems (CHI EA '10)*. New York, USA: ACM Press. Retrieved (<http://dl.acm.org/citation.cfm?id=1753881>).
74. Henriques, D. Y. P., W. P. Medendorp, A. Z. Khan, and J. D. Crawford. 2002. "Visuomotor Transformations for Eye-Hand Coordination." *Progress in brain research* 140:329–40.

75. Henze, Niels, Wilko Heuten, and Susanne Boll. 2006. "Non-Intrusive Somatosensory Navigation Support for Blind Pedestrians." Pp. 459–64 in *Proc of Eurohaptics*, vol. 2006. Retrieved (<http://lsc.univ-evry.fr/~eurohaptics/upload/cd/papers/f76.pdf>).
76. Hermann, Thomas, Andy Hunt, John G. Neuhoff, and Paul Vickers. 2011. "Chapitre 18-Sonification for Process Monitoring." Pp. 455–91 in *The Sonification Handbook*.
77. Herse, Peter. 2005. "Retinitis Pigmentosa: Visual Function and Multidisciplinary Management." *Clinical & experimental optometry : journal of the Australian Optometrical Association* 88(5):335–50.
78. Heuten, Wilko, Niels Henze, Susanne Boll, and Martin Pielot. 2008. "Tactile Wayfinder: A Non-Visual Support System for Wayfinding." Pp. 172–81 in *Proceedings of the 5th Nordic conference on Human-computer interaction building bridges - NordiCHI '08*. New York, USA: ACM Press. Retrieved (<http://portal.acm.org/citation.cfm?doid=1463160.1463179>).
79. Hinckley, Ken and Hyunyoung Song. 2011. "Sensor Synaesthesia: Touch in Motion, and Motion in Touch." Pp. 801–10 in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '11*. New York, NY, USA: ACM. Retrieved (<http://doi.acm.org/10.1145/1978942.1979059>).
80. Hoggan, Eve, Andrew Crossan, Stephen A. Brewster, and Topi Kaaresoja. 2009. "Audio or Tactile Feedback: Which Modality When?" Pp. 2253–56 in *Proceedings of the 27th international conference on Human factors in computing systems - CHI 09*. New York, USA: ACM Press. Retrieved (<http://dl.acm.org/citation.cfm?doid=1518701.1519045>).
81. Hoshiyama, Minoru, Ryusuke Kakigi, and Yohei Tamura. 2004. "Temporal Discrimination Threshold on Various Parts of the Body." *Muscle and Nerve* 29(2):243–47.
82. Hutchins, Edwin, James Hollan, and Donald Norman. 1985. "Direct Manipulation Interfaces." *Human-Computer Interaction* 1(4):311–38.
83. ISO 7029:2000. 2000. *ISO 7029:2000 : Acoustique -- Distribution Statistique Des Seuils D'audition En Fonction de L'âge*. Retrieved ([http://www.iso.org/iso/fr/catalogue\\_detail.htm?csnumber=26314](http://www.iso.org/iso/fr/catalogue_detail.htm?csnumber=26314)).
84. Jacquet, Emmanuelle, Jérôme Chambert, and Djamel Remache. 2013. "Comportement Vieillessement et Peau." *MECAMAT 2013 <hal-01002388>* (Jan 2013):1–10. Retrieved (<https://hal.archives-ouvertes.fr/hal-01002388/document>).
85. Jahnke, John C. 1963. "Serial Position Effects in Immediate Serial Recall." *Journal of Verbal Learning and Verbal Behavior* 2(3):284–87. Retrieved (<http://www.sciencedirect.com/science/article/pii/S002253716380095X>).
86. Johansson, R. S. and a B. Vallbo. 1979. "Tactile Sensibility in the Human Hand: Relative and Absolute Densities of Four Types of Mechanoreceptive Units in Glabrous Skin." *The Journal of physiology* 286(1):283–300.
87. Johnson, K. O. and J. R. Phillips. 1981. "Tactile Spatial Resolution. I. Two-Point Discrimination, Gap Detection, Grating Resolution, and Letter Recognition." *Journal of neurophysiology* 46(6):1177–92. Retrieved (<http://jn.physiology.org/content/46/6/1177.full-text.pdf+html>).
88. Jones, Chanon M., Rob Gray, Charles Spence, and Hong Z. Tan. 2008. "Directing Visual Attention with Spatially Informative and Spatially Noninformative Tactile Cues." *Experimental brain research* 186(4):659–69. Retrieved December 2, 2014 (<http://www.ncbi.nlm.nih.gov/pubmed/18224305>).
89. Jones, Lynette A., Mealani Nakamura, and Brett Lockyer. 2004. "Development of a Tactile Vest." *Proceedings - 12th International Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems, HAPTICS* 82–89.
90. Jones, Matt et al. 2008. "ONTRACK: Dynamically Adapting Music Playback to Support Navigation." *Personal and Ubiquitous Computing* 12(7):513–25.
91. Kane, Shaun K. et al. 2011. "Access Overlays: Improving Non-Visual Access to Large Touch Screens for Blind Users." Pp. 273–82 in *Proceedings of the 24th annual ACM symposium on User interface software and technology - UIST '11*. New York, USA: ACM Press. Retrieved (<http://dl.acm.org/citation.cfm?id=2047196.2047232>).
92. Katz, Brian F. G. et al. 2012. "NAVIG: Augmented Reality Guidance System for the Visually Impaired: Combining Object Localization, GNSS, and Spatial Audio." *Virtual Reality* 16(4):253–69.
93. Katz, Brian FG G. et al. 2012. "NAVIG: Guidance System for the Visually Impaired Using Virtual Augmented Reality." *Technology and Disability* 24(2):163–78.

94. Keim, Daniel et al. 2008. "Visual Analytics: Definition, Process, and Challenges." Pp. 154–75 in *Information Visualization: Human-Centered Issues and Perspectives*, vol. 4950, *Lecture Notes in Computer Science*, edited by Andreas Kerren, John T. Stasko, Jean-Daniel Fekete, and Chris North. Springer. Retrieved ([http://dx.doi.org/10.1007/978-3-540-70956-5\\_7](http://dx.doi.org/10.1007/978-3-540-70956-5_7)).
95. Kildal, Johan, Susanna Paasoara, and Viljakaisa Aaltonen. 2012. "Kinetic Device: Designing Interactions with a Deformable Mobile Interface." Pp. 1871–76 in *CHI '12 Extended Abstracts on Human Factors in Computing Systems (CHI EA '12)*. New York, USA: ACM Press. Retrieved (<http://dl.acm.org/citation.cfm?doid=2212776.2223721>).
96. Kim, Seoktae, Hyunjung Kim, Boram Lee, Tek-Jin Nam, and Woohun Lee. 2008. "Inflatable Mouse: Volume-Adjustable Mouse with Air-Pressure-Sensitive Input and Haptic Feedback." Pp. 211–24 in *Proceeding of the twenty-sixth annual CHI conference on Human factors in computing systems - CHI '08*. New York, USA: ACM Press. Retrieved ([http://dl.acm.org/ft\\_gateway.cfm?id=1357090&ftid=498849&dwn=1&CFID=174223433&CFTOKEN=32504795](http://dl.acm.org/ft_gateway.cfm?id=1357090&ftid=498849&dwn=1&CFID=174223433&CFTOKEN=32504795)) (<http://portal.acm.org/citation.cfm?doid=1357054.1357090>).
97. Kozlowski, J. M. D. and A. E. Jalkh. 1985. "An Improved Negative-Lens Field Expander for Patients With Concentric Field Constriction." *Archives of Ophthalmology* 103(3):326–326.
98. Kramer, Gregory. 1993. *Auditory Display: Sonification, Audification, and Auditory Interfaces*. Perseus Publishing.
99. Krefman, R. A. 1981. "Reversed Telescopes on Visual Efficiency Scores in Field-Restricted Patients." *American journal of optometry and physiological optics* 58(2):159–62.
100. Kuyk, T. K., L. Liu, and P. S. W. Fuhr. 2005. "Feature Search in Persons with Severe Visual Impairment." *Vision Research* 45:3224–34.
101. Larsen, Camilla H., David S. Lauritsen, Jacob J. Larsen, Marc Pilgaard, and Jacob B. Madsen. 2013. "Differences in Human Audio Localization Performance between a HRTF- and a Non-HRTF Audio System." *Proceedings of the 8th Audio Mostly Conference on - AM '13* 5:1–5:8. Retrieved (<http://dl.acm.org/citation.cfm?doid=2544114.2544118>).
102. Lederman, S. J. and Roberta L. Klatzky. 2009. "Haptic Perception: A Tutorial." *Attention, perception & psychophysics* 71(7):1439–59. Retrieved (<http://link.springer.com/article/10.3758/APP.71.7.1439#page-1>).
103. Lehtinen, Ville, Antti Oulasvirta, Antti Salovaara, and Petteri Nurmi. 2012. "Dynamic Tactile Guidance for Visual Search Tasks." *Proceedings of the 25th annual ACM symposium on User interface software and technology - UIST '12* 445. Retrieved (<http://dl.acm.org/citation.cfm?id=2380173>).
104. Lenay, Charles, Olivier Gapenne, Sylvain Hanneton, Chaterine Marque, and Christelle Genouëlle. 2003. "Sensory Substitution: Limits and Perspectives." P. 320 in *Touching for knowing: cognitive psychology of haptic manual perception*. Yvette, Hatwell. Retrieved (<http://books.google.com/books?hl=fr&lr=&id=GSOHmpdyobAC&pgis=1>).
105. Leplâtre, Grégory and Stephen A. Brewster. 1998. "An Investigation of Using Music to Provide Navigation Cues." Pp. 19–19 in *Proceedings of the 1998 International Conference on Auditory Display, ICAD'98*. Swinton, UK, UK: British Computer Society. Retrieved (<http://dl.acm.org/citation.cfm?id=2227605.2227624>).
106. Letowski, Tomasz R. and Szymon T. Letowski. 2012. *Auditory Spatial Perception : Auditory Localization*. Retrieved ([http://www.arl.army.mil/www/default.cfm?technical\\_report=6466](http://www.arl.army.mil/www/default.cfm?technical_report=6466)).
107. Liao, Chunyuan, François Guimbretière, and Corinna E. Loeckenhoff. 2006. "Pen-Top Feedback for Paper-Based Interfaces." Pp. 201–10 in *Proceedings of the 19th Annual ACM Symposium on User Interface Software and Technology, UIST '06*. New York, NY, USA: ACM. Retrieved (<http://doi.acm.org/10.1145/1166253.1166285>).
108. Liljedahl, Mats et al. 2012. "Testing Two Tools for Multimodal Navigation." *Advances in Human-Computer Interaction* 2012.
109. Lindeman, RW, Yasuyuki Yanagida, John L. Sibert, and Robert Lavine. 2003. "Effective Vibrotactile Cueing in a Visual Search Task." *Proceedings of the ninth IFIP TC13 international conference on Human-computer interaction (INTERACT 2003)* 89–96. Retrieved February 10, 2014 ([http://books.google.com/books?hl=en&lr=&id=PTg0fVYqgCcC&oi=fnd&pg=PA89&dq=Effective+Vibrotactile+Cueing+in+a+Visual+Search+Task&ots=O8SKwzlyA\\_&sig=zW0TJW7PWU\\_5qhyYRH6QPBVjscl](http://books.google.com/books?hl=en&lr=&id=PTg0fVYqgCcC&oi=fnd&pg=PA89&dq=Effective+Vibrotactile+Cueing+in+a+Visual+Search+Task&ots=O8SKwzlyA_&sig=zW0TJW7PWU_5qhyYRH6QPBVjscl))

110. Lowe, John and Neville Drasdo. 1992. "Patient's Responses to Retinitis Pigmentosa." *Optometry & Vision Science* 69(3):182–85.
111. Makous, J. C., R. M. Friedman, and C. J. Vierck. 1995. "A Critical Band Filter in Touch." *The Journal of neuroscience : the official journal of the Society for Neuroscience* 15(April):2808–18.
112. Martin, J. C. 1994. "Cadre D'étude de La Multimodalité Fondé Sur Les Types et Buts de Coopération Entre Modalités." Pp. 97–106 in *Actes des troisiemes Journées Internationales sur "l'Interface des Mondes Réels et Virtuels."*
113. Marzo, Asier et al. 2015. "Holographic Acoustic Elements for Manipulation of Levitated Objects." *Nature Communications* 6. Retrieved (<http://dx.doi.org/10.1038/ncomms9661>).
114. Matscheko, Michael, Alois Ferscha, Andreas Riener, and Manuel Lehner. 2010. "Tactor Placement in Wrist Worn Wearables." *International Symposium on Wearable Computers (ISWC) 2010* 1–8.
115. Medendorp, W. Pieter, Sabine M. Beurze, Stan Van Pelt, and Jurrian Van Der Werf. 2008. "Behavioral and Cortical Mechanisms for Spatial Coding and Action Planning." *Cortex* 44:587–97.
116. Meijer, P. B. L. 1992. "An Experimental System for Auditory Image Representations." *IEEE Transactions on Biomedical Engineering* 39(2):112–21.
117. Mershon, D. H. and J. N. Bowers. 1979. "Absolute and Relative Cues for the Auditory Perception of Egocentric Distance." *Perception* 8(3):311–22.
118. Mershon, Donald H. and L. Edward King. 1975. "Intensity and Reverberation as Factors in." 18(6):409–15.
119. Mershon, Donald H. and L. Edward King. 1975. "Intensity and Reverberation as Factors in the Auditory Perception of Egocentric Distance." *Perception & Psychophysics* 18(6):409–15. Retrieved (<http://www.springerlink.com/index/10.3758/BF03204113>).
120. Miller, George A. A. 1956. "The Magical Number Seven, plus or Minus Two: Some Limits on Our Capacity for Processing Information." *Psychological Review* 63(2):81–97. Retrieved (<http://content.apa.org/journals/rev/63/2/81>).
121. Mills, A. W. 1958. "On the Minimum Audible Angle." *The Journal of the Acoustical Society of America* 30(4):237–46. Retrieved (<http://scitation.aip.org/content/asa/journal/jasa/30/4/10.1121/1.1909553>).
122. Ministère de l'éducation nationale - DGESCO. 2012. "Développer Sa Mémoire, Techniques de Mémorisation." 2–7. Retrieved ([http://cache.media.eduscol.education.fr/file/Accompagnement\\_personnalise/97/3/LyceesGT\\_Ressource\\_AP\\_Developper-memoire-technique-memorisation\\_sequence1\\_215973.pdf](http://cache.media.eduscol.education.fr/file/Accompagnement_personnalise/97/3/LyceesGT_Ressource_AP_Developper-memoire-technique-memorisation_sequence1_215973.pdf)).
123. Murdock, Bennet B., Jr. 1962. "The Serial Position Effect of Free Recall." *Journal of Experimental Psychology* 64(5):482–88.
124. Nakamura, Mealani and Lynette Jones. 2003. "An Actuator for the Tactile Vest " A Torso-Based Haptic Device." P. 333 – in *Proceedings of the 11th Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems (HAPTICS'03), HAPTICS '03*. Washington, DC, USA: IEEE Computer Society. Retrieved (<http://dl.acm.org/citation.cfm?id=795683.797550>).
125. Nakashima, Ryoichi and Satoshi Shioiri. 2014. "Why Do We Move Our Head to Look at an Object in Our Peripheral Region? Lateral Viewing Interferes with Attentive Search." *PLoS ONE* 9(3).
126. Nelson, P., P. Aspinall, and C. O'Brien. 1999. "Patients' Perception of Visual Impairment in Glaucoma: A Pilot Study." *The British journal of ophthalmology* 83(5):546–52.
127. Ngo, Mary Kim and Charles Spence. 2010. "Auditory, Tactile, and Multisensory Cues Facilitate Search for Dynamic Visual Stimuli." *Attention, perception & psychophysics* 72(6):1654–65.
128. Nicolau, Hugo, João Guerreiro, Tiago Guerreiro, and Luís Carriço. 2013. "UbiBraille: Designing and Evaluating a Vibrotactile Braille-Reading Device." P. Article No. 23, 8 pages in *Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility, (ASSETS '13)*. ACM, New York. Retrieved (<http://dl.acm.org/citation.cfm?doid=2513383.2513437>).
129. Nigay, Laurence and Joëlle Coutaz. 1993. "A Design Space for Multimodal Systems: Concurrent Processing and Data Fusion." Pp. 172–78 in *Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems (CHI '93), CHI '93*. New York, USA: ACM Press. Retrieved (<http://portal.acm.org/citation.cfm?doid=169059.169143>).

130. Nigay, Laurence and Joëlle Coutaz. 1996. "Espaces Conceptuels Pour Multimédia et Multimodale L' Interaction." *Techniques et Sciences Informatiques, Spécial Multimédia et Collecticiel, AFCET & Hermes Publ.* 15(9), 1195–1225.
131. Nigay, Laurence and Joëlle Coutaz. 1997. "Multifeature Systems: The CARE Properties and Their Impact on Software Design." in *Intelligence and multimodality in multimedia interfaces*. Retrieved (<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.12.7140&rep=rep1&type=pdf>).
132. Nikolakis, Georgios and Dimitrios Tzovaras. 2004. "CyberGrasp and PHANTOM Integration: Enhanced Haptic Access for Visually Impaired Users." *Conference Speech and Computer* 507–13. Retrieved ([http://www.isca-speech.org/archive\\_open/specom\\_04/spc4\\_507.html](http://www.isca-speech.org/archive_open/specom_04/spc4_507.html)).
133. Oh, Uran, Shaun K. Kane, and Leah Findlater. 2013. "Follow That Sound: Using Sonification and Corrective Verbal Feedback to Teach Touchscreen Gestures." P. Article 13 , 8 pages in *Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility, (ASSETS '13)*. ACM, New York. Retrieved (<http://dl.acm.org/citation.cfm?id=2513455>).
134. Oron-Gilad, Tal, Joshua L. Downs, Richard D. Gilson, and Peter a. Hancock. 2007. "Vibrotactile Guidance Cues for Target Acquisition." *IEEE Transactions on Systems, Man and Cybernetics, Part C (Applications and Reviews)* 37(5):993–1004.
135. Palladino, Dianne K. and Bruce N. Walker. 2007. "Learning Rates for Auditory Menus Enhanced with Spearcons Versus Earcons." Pp. 274–79 in *International Conference on Auditory Display*. Montréal, Canada: Georgia Institute of Technology International Community on Auditory Display. Retrieved (<http://hdl.handle.net/1853/50011>).
136. Panëels, Sabrina A., Adriana Olmos, Jeffrey R. Blum, and Jeremy R. Cooperstock. 2013. "Listen to It Yourself!: Evaluating Usability of 'What's Around Me?' For the Blind." Pp. 2107–16 in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '13*. ACM, New York. Retrieved (<http://dl.acm.org/citation.cfm?id=2481290>).
137. Pasquero, Jérôme and Vincent Hayward. 2003. "STRess : A Practical Tactile Display System with One Millimeter Spatial Resolution and 700 Hz Refresh Rate." Pp. 94–110 in *Proc Eurohaptics*, vol. 2003. Dublin, Ireland. Retrieved (<http://cim.mcgill.ca/~haptic/pub/JP-VH-EH-03.pdf>).
138. Patterson, Roy D. and Civil Aviation Authority. 1982. *Guidelines for Auditory Warning Systems on Civil Aircraft*. Civil Aviation Authority.
139. Peterson, Lloyd and Margaret Jean Peterson. 1959. "Short-Term Retention of Individual Verbal Items." *Journal of Experimental Psychology* 58(3):193–98. Retrieved (<http://content.apa.org/journals/xge/58/3/193>).
140. Pielot, Martin, Benjamin Poppinga, and Susanne Boll. 2010. "PocketNavigator: Vibro-Tactile Waypoint Navigation for Everyday Mobile Devices." *MobileHCI 2010* 423–26. Retrieved (<http://portal.acm.org/citation.cfm?id=1851696>).
141. Pielot, Martin, Benjamin Poppinga, Wilko Heuten, and Susanne Boll. 2011. "A Tactile Compass for Eyes-Free Pedestrian Navigation." Pp. 640–56 in *Human-Computer Interaction - Interact 2011, LNCS 6947*, edited by Pedro Campos et al. Lisbon, Portugal: Springer Berlin Heidelberg.
142. Pielot, Martin, Benjamin Poppinga, Wilko Heuten, and Susanne Boll. 2012. "PocketNavigator: Studying Tactile Navigation Systems in-Situ." Pp. 3131–40 in *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems - CHI '12*. ACM, New York. Retrieved (<http://dl.acm.org/citation.cfm?doid=2207676.2208728>).
143. Prasad, Manoj, Paul Taelle, Daniel Goldberg, and Tracy a. Hammond. 2014. "HaptiMoto: Turn-by-Turn Haptic Route Guidance Interface for Motorcyclists." Pp. 3597–3606 in *Proceedings of the 32nd annual ACM conference on Human factors in computing systems - CHI '14*. ACM, New York. Retrieved (<http://dl.acm.org/citation.cfm?id=2556288.2557404>).
144. Proffitt, Dennis R., Jeanine Stefanucci, Tom Banton, and William Epstein. 2003. "The Role of Effort in Perceiving Distance." *Psychological Science* 14(2):106–12.
145. Qi Wang and Vincent Hayward. 2006. "Compact, Portable, Modular, High-Performance, Distributed Tactile Transducer Device Based on Lateral Skin Deformation." Pp. 67–72 in *2006 14th Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*. IEEE. Retrieved (<http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1627091>).

146. Raggenbass, Mario. 2011. "Somesthésie." *Neurosciences Fondamentales, Centre Médical Universitaire CH-1211 Genève* 4 56. Retrieved (<http://www.unige.ch/medecine/neuf/files/7214/0369/8105/sometdouleur.pdf>).
147. Renaudin, Valérie and Pierre-Yves Gilliéron. 2011. "Personal Robust Navigation in Challenging Applications." *Journal of Navigation* 64(02):235–49.
148. Révész, Géza. 1950. *Psychology and Art of the Blind*. New York: Longmans.
149. Roudaut, Anne and Eric Lecolinet. 2007. "Un Espace de Classification Pour L'interaction Sur Dispositifs Mobiles." *Proceedings of the 19th International Conference of the Association Francophone d'Interaction Homme-Machine on - IHM '07* 99. Retrieved (<http://portal.acm.org/citation.cfm?doid=1541436.1541456>).
150. Rümelin, Sonja, Enrico Rukzio, and Robert Hardy. 2011. "NaviRadar: A Novel Tactile Information Display for Pedestrian Navigation." Pp. 293–302 in *Proceedings of the 24th annual ACM symposium on User interface software and technology (UIST '11), UIST '11*. New York, USA: ACM Press. Retrieved (<http://doi.acm.org/10.1145/2047196.2047234>).
151. Schloerb, David W., Orly Lahav, Joseph G. Desloge, and Mandayam A. Srinivasan. 2010. "BlindAid: Virtual Environment System for Self-Reliant Trip Planning and Orientation and Mobility Training." Pp. 363–70 in *2010 IEEE Haptics Symposium, HAPTICS 2010*. IEEE. Retrieved ([http://ieeexplore.ieee.org/xpl/login.jsp?tp=&arnumber=5444631&url=http://ieeexplore.ieee.org/xpl/abs\\_all.jsp?arnumber=5444631](http://ieeexplore.ieee.org/xpl/login.jsp?tp=&arnumber=5444631&url=http://ieeexplore.ieee.org/xpl/abs_all.jsp?arnumber=5444631)).
152. Serrano, Marcos. 2010. "Interaction Multimodale En Entrée : Conception et Prototypage." 304.
153. Shaer, Orit and Eva Hornecker. 2010. "Tangible User Interfaces: Past, Present, and Future Directions." *Found. Trends Hum.-Comput. Interact.* 3(1&#8211;2):1–137. Retrieved (<http://dx.doi.org/10.1561/11000000026>).
154. Siegfried, Adrien. 2015. *HaptiQ: A Haptic Tracking Device with Vectorial Guidance for Graph Exploration*. Toulouse.
155. Smith, Nicholas D., David P. Crabb, and David F. Garway-Heath. 2011. "An Exploratory Study of Visual Search Performance in Glaucoma." *Ophthalmic & physiological optics: the journal of the British College of Ophthalmic Opticians (Optometrists)* 31(3):225–32.
156. Stevens, S. S. 1957. "On the Psychophysical Law." *Psychological Review* 64(3):153–81.
157. Stuart, Meg, a Bulent Turman, Jacqueline Shaw, Natalie Walsh, and Vincent Nguyen. 2003. "Effects of Aging on Vibration Detection Thresholds at Various Body Regions." *BMC geriatrics* 3(January):1.
158. Sucu, Burkay and Eelke Folmer. 2013. "Haptic Interface for Non-Visual Steering." *Proceedings of the 2013 international conference on Intelligent user interfaces - IUI '13* 427. Retrieved (<http://dl.acm.org/citation.cfm?doid=2449396.2449451>).
159. Sucu, Burkay and Eelke Folmer. 2014. "The Blind Driver Challenge: Steering Using Haptic Cues." Pp. 3–10 in *Proceedings of the 16th international ACM SIGACCESS conference on Computers & accessibility - ASSETS '14*. ACM, New York. Retrieved (<http://dl.acm.org/citation.cfm?doid=2661334.2661357>).
160. Talbot, Martin and William Cowan. 2009. "On the Audio Representation of Distance for Blind Users." Pp. 1839–48 in *Proceedings of the 27th international conference on Human factors in computing systems - CHI 09*. New York, USA: ACM Press. Retrieved (<http://dl.acm.org/citation.cfm?doid=1518701.1518984>).
161. Tan, Hong Z. and Alex Pentland. 1997. "Tactual Displays for Wearable Computing." *Personal Technologies* 1(4):225–30. Retrieved (<http://link.springer.com/article/10.1007/BF01682025>).
162. Toney, Aaron, Lucy Dunne, B. H. Thomas, and S. P. Ashdown. 2003. "A Shoulder Pad Insert Vibrotactile Display." in *Proceedings of the 7th IEEE International Symposium on Wearable Computers (ISWC '03)*.
163. Trocka-Leszczynska, Elzbieta and Joanna Jablonska. 2013. "Design of Modern Hotels – Humanization of the Residential Environment." Pp. 358–67 in *Universal Access in Human-Computer Interaction. Design Methods, Tools, and Interaction Techniques for eInclusion*, vol. 8009 LNCS, edited by Constantine Stephanidis and Margherita Antona. Springer Berlin Heidelberg.
164. Tsukada, K. and M. Yasumura. 2004. *ActiveBelt: Belt-Type Wearable Tactile Display for Directional Navigation*. edited by Nigel Davies, Elizabeth D Mynatt, and Itiro Siio. Berlin, Heidelberg: Springer Berlin Heidelberg. Retrieved ([http://link.springer.com/chapter/10.1007/978-3-540-30119-6\\_23](http://link.springer.com/chapter/10.1007/978-3-540-30119-6_23)).

165. Turano, Kathleen a., Dylan Yu, Lei Hao, and John C. Hicks. 2005. "Optic-Flow and Egocentric-Direction Strategies in Walking: Central vs Peripheral Visual Field." *Vision Research* 45(25-26):3117–32.
166. Turk, Matthew and George Robertson. 2000. "Perceptual User Interfaces." *Communications of the ACM* 43(3):33–34. Retrieved ([http://link.springer.com/chapter/10.1007/978-1-4471-0259-5\\_4](http://link.springer.com/chapter/10.1007/978-1-4471-0259-5_4)).
167. Vargas-Martin, F. and E. Peli. 2002. "Augmented-View for Restricted Visual Field: Multiple Device Implementations." *Optometry & Vision Science* 79(11):715–23.
168. Velazquez, R., E. Pissaloux, E. Fontaine, and M. Hafez. 2007. "Vers Une Aide Pour L'interaction Avec L'espace Pour Les Déficients Visuels Basée Sur La Perception Tactile." *JESA (Journal Européen des Systèmes Automatisés, N° Spécial « Robotique et Handicap »)* 41/2:179–98.
169. Vieira, João, Maurício Sousa, Artur Arsénio, and Joaquim Jorge. 2015. "Augmented Reality for Rehabilitation Using Multimodal Feedback." Pp. 38–41 in *Proceedings of the 3rd 2015 Workshop on ICTs for Improving Patients Rehabilitation Research Techniques, REHAB '15*. New York, NY, USA: ACM. Retrieved (<http://doi.acm.org/10.1145/2838944.2838954>).
170. Voinier, T. and F. Brioloe. 1992. "Simulation de L'espace Sonore Sous écouteurs." *Colloque C1, Supplément au J. de Phys. III* 2(avril):121–24. Retrieved (<http://demeter.revue.univ-lille3.fr/lodel9/index.php?id=466>).
171. Wagner, Julie, Mathieu Nancel, Sean G. Gustafson, Stephane Huot, and Wendy E. Mackay. 2013. "Body-Centric Design Space for Multi-Surface Interaction." Pp. 1299–1308 in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '13*. ACM, New York. Retrieved (<http://dl.acm.org/citation.cfm?doid=2470654.2466170>).
172. Walker, Bruce N., Amanda Nance, and Jeffrey Lindsay. 2006. "Spearcons : Speech-Based Earcons Improve Navigation Performance in Auditory Menus." Pp. 63–68 in *Proceedings of the International Conference on Auditory Display, ICAD 2006*. London, England.
173. Weibel, Samuel et al. 2006. "SharpWComp 1.0 , Inter Deposit Digital Number." Retrieved (<http://stephane.lavirotte.com/research/biblio/hourdin-lavirotte-et-al2006a.pdf>).
174. Witt, Jessica K., Dennis R. Proffitt, and William Epstein. 2004. "Perceiving Distance: A Role of Effort and Intent." *Perception* 33(5):577–90.
175. Wolfe, Jeremy M. 1998. "What Can 1,000,000 Trials Tell Us about Visual Search?" *Psychological Science* 9(1):33–39.
176. Woods, Adam J., John W. Philbeck, and Jerome V Danoff. 2009. "The Various Perceptions of Distance: An Alternative View of How Effort Affects Distance Judgments." *Journal of Experimental Psychology: Human Perception and Performance* 35(4):1104–17. Retrieved (<http://doi.apa.org/getdoi.cfm?doi=10.1037/a0013622>).
177. Woods, Russell L., Robert G. Giorgi, Eliot L. Berson, and Eli Peli. 2010. "Extended Wearing Trial of Trifield Lens Device for 'tunnel Vision.'" *Ophthalmic and Physiological Optics* 30(3):240–52. Retrieved (<http://doi.wiley.com/10.1111/j.1475-1313.2010.00718.x>).
178. Yamano, Shingo, Takamitsu Hamajo, Shunsuke Takahashi, and Keita Higuchi. 2012. "EyeSound: Single-Modal Mobile Navigation Using Directionally Annotated Music." Pp. 22:1–22:4 in *Proceedings of the 3rd Augmented Human International Conference, AH '12*. New York, NY, USA: ACM.
179. Ye, Hanlu, Meethu Malu, Uran Oh, and Leah Findlater. 2014. "Current and Future Mobile and Wearable Device Use by People With Visual Impairments." *CHI 2014* 3123–32.
180. Young, Rockefeller S. L. and Gerald A. Fishman. 1982. "Sensitivity Losses in a Long Wavelength Sensitive Mechanism of Patients with Retinitis Pigmentosa." *Vision Research* 22(1):163–72.
181. Zahorik, Pavel. 2002. "Assessing Auditory Distance Perception Using Virtual Acoustics." *The Journal of the Acoustical Society of America* 111(4):1832–46. Retrieved (<http://www.ncbi.nlm.nih.gov/pubmed/12002867>).
182. Zahorik, Pavel, Douglas S. Brungart, and Adelbert W. Bronkhorst. 2005. "Auditory Distance Perception in Humans: A Summary of Past and Present Research." *Acta Acustica united with Acustica* 91(3):409–20.
183. Zeng, Limin, Denise Prescher, and Gerhard Weber. 2012. "Exploration and Avoidance of Surrounding Obstacles for the Visually Impaired." Pp. 111–18 in *Proceedings of the 14th international ACM SIGACCESS conference on Computers and accessibility (ASSETS '12)*. New York: ACM Press. Retrieved (<http://dl.acm.org/citation.cfm?id=2384936>).

184. Zeng, Tao, Frédéric Giraud, Betty Lemaire-Semail, and Michel Amberg. 2010. "Analysis of a New Haptic Display Coupling Tactile and Kinesthetic Feedback to Render Texture and Shape." Pp. 87–93 in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 6192 LNCS.
185. Zhao, Haixia, Catherine Plaisant, Ben Shneiderman, and Jonathan Lazar. 2008. "Data Sonification for Users with Visual Impairment." *ACM Transactions on Computer-Human Interaction* 15(1):1–28.
186. Zwicker, E. 1975. "Scaling." Pp. 401–48 in *Auditory System SE - 9*, vol. 5 / 2, *Handbook of Sensory Physiology*, edited by Wolf D. Keidel and William D. Neff. Springer Berlin Heidelberg. Retrieved ([http://dx.doi.org/10.1007/978-3-642-65995-9\\_9](http://dx.doi.org/10.1007/978-3-642-65995-9_9)).



# **Partie 9 :**

# **Annexes**



## 9.1 Sommaire étendu

<b>1</b>	<b><u>INTRODUCTION</u></b>	<b>15</b>
1.1	PLAN DE LECTURE	17
<b>2</b>	<b><u>ETAT DE L'ART SUR LA TRANSMISSION D'INFORMATION BASEE SUR LES MODALITES TACTILES ET/OU AUDITIVES</u></b>	<b>23</b>
2.1	MODALITES TACTILES	23
2.1.1	STIMULUS TACTILE	23
2.1.2	PERCEPTION CHEZ L'HOMME : MODALITE SENSORIELLE TACTILE	24
2.1.3	ARTEFACTS TECHNOLOGIQUES POUR PRODUIRE DES STIMULI TACTILES	28
2.1.4	MISE EN ŒUVRE DE LA MODALITE SENSORIELLE TACTILE DANS DES SITUATIONS INTERACTIVES	29
2.1.5	APPORT DU TACTILE AUX SITUATIONS D'INTERACTION NON VISUELLES.	44
2.2	MODALITES AUDITIVES	46
2.2.1	STIMULUS AUDITIF	46
2.2.2	PERCEPTION CHEZ L'HOMME : MODALITE SENSORIELLE AUDITIVE	46
2.2.3	LES MAPPINGS SONORES	54
2.2.4	ARTEFACTS TECHNOLOGIQUES POUR PRODUIRE DES STIMULI AUDITIFS	55
2.2.5	MISE EN ŒUVRE DE LA MODALITE AUDITIVE DANS DES SITUATIONS INTERACTIVES	56
2.2.6	APPORT DE L'AUDITIF AUX SITUATIONS D'INTERACTION NON VISUELLE	64
2.3	MULTIMODALITE	66
2.3.1	MODALITES ET MULTIMODALITE	66
2.3.2	LES PROPRIETES DE LA MULTIMODALITE	66
2.3.3	MISE EN ŒUVRE DE LA MULTIMODALITE DANS DES SITUATIONS INTERACTIVES	70
2.3.4	CONCLUSION SUR LA MULTIMODALITE	74
2.4	CONCLUSION DE L'ETAT DE L'ART	74
<b>3</b>	<b><u>UTILISATION/DEFINITION D'UN CADRE ANALYTIQUE</u></b>	<b>79</b>
3.1	MOTIVATION	79
3.2	PRESENTATION DES FACTEURS DE CONCEPTION DE NOTRE CADRE ANALYTIQUE	80
3.2.1	CARTOGRAPHIE DES FACTEURS PERTINENTS POUR UNE TRANSMISSION D'INFORMATION	82
3.2.2	CONCLUSION SUR L'EXTRACTION DES FACTEURS DE CONCEPTION	90
3.3	ORGANISATION DES FACTEURS DE CONCEPTION DANS NOTRE CADRE D'ANALYSE	90
3.3.1	CONCLUSION DE L'ORGANISATION DES FACTEURS DE CONCEPTION	94
3.4	ILLUSTRATION DE GENERATION DE TECHNIQUES AVEC NOTRE CADRE D'ANALYSE	95
3.4.1	PRESENTATION DE TECHNIQUES D'INTERACTION NON VISUELLE OPTIMISEES(?) POUR TRANSMETTRE UNE VALEUR	96
3.4.2	DISCUSSION ET CONCLUSION SUR LA GENERATION DES TECHNIQUES	107
3.5	CONCLUSION SUR L'APPORT DU CADRE D'ANALYSE POUR CONCEVOIR NOS TECHNIQUES	111
<b>4</b>	<b><u>PREMIERE ETUDE : TECHNIQUES D'INTERACTION NON VISUELLE POUR ESTIMER UNE VALEUR</u></b>	<b>115</b>

<b>4.1</b>	<b>MOTIVATION</b>	<b>115</b>
<b>4.2</b>	<b>EXPERIENCE 1.1 : TECHNIQUES D'INTERACTION NON VISUELLE POUR ESTIMER UNE VALEUR</b>	<b>116</b>
4.2.1	HUIT TECHNIQUES ET SEPT « ALTERNATIVES »	117
4.2.2	PROTOCOLE EXPERIMENTAL (EXP. 1.1)	120
4.2.3	RESULTATS DE L'EXPERIENCE 1.1	121
4.2.4	DISCUSSION DE L'EXPERIENCE 1.1	131
4.2.5	CONCLUSION DE L'EXPERIENCE 1.1	136
<b>4.3</b>	<b>EXPERIENCE 1.2 : ETUDE SPECIFIQUE DES FACTEURS DE CONCEPTION A/P ET AR/SR</b>	<b>137</b>
4.3.1	MISE EN PLACE DE L'EXPERIMENTATION	137
4.3.2	PRESENTATION DES QUATRE NOUVELLES TECHNIQUES	138
4.3.3	PROTOCOLE EXPERIMENTAL (EXP. 1.2)	144
4.3.4	RESULTATS DE L'EXPERIENCE 1.2	144
4.3.5	DISCUSSION DE L'EXPERIENCE 1.2	153
4.3.6	CONCLUSION DE L'EXPERIENCE 1.2	155
<b>4.4</b>	<b>CONCLUSION DE LA PREMIERE ETUDE</b>	<b>156</b>
<b>5</b>	<b><u>DEUXIEME ETUDE : INTERACTIONS NON VISUELLES POUR FACILITER LE RENDU D'UNE SEQUENCE DE COUPLES DIRECTION-DISTANCE</u></b>	<b><u>161</u></b>
<hr/>		
<b>5.1</b>	<b>INTRODUCTION A LA DEUXIEME ETUDE</b>	<b>161</b>
<b>5.2</b>	<b>RAPPEL SUR L'ETAT DE L'ART : TRANSMISSION DE DIRECTION ET DE DISTANCE</b>	<b>163</b>
<b>5.3</b>	<b>PRE-ETUDES ET CHOIX DE DESIGN</b>	<b>163</b>
<b>5.4</b>	<b>DESCRIPTION DES QUATRE TECHNIQUES RETENUES</b>	<b>165</b>
<b>5.5</b>	<b>HYPOTHESES DE TRAVAIL</b>	<b>172</b>
<b>5.6</b>	<b>PROTOCOLE EXPERIMENTAL DE L'EXPERIENCE 2.1</b>	<b>172</b>
<b>5.7</b>	<b>ANALYSE DES RESULTATS DES SUJETS NON-DV</b>	<b>174</b>
<b>5.8</b>	<b>ANALYSE DES RESULTATS PRELIMINAIRES CONCERNANT LES SUJETS DV</b>	<b>179</b>
<b>5.9</b>	<b>ANALYSE SPECIFIQUE : MEMORISATION DES SEGMENTS SELON LEUR POSITION DANS L'ITINERAIRE</b>	<b>181</b>
<b>5.10</b>	<b>DISCUSSION ET CONCLUSION DE LA DEUXIEME ETUDE</b>	<b>182</b>
<b>6</b>	<b><u>TROISIEME ETUDE : ASSISTANCE TACTILE A LA LOCALISATION DE CIBLES PERIPHERIQUES POUR DES PERSONNES A VISION TUBULAIRE</u></b>	<b><u>187</u></b>
<hr/>		
<b>6.1</b>	<b>ÉTAT DE L'ART SUR LES TECHNOLOGIES D'ASSISTANCE A LA RECHERCHE VISUELLE</b>	<b>188</b>
<b>6.2</b>	<b>DISPOSITIFS TACTILES POUR LA PERCEPTION SPATIALE</b>	<b>191</b>
<b>6.3</b>	<b>EXPERIENCE 3.1 : CONCEPTION ET COMPARAISON D'INDICES TACTILES PERMETTANT DE LOCALISER UNE CIBLE PERIPHERIQUE</b>	<b>193</b>
6.3.1	CONCEPTION DES INDICES TACTILES	193
6.3.2	PRESENTATION DES TECHNIQUES AVEC NOTRE CADRE ANALYTIQUE	197
6.3.3	ÉVALUATION EXPERIMENTALE (EXP. 3.1)	204
6.3.4	RESULTATS DE L'EXPERIENCE 3.1	206
6.3.5	CONCLUSIONS DE L'EXPERIENCE 3.1	207
<b>6.4</b>	<b>EXPERIENCE 3.2 : RECHERCHE VISUELLE D'UNE PETITE CIBLE DANS UN SALON VIRTUEL ENCOMBRE</b>	<b>208</b>
6.4.1	PARTICIPANTS	208
6.4.2	ENVIRONNEMENT VIRTUEL, DISPOSITIF TACTILE ET STIMULI	208
6.4.3	PROCEDURE EXPERIMENTALE (EXP. 3.2)	211

6.4.4	L'ANALYSE DES DONNEES	212
6.4.5	RESULTATS DE L'EXPERIENCE 3.2	212
6.4.6	CONCLUSIONS DE L'EXPERIENCE 3.2	215
<b>6.5</b>	<b>EXPERIENCE 3.3 : ETUDE DE CAS SUR UN SUJET AVEC UNE VISION TUBULAIRE</b>	<b>215</b>
6.5.1	PARTICIPANT	215
6.5.2	ENVIRONNEMENT VIRTUEL, DISPOSITIF TACTILE ET CIBLES	215
6.5.3	PROCEDURE EXPERIMENTALE (EXP. 3.3)	215
6.5.4	RESULTATS DE L'EXPERIENCE 3.3	216
<b>6.6</b>	<b>DISCUSSION ET CONCLUSION DE LA TROISIEME ETUDE</b>	<b>216</b>
<b>7</b>	<b>RESUME DES CONTRIBUTIONS, CONCLUSION GENERALE ET PERSPECTIVES</b>	<b>223</b>
<b>7.1</b>	<b>RESUME DES RESULTATS DE LA PREMIERE ETUDE</b>	<b>223</b>
<b>7.2</b>	<b>RESUME DES RESULTATS DE LA DEUXIEME ETUDE</b>	<b>228</b>
<b>7.3</b>	<b>RESUME DES RESULTATS DE LA TROISIEME ETUDE</b>	<b>230</b>
<b>7.4</b>	<b>CONCLUSION GENERALE</b>	<b>231</b>
<b>7.5</b>	<b>PERSPECTIVES</b>	<b>235</b>
<b>8</b>	<b>BIBLIOGRAPHIE</b>	<b>241</b>
<b>9</b>	<b>ANNEXES</b>	<b>253</b>
<b>9.1</b>	<b>SOMMAIRE ÉTENDU</b>	<b>255</b>
<b>9.2</b>	<b>TABLE DES FIGURES</b>	<b>258</b>
<b>9.3</b>	<b>TABLE DES TABLEAUX</b>	<b>264</b>
<b>9.4</b>	<b>CAHIER DES SUJETS DE L'EXP. 1.2</b>	<b>265</b>
<b>9.5</b>	<b>ARTICLES ET POSTERS</b>	<b>279</b>
9.5.1	EYES-FREE INTERACTION TECHNIQUES FOR DISTANCE ESTIMATION	280
9.5.2	ASSISTANCE TACTILE A LA LOCALISATION DE CIBLES PERIPHERIQUES POUR DES PERSONNES A VISION TUBULAIRE 300	
9.5.3	INTERACTIONS NON-VISUELLES POUR FACILITER LE RENDU D'UNE SEQUENCE D'INFORMATIONS	320
9.5.4	INFLUENCE DE LA PRESENTATION DES MOTS POUR DES LECTEURS DYSLEXIQUES	332
9.5.5	AN MDE-BASED FRAMEWORK TO SUPPORT THE DEVELOPMENT OF MIXED INTERACTIVE SYSTEMS	340

## 9.2 Table des figures

Figure 2-1 : Plus le seuil est bas, plus la partie du corps est « sensible » ; à gauche, seuil de perception de vibrations et de pressions ; à droite, seuil de discrimination d'une distance entre deux points de contact (« gap detection ») et seuil de discrimination d'un second point de contact autour d'un premier (« two-point discrimination ») ; source : (Johnson and Phillips 1981) _____	26
Figure 2-2 : exemples de moteur vibrant ; de la gauche vers la droite : VPM2, Samsung A400 vibration motors, LilyPad Vibe Boards et le C2 tactor et Vibrating "Eyeronman" Vest. _____	28
Figure 2-3 : exemple de dispositifs capables de produire un stimulus de pression ; de gauche vers la droite : PHANToM 6/6 DOF, PHANToM Omni 6/3 DOF, Dexmo F2, Mechanism of the Tactile Vest actuator (Nakamura and Jones 2003), « STReSS » (Pasquero and Hayward 2003) et un Haptic Glove. _____	29
Figure 2-4 : Illustration des dispositifs PHANToM™ (à gauche) et CyberGrasp™ (à droite). _____	31
Figure 2-5 : Dispositifs utilisant la modalité haptique de pression. A gauche, illustration d'un « Hands Omni haptic glove », à l'intérieur des coussinets se gonflent afin de reproduire par exemple, la prise en main d'un objet. A droite, illustration d'un HaptiQ, qui transmet une direction à l'aide de picots sous la main. _____	33
Figure 2-6 : Illustration du TéléTact II ( <a href="https://perso.limsi.fr/bellik/TeleTact/TeleTact.htm">https://perso.limsi.fr/bellik/TeleTact/TeleTact.htm</a> ) _____	43
Figure 2-7 : Schéma récapitulatif des trois modalités tactile : haptique, vibrotactile et tangible ; Au centre, leur caractéristique principale les distinguant entre elles. La barre noire indique une séparation importante des modalités et les barres en pointillées indiquent que la séparation est moins nette. Pour rappel, « blind walking » signifie « marche à l'aveugle » et « blind reaching » signifie « atteindre à l'aveugle ». _____	45
Figure 2-8 : Dimensions physiques et propriétés d'un son composé et d'un son pur. _____	46
Figure 2-9 : niveaux sonores (en dB) illustrés par des environnements sonores et les seuils de dangerosité, issue de <a href="http://audioson.wordpress.com">audioson.wordpress.com</a> . _____	47
Figure 2-10 : perception et durée d'un son ; la distinction entre deux sons. _____	48
Figure 2-11 : illustration d'une différence interaurale de temps (ITD) avec une source sonore à 90° sur l'azimut (i.e. à droite, flèche verte) ; la différence est symbolisée rapidement par la flèche rouge (qui traverse la tête) ; l'ITD est de 0,7 ms. _____	49
Figure 2-12 Champ audible chez l'Homme. Ce graphique montre l'intensité nécessaire à la perception (en dB) en fonction des fréquences (en Hz). Aux fréquences audibles, l'aire en vert clair délimitée par le seuil d'audition et le seuil d'inconfort (lignes noire sup. et inf.) définit le champ auditif, domaine de sensation auditive ou aire d'audition. L'aire en vert foncé est la zone conversationnelle et définit les sons utilisés pour la communication par la voix humaine (Unité INSERM en 1999 : « Promenade autour de la cochlée »). _____	53
Figure 2-13 : oreillette monophonique Bluetooth (sans fil)(source : <a href="https://fr.wikipedia.org/wiki/écouteur">https://fr.wikipedia.org/wiki/écouteur</a> ) _____	56
Figure 2-14 : Un casque audio des années 1970(source : <a href="https://fr.wikipedia.org/wiki/Casque_audio">https://fr.wikipedia.org/wiki/Casque_audio</a> ) _____	56
Figure 2-15 à gauche : schéma de coupe d'un hautparleur ; à droite : un hautparleur. (source : <a href="https://fr.wikipedia.org/wiki/Haut-parleur">https://fr.wikipedia.org/wiki/Haut-parleur</a> ) _____	56
Figure 2-16 Schéma récapitulatif des propriétés perceptibles d'un son. _____	65
Figure 2-17 : Illustration de transmissions concurrentes de plusieurs informations indépendantes : à gauche, deux messages vocaux en sortie, l'utilisateur devra percevoir correctement les deux messages qui s'entremêlent (au point de vue perceptif) ; à droite, deux informations « numériques » en entrée qui sont reçues en même temps par le système, mais affichées en séquence : « AB » ou « BA » ? Le système ne peut savoir quel caractère écrire en premier et encore moins les écrire en simultané. _____	68
Figure 2-18 : Illustrations de transmission synergique de plusieurs informations dépendantes : à gauche, un message vocal sonifié en sortie. L'utilisateur devra percevoir correctement les deux sous-unités pour comprendre l'ensemble de l'information. A droite, le système devra interpréter deux informations « numériques » en entrée qui sont reçues en même temps comme la commande « copier ». En bas, l'utilisateur pourra comprendre l'information par l'une ou l'autre modalité, car elles sont redondantes. _____	69
Figure 2-19 : Illustration d'une transmission alternée de plusieurs sous-unités d'information : l'utilisateur pointe un bouton avec sa souris (curseur) et dit la commande vocale « clic » pour ordonner au système de cliquer sur le bouton. _____	69

- Figure 2-20 : illustration du système SleeveAR, l'utilisateur effectue un mouvement, il ne peut pas lire les informations visuelles (bras droit) qui sont hors de son champ de vision car le sujet gardait la tête droite pour ce mouvement de rééducation, un guidage audio et haptique prend le relais. \_\_\_\_\_ 71
- Figure 2-21 : Illustration de la preuve de concept d'un système collaboratif de gestion de crise inspirée du GeoMIP. Les interactions permettent une transmission d'informations complémentaires en synergie ou en alternance. \_\_\_\_\_ 71
- Figure 2-22 : Illustration issue de (Liao et al. 2006) des interactions du « Pen-top » : à gauche, un utilisateur novice recevant des informations relatives à sa navigation ; au milieu, la même navigation mais dans le mode expert (donc moins d'indices) ; à droite, la recherche du mot « computer » dans un texte. \_\_\_\_\_ 72
- Figure 2-23 : Illustration issue de (Cassell et al. 1999) : « Rea le robot », en « pleine conversation » avec un participant. \_\_\_\_\_ 73
- Figure 3-1 A gauche, une INFORMATION MAPPÉE par une « technique d'interaction sans référence, passive, signal en sortie du système vers l'utilisateur, avec un effecteur capable de transmettre un signal audio de manière discrète (ponctuelle) et dont la force (intensité) du signal code une information » ; à droite, une INFORMATION MAPPÉE par une « technique d'interaction avec 3 références et active, signal en entrée du système de la part de l'utilisateur, avec un objet tangible capable de transmettre un « signal tactile » de manière continue et dont les pulsations du signal codent une information ; l'utilisateur recevra un feedback de son action, avec un effecteur capable de transmettre un signal auditif discret (ponctuel) basé sur un langage oral naturel (parole) ou abstrait (earcons) ». \_\_\_\_\_ 92
- Figure 3-2 : illustration des propriétés CASE adaptées à notre contexte de transmission d'information(s). Les flèches représentent des lignes temporelles, les formes géométriques représentent des modalités (des techniques d'interaction) et les lettres représentent des informations. 1 et 3, sont des transmissions séquentielles ; 2 et 4 sont des transmissions parallèles ; 1 et 2 sont des transmissions d'informations dépendantes avec des modalités différentes ; 3 et 4 sont des transmissions d'informations indépendantes. \_\_\_\_\_ 93
- Figure 3-3 : illustration du prototype palm gauge, à gauche le concept, à droite le prototype vu de haut. Exemple d'utilisation : « Je perçois la rotation d'une aiguille sous ma main, elle s'arrête puis revient à sa position d'origine, j'estime la valeur transmise » \_\_\_\_\_ 97
- Figure 3-4 : illustration de l'analyse de la technique Palm Gauge. \_\_\_\_\_ 97
- Figure 3-5 : Illustration de la technique « single bip ». Exemple d'utilisation : « J'entends un son pendant une certaine durée, puis plus rien. J'estime la valeur transmise » \_\_\_\_\_ 98
- Figure 3-6 : illustration de l'analyse de la technique Single Stimulus. \_\_\_\_\_ 98
- Figure 3-7 : illustration de la technique dual stimuli. Exemple d'utilisation : « je perçois un premier son d'une durée fixe, puis un second d'une durée variable, j'estime la valeur transmise par la partie variable ». \_\_\_\_\_ 99
- Figure 3-8 : illustration de l'analyse de la technique Dual Stimuli. \_\_\_\_\_ 100
- Figure 3-9 : Illustration du prototype "ClicOnMe" pris en main comme un chronomètre ; à gauche, le concept, à droite, le prototype tenu en main. Exemple d'utilisation : « J'appuis sur le bouton de la souris, j'attends. Je perçois un feedback au bout d'une certaine durée, j'estime la valeur transmise ». \_\_\_\_\_ 101
- Figure 3-10 : illustration de l'analyse de la technique ClicOnMe. \_\_\_\_\_ 101
- Figure 3-11 : illustration de la conception de la technique « pipe ». Exemple d'utilisation : « Je tourne doucement le cube vers l'avant. Lors de ma rotation, j'entends un bip, j'estime la valeur transmise ». \_\_\_\_\_ 102
- Figure 3-12 : illustration de l'analyse de la technique « Pipe ». \_\_\_\_\_ 103
- Figure 3-13 : illustration de la technique « volant ». Exemple d'utilisation « Je tourne doucement le volant à droite ou à gauche. Lors de ma rotation, j'entends un bip, j'estime la valeur transmise ». \_\_\_\_\_ 104
- Figure 3-14 : illustration de l'analyse de la technique Volant. \_\_\_\_\_ 104
- Figure 3-15 : illustration de la technique Pédale. Exemple d'utilisation : « J'incline doucement la pédale vers le bas. Lors de mon inclinaison, j'entends un bip, j'estime la valeur transmise ». \_\_\_\_\_ 105
- Figure 3-16 : illustration de l'analyse de la technique Pédale. \_\_\_\_\_ 106
- Figure 3-17 : illustration de la technique « Slider ». Exemple d'utilisation : « Je déplace doucement le slider vers l'extrémité opposée du rail. Lors de mon déplacement, j'entends un bip, j'estime la valeur transmise ». \_\_\_\_\_ 106
- Figure 3-18 : illustration de l'analyse de la technique Slider. \_\_\_\_\_ 107
- Figure 3-19 : Récapitulatif des huit techniques d'interaction générées. A gauche les passives, à droite les actives. Les traits représentent les designs de conception des techniques ; le trait épais met en valeur le fait que les

- techniques actives ont un design similaire sauf pour la présence de référence et le type de codage ; les cases non traversées représentent les possibilités de conception non explorées ici ; nous pouvons également voir les facteurs communs et spécifiques aux techniques. \_\_\_\_\_ 109
- Figure 4-1 : illustration des différences entre les techniques de base et leur alternative. A gauche les « Passive », la différence se situe au niveau de la modalité, à droite les « Active », la différence se situe au niveau de la modalité du feedback. \_\_\_\_\_ 118
- Figure 4-2 : les techniques sont équivalentes en termes d'efficacité (%succès) et ce, peu importe la tolérance à l'erreur. \_\_\_\_\_ 123
- Figure 4-3 : A gauche, l'erreur (précision) et l'écart-type (dispersion) pour chaque technique. Les techniques « ClicOnMe », « Pipe » et « SingleStim. » semblent graphiquement moins précises et plus dispersées que les autres techniques. Les lignes en tiret montrent les limites -5 et +5. A droite, les valeurs des moyennes et des écarts-types, pour chacune des techniques, ordonnés de manière croissante selon les écarts-types. \_\_\_\_\_ 125
- Figure 4-4 : A gauche (A), Scores de satisfaction pour chaque technique. La ligne grise à 72 sur l'axeY, représente trois catégories SUS : bonne (au-dessus), acceptables (sur), et non acceptables (en dessous). Trois techniques sont considérées comme étant bonnes ou acceptables : « Slider », « Pédale » et « PalmGauge ». A droite (B), l'ordre de préférence –normalisé– des techniques par les utilisateurs, sachant qu'une préférence de 1 est la meilleure et 0 la pire. Notons que les trois meilleures techniques (« PalmGauge », « Slider » et « Pédale ») sont sur le podium de la satisfaction (A) et de la préférence (B), tandis que les pires techniques sont « ClicOnMe », « SingleStim. » et « Pipe » pour (A) et (B). Rappelons que ces trois dernières techniques sont « SansRéférence ». \_\_\_\_\_ 126
- Figure 4-5 : Régression linéaire de la moyenne de l'erreur (12 sujets) par technique. Les points noirs représentent la moyenne de l'erreur de chaque essai pour tous les sujets pour une technique. Notons que les techniques (« ClicOnMe », « Pipe » et « SingleStim. ») qui dérivent significativement n'ont pas de références explicites. Les [\*] montrent les dérives significatives (\*\*  $p < 0,01$  ; \*\*\*,  $p < 0,001$ ) et les chiffres en dessous des régressions linéaires montrent leur coefficient. \_\_\_\_\_ 127
- Figure 4-6 : A gauche, l'erreur (précision) et l'écart-type (dispersion) pour chaque groupe de techniques (Active/Passive). Les lignes montrent les limites -5,0 et +5 ; les [\*] montrent les différences significatives (\*\*\*,  $p < 0,001$ ). A droite, les valeurs des moyennes et des écarts-types, pour les groupes « Active » et « Passive », ordonnés de manière croissante selon les écarts-types. \_\_\_\_\_ 128
- Figure 4-7 : Régression linéaire de la moyenne de l'erreur (12 sujets) par catégorie de technique Active/Passive. Les points noirs représentent la moyenne de l'erreur de chaque essai pour tous les sujets pour un groupe de techniques. Les [\*] montrent la p.value (\*\*  $p < 0,01$  ; \*\*\*,  $p < 0,001$ ) et les chiffres en dessous des régressions montrent leur coefficient. \_\_\_\_\_ 129
- Figure 4-8 : A gauche, l'erreur (précision) et l'écart-type (dispersion) pour chaque catégorie de technique AvecRéférences/SansRéférence. Les lignes montrent les limites -5,0 et +5 ; les [\*] montrent les différences significatives (\*\*\*,  $p < 0,001$ ). A droite, les valeurs des moyennes et des écarts-types, pour les groupes de techniques AvecRéférences et SansRéférence, ordonnés de manière croissante selon les écarts-types. \_\_\_\_\_ 130
- Figure 4-9 : Régression linéaire de la moyenne de l'erreur (12 sujets) par catégorie de technique AvecRéférences/SansRéférence. Les points noirs représentent la moyenne de l'erreur de chaque essai pour tous les sujets pour un groupe de techniques. Les [\*] montrent la p.value (\*\*\*,  $p < 0,001$ ) et les chiffres en dessous des régressions montrent leur coefficient. \_\_\_\_\_ 131
- Figure 4-10 : illustration de la technique « Règle lisse ». Exemple d'utilisation : « je fais une double tape, j'entends un bip, je déplace doucement mon doigt vers la droite. Lors de mon déplacement, j'entends un bip, j'estime la valeur transmise ». \_\_\_\_\_ 139
- Figure 4-11 : illustration de l'analyse de la technique Règle-Lisse. \_\_\_\_\_ 139
- Figure 4-12 : illustration de la technique « Règle graduée ». Exemple d'utilisation : « je fais une double tape, j'entends un bip, je déplace doucement mon doigt vers la droite en passant sur les clous (réf. physiques). Lors de mon déplacement, j'entends un bip, j'estime la valeur transmise, m'aidant des références sur la règle ». \_\_\_\_\_ 140
- Figure 4-13 : illustration de l'analyse de la technique Règle-Graduée. \_\_\_\_\_ 141
- Figure 4-14 : illustration de la technique BipSolo. Exemple d'utilisation : « J'entends un son pendant une certaine durée, puis plus rien. J'estime la valeur transmise » \_\_\_\_\_ 141
- Figure 4-15 : illustration de l'analyse de la technique BipSolo. \_\_\_\_\_ 142

- Figure 4-16 : illustration de la technique « BipMultiple ». Exemple d'utilisation : « j'entends un son, dans lequel il y a plusieurs bips, chaque bip marque une vingtaine, j'estime la valeur transmise (dans l'exemple, 100, car 5 bips) ». \_\_\_\_\_ 143
- Figure 4-17 : illustration de l'analyse de la technique BipMultiple. \_\_\_\_\_ 143
- Figure 4-18 : A gauche, l'erreur (précision) et l'écart-type (dispersion) pour chaque groupe de techniques (Active/Passive). Les lignes montrent les limites -5,0 et +5 ; les [\*] montrent les différences significatives (\*\*,  $p < 0,01$ ). A droite, les valeurs des moyennes et des écarts-types, pour les groupes « Active » et « Passive », ordonnées de manière croissante selon les écarts-types. \_\_\_\_\_ 145
- Figure 4-19 : Régression linéaire de la moyenne de l'erreur (12 sujets) par catégorie de technique Active/Passive. Les points noirs représentent la moyenne de l'erreur de chaque essai pour tous les sujets pour un groupe de techniques. Les [\*] montrent la p.valeur (\*\*  $p < 0,01$  ; \*\*\*,  $p < 0,001$ ) et les chiffres en dessous des régressions montrent leur coefficient. \_\_\_\_\_ 146
- Figure 4-20 : A gauche, l'erreur (précision) et l'écart-type (dispersion) pour chaque catégorie de technique AvecRéférences/SansRéférence. Les lignes montrent les limites -5,0 et +5 ; les [\*] montrent les différences significatives (\*\*,  $p < 0,01$ ). A droite, les valeurs des moyennes et des écarts-types, pour les groupes de techniques AvecRéférences et SansRéférence, ordonnés de manière croissante selon les écarts-types. \_\_\_\_\_ 147
- Figure 4-21 : Régression linéaire de la moyenne de l'erreur (12 sujets) par catégorie de technique AvecRéférences/SansRéférence. Les points noirs représentent la moyenne de l'erreur de chaque essai pour tous les sujets pour un groupe de techniques. Les [\*] montrent la p.valeur (\*\*\*,  $p < 0,001$ ) et les chiffres en dessous des régressions montrent leur coefficient. \_\_\_\_\_ 148
- Figure 4-22 : A gauche, l'erreur (précision) et l'écart-type (dispersion) pour chaque technique. Les techniques « +SR » semblent graphiquement moins précises et plus dispersées que les techniques « +AR ». Les lignes en pointillées montrent les limites -5 et +5. A droite, les valeurs des moyennes et des écarts-types, pour chacune des conditions, ordonnés de manière croissante selon les écarts-types. \_\_\_\_\_ 149
- Figure 4-23 : A gauche (A), Scores de satisfaction pour chaque technique. La ligne rouge à 70 sur l'axeY, représente trois catégories SUS : bonne (au-dessus), acceptables (sur), et non acceptables (en dessous). Une seule technique est considérée comme étant bonne : « A+AR ». A droite (B), l'ordre de préférence –normalisé– des techniques par les utilisateurs, sachant qu'une préférence de 1 est la meilleure et 0 la pire. Notons que la meilleure technique est « A+AR » est sur le podium de la satisfaction (A) et de la préférence (B), tandis que la pire technique est « P+SR » pour (A) et (B). \_\_\_\_\_ 150
- Figure 4-24 : Régression linéaire de la moyenne de l'erreur (12 sujets) par technique. Les points noirs représentent la moyenne de l'erreur de chaque essai pour tous les sujets pour une technique. Notons que les techniques sans référence dérivent significativement. Les [\*] montrent les dérives significatives (\*\*  $p < 0,01$  ; \*\*\*,  $p < 0,001$ ) et les chiffres en dessous des régressions linéaires montrent leur coefficient. \_\_\_\_\_ 151
- Figure 4-25 : les techniques « SansRéférence » nécessitent une tolérance trois fois plus grande pour obtenir un pourcentage de succès similaire à ceux de techniques « AvecRéférences ». \_\_\_\_\_ 152
- Figure 5-1 : Exemple d'un itinéraire de 3 segments (couples) ; à gauche la version Google Maps et à droite la version simplifiée transmise par nos techniques, soit 6 valeurs : Droite-90 ; Gauche-50 ; Droite-6 \_\_\_\_\_ 162
- Figure 5-2 : illustration de la technique TTS. Exemple d'utilisation : liste des segments d'un itinéraire de taille trois prononcée par une synthèse vocale. Les «...» indiquent que la taille de l'itinéraire peut être augmentée. \_\_\_\_\_ 166
- Figure 5-3 : illustration de l'analyse de la technique TextToSpeech (TTS). \_\_\_\_\_ 167
- Figure 5-4 : illustration de la technique TTS-Lat. Exemple d'utilisation : liste des segments d'un itinéraire de taille trois (Droite-90 ; Gauche-50 ; Droite-6) prononcée par une synthèse vocale dont la source sonore est spatialisée. Les «...» indiquent que la taille de l'itinéraire peut être augmentée. \_\_\_\_\_ 168
- Figure 5-5 : illustration de l'analyse de la technique TextToSpeech Latéralisé (TTS-Lat). \_\_\_\_\_ 169
- Figure 5-6 : illustration de la technique PG-TTS. Exemple d'utilisation : liste des segments d'un itinéraire de taille trois (Droite-90 ; Gauche-50 ; Droite-6), où les directions sont données par une direction tactile et les distances par synthèse vocale. Les «...» indiquent que la taille de l'itinéraire peut être augmentée. \_\_\_\_\_ 169
- Figure 5-7 : illustration de l'analyse de la technique Palm Gauge-TextToSpeech (PG-TTS). \_\_\_\_\_ 170
- Figure 5-8 : illustration de la technique Mélodie. Exemple d'utilisation : liste des segments d'un itinéraire de taille trois (Droite-90 ; Gauche-50 ; Droite-6) sous la forme d'une mélodie. La note et sa latéralisation indique la direction. Sa durée indique la distance. Les «...» indiquent que la taille de l'itinéraire peut être augmentée. \_\_\_\_\_ 171

Figure 5-9 : illustration de l'analyse de la technique Mélodie.	172
Figure 5-10 : illustration du protocole pour un bloc.	173
Figure 5-11 : Moyenne du nombre d'éléments correctement restitués (R1) par taille (5 à gauche et 7 à droite) et par technique. Nous pouvons voir que les moyennes de réponses correctes pour MELODIE et TTS étaient proches de 7. Par contre les moyennes pour PG-TTS et TTS-Lat étaient supérieures, avec un écart-type qui passait la barre des 9 éléments corrects.	175
Figure 5-12 : Efficacité des techniques chez les non-DV, selon deux tailles (5 segments à gauche et 7 à droite). (Wilcoxon signed rank test ; * : $p < 0,05$ ).	176
Figure 5-13 : Robustesse de la mémorisation selon la taille de l'itinéraire (5 segments à gauche et 7 à droite). (Wilcoxon signed rank test, * : $p < 0,05$ ).	177
Figure 5-14 : Ordre moyen de préférence des techniques des sujets Non-DV ; Mélodie et TTS-Lat sont préférées à TTS et PG-TTS (Wilcoxon signed rank test ; * : $p < 0,05$ )	179
Figure 5-15 : Efficacité des techniques chez les DV mesurée lors du premier rappel, pour des itinéraires de 5 (gauche) ou 7 (droite) segments. Les lignes de tirets noirs, épais, illustrent la limite de l'empan mnésique à $7 \pm 2$ éléments.	180
Figure 5-16 : Efficacité des techniques pour les deux populations. Le graphique illustre le pourcentage du nombre d'itinéraires corrects selon les directions (haut) et les distances (bas) observé lors de la première restitution. Les itinéraires de taille 5 sont illustrés à gauche, et ceux de taille 7 à droite. Les lignes de tirets indiquent un seuil d'efficacité à 75%.	180
Figure 5-17 : Robustesse des techniques pour les deux populations. Le graphique illustre le pourcentage du nombre d'itinéraires conformes selon les directions (haut) et les distances (bas). Les itinéraires de taille 5 sont illustrés à gauche, et ceux de taille 7 à droite. Les lignes de tirets indiquent un seuil de robustesse à 75%.	181
Figure 5-18 : Moyenne des items correctement restitués (R1) par les Non-DV, selon la position du segment dans l'itinéraire. Les graphiques illustrent les directions (rouge) et les distances (bleu) par technique et par taille d'itinéraire (5 en haut et 7 en bas). Seule TTS-Lat, pour la taille 5, résiste à l'effet de position. Rappelons que Mélodie est la seule technique utilisant le codage DURÉE pour la distance, les autres techniques utilisent la parole.	182
Figure 6-1 : Illustration d'une vue tubulaire : à gauche, « Le déjeuner des canotiers » par Pierre-Auguste Renoir en 1880 ; à droite ce même tableau vu avec une vision tubulaire schématisée.	188
Figure 6-2 : illustration de la technique de superposition de (Vargas-Martin and Peli 2002) : un utilisateur regarde un visage dans une salle et celle-ci est superposée dans le champs de vision résiduel de l'utilisateur avec de la réalité augmentée.	189
Figure 6-3 : illustration de l'encodage d'une image en signaux auditifs par «The vOICe», issue de (Auvray et al. 2007) ; L'encodage est basé sur trois propriétés du son, liées à trois propriétés d'une image : fréquence, durée, intensité (resp. ligne de pixels, colonne de pixel, nuance de gris). Les utilisateurs devaient « estimer » une image avec du son.	190
Figure 6-4 : illustration des prototypes issues de (Bach-Y-Rita et al. 1969) (gauche), (P Bach-y-Rita et al. 1998)(droite).	190
Figure 6-5 : illustration issue de (Lehtinen et al. 2012). La main des utilisateurs est continuellement guidée vers une lettre B parmi un ensemble de lettres P distractives avec deux messages : l'un signifiant « UP », pour signifier à l'utilisateur qu'il devait monter sa main et, l'autre « RIGHT » pour qu'il décale sa main vers la droite.	192
Figure 6-6 : illustration des deux systèmes de coordonnées indiquant une cible rouge : à gauche (A), le système dit cartésien, l'azimut représente la coordonnée 1 (bleu) et l'élévation représente la coordonnée 2 (orange) ; à droite (B), le système dit polaire, l'inclinaison représente la coordonnée 1 (violet) et la distance radiale représente la coordonnée 2 (vert).	195
Figure 6-7 : Codage tactile d'une même cible (point rose) avec quatre techniques différentes. Deux codes tactiles ont été utilisés : le Singlestimulus (A, B) et le train de pulsations (C, D). Les coordonnées de la cible sont transmises dans les SC cartésien (A, C) ou polaire (B, D). Les deux coordonnées sont rendues par deux stimulations successives (S1, S2) représentées avec des barres grises. Dans cette expérience, la position de la cible était choisie de façon pseudo-aléatoire dans toute la zone bleue.	196

- Figure 6-8 : illustration de la technique C/Si. Exemple illustré : la cible est 29,25° à droite et 14,25° en haut. La succession des stimuli était : flèche pointant à droite sur l'écran, puis une vibration de 975 ms, puis 500 ms plus tard, flèche pointant en haut, puis une vibration de 475 ms. \_\_\_\_\_ 198
- Figure 6-9 : illustration de l'analyse de la technique Singlestimulus dans un SC Cartésien (C/Si). \_\_\_\_\_ 199
- Figure 6-10 : illustration de la technique C/Tr. Exemple illustré : la cible est 27° à droite et 18° en haut. La succession des stimuli était : flèche pointant à droite sur l'écran, puis un train de trois pulsations, puis 500 ms plus tard, flèche pointant en haut, puis un train de deux pulsations. \_\_\_\_\_ 200
- Figure 6-11 : illustration de l'analyse de la technique Train de pulsations dans un SC Cartésien (C/Tr). \_\_\_\_\_ 200
- Figure 6-12 : illustration de la technique P/Si. Exemple illustré : la cible est 29,25° à droite et 14,25° en haut. La succession des stimuli était : barre haute sur l'écran, puis une vibration de 975 ms, puis 500 ms plus tard, un cercle rouge, puis une vibration de 475 ms. \_\_\_\_\_ 201
- Figure 6-13 : illustration de l'analyse de la technique Singlestimulus dans un SC Polaire (P/Si). \_\_\_\_\_ 202
- Figure 6-14 : illustration de la technique P/Tr. Exemple illustré : la cible est 27° d'inclinaison et 18° du centre (distance radiale). La succession des stimuli était : barre haute sur l'écran, puis un train de trois pulsations, puis 500 ms plus tard, un cercle rouge, puis un train de deux pulsations. \_\_\_\_\_ 203
- Figure 6-15 : illustration de l'analyse de la technique Train de pulsations dans un SC Polaire (P/Tr) \_\_\_\_\_ 203
- Figure 6-16 : Photos d'un sujet voyant équipé avec l'interface tactile au poignet et le casque de réalité virtuelle permettant de simuler la vision tubulaire. Les mouvements de la tête sont enregistrés par un système infrarouge, visible en A. Dans l'expérience 3.1, les sujets ne disposent que d'un seul vibreur, placé sur le poignet par une bande de fixation élastique et cohésive (A). Dans les expériences 3.2 et 3.3, quatre vibreurs sont placés sur le poignet gauche (B). Le quatrième vibreur, placé sur la partie antérieure du poignet n'est pas visible sur la photo. \_\_\_\_\_ 205
- Figure 6-17 : efficacité (A) et satisfaction (B) pour les quatre techniques conçues. A) L'erreur moyenne de pointage pour chacune des techniques est représentée par des barres de couleur. La moyenne de l'erreur de pointage par sujet est représentée avec des points blancs. La technique C/Tr montre l'erreur moyenne la plus faible (8,2°) avec l'écart-type le plus faible (1,2°). B) Le score moyen de satisfaction (SUS) pour chacune de ces techniques est représenté par des barres de couleur. Le score de satisfaction par sujet est représenté avec des points blancs. La technique C/Tr reçoit le score le plus élevé avec 81,3±12,0. \_\_\_\_\_ 207
- Figure 6-18 : A – Portion de l'environnement visuel. La cible rose apparaît en bas à gauche. B – Exemple de la simulation artificielle de la vision tubulaire de 10°. \_\_\_\_\_ 209
- Figure 6-19 : illustration de la technique d'assistance à la recherche visuelle. Exemple pour une cible située à 27° à droite et à 18° en haut : 3 vibrations sur le vibreur droit, puis 500 ms plus tard, 2 vibrations sur le vibreur haut. \_\_\_\_\_ 210
- Figure 6-20 : illustration de l'analyse de la technique d'assistance à la recherche visuelle, basée sur la technique de l'exp. 3.1 C/Tr. \_\_\_\_\_ 211
- Figure 6-21 : Exemples de mouvements de tête observés pendant une recherche visuelle dans les conditions contrôle, sans indice (A, B, C) et avec la technique d'assistance, avec indice, (D, E, F). A - Balayage sur l'axe vertical. B – Carte de densité d'une recherche avec un balayage principalement sur l'axe horizontal. C – Carte de densité d'une exploration en cercles concentriques. D – Mouvements ordonnés le long de l'axe horizontal puis vertical suivant l'ordre de transmission des deux indices tactiles. E – Carte de densité d'un mouvement commençant essentiellement par l'axe horizontal. F- Carte de densité d'un mouvement intégré. Ici, le sujet a été capable d'intégrer les deux coordonnées avant de bouger, puis de déplacer la tête en faisant des mouvements obliques vers les cibles. Chaque carte de densité représente les mouvements de tête d'un sujet. La position de la cible et les mouvements de tête ont été normalisés avec une transformation symétrique sur l'axe horizontal puis vertical, en plaçant les cibles dans le coin supérieur droit. Le rouge indique des densités élevées et le bleu des densités nulles. \_\_\_\_\_ 213
- Figure 6-22 : Performance de la recherche visuelle (A – temps de recherche et B- distance parcourue) avec et sans indices tactiles. En gris, la condition contrôle sans indice ; en cyan, la technique d'assistance, avec indices. \_\_\_\_\_ 214
- Figure 6-23 : Performance de la recherche visuelle en temps de recherche, avec et sans indices tactiles. En gris, la condition contrôle sans indices ; en cyan, la technique d'assistance, avec indices. \_\_\_\_\_ 216

### 9.3 Table des tableaux

Tableau 3-1 : exemples de dispositifs utilisés dans des techniques d'interaction de l'état de l'art, avec la partie du corps associée. _____	83
Tableau 3-2 : exemples de signaux continus et discrets. _____	87
Tableau 3-3 : tableau combinatoire ( $\cap$ ) afin de lever toute ambiguïté sur les termes : active, passive et les types d'utilisation. Notons l'absence de proposition dans la case [-/OUT $\cap$ Active], cette case peut être remplie par des interactions non connues des utilisateurs. Prenons l'exemple de l'interaction suivante : « l'utilisateur entre dans une pièce ; plus l'utilisateur va vers le fond de la pièce, plus l'éclairage sera amplifié ; et inversement » ; [??] Si l'utilisateur n'a pas conscience de cette interaction, alors pour lui et du point de vue conception, cela sera une [-/OUT $\cap$ Active], par contre, s'il en a conscience cela sera une [IN/FEEDBACK $\cap$ Active]. _____	89
Tableau 3-4 : tableau de synthèse composé de huit paramètres de base _____	91
Tableau 3-5 : synthèse présentant les choix de conception pour les huit techniques d'interaction. AR signifie AvecRéférences, 0,1,2,3 = le nombre de référence et le « + » signifie plus de 3. _____	96
Tableau 3-6 : lois pour diverses sensations/perceptions : $S = k.I^a$ , (voir Loi de Stevens). _____	110
Tableau 4-1 : Synthèse des différences entre les techniques de base et leur alternative. Notons que seule PalmGauge n'a pas d'alternative. _____	117
Tableau 4-2 : tableau de synthèse des 15 techniques d'interaction de l'expérience 1.1 _____	119
Tableau 4-3 : résultats du test post-hoc pour l'efficacité. Notons que nous effectuons nos tests statistiques à partir de la moyenne de l'erreur relative (signée), or, il se trouve que le « Pipe » et le « PalmGauge » ont une moyenne similaire et des écarts-types très différents, pourtant dans ce tableau, ils ne sont pas significativement différents. Ce test ne montre donc pas complètement les différences entre les techniques. _____	124
Tableau 4-4 : résultats du test post-hoc pour la satisfaction (score SUS). Notons, par exemple, que le « Pipe » n'est pas différent du « ClicOnMe », ni du « SingleStim. », ces trois techniques sont « SansRéférence ». _____	126
Tableau 4-5 : synthèse des résultats pour les huit techniques que nous avons conçu et évalué. _____	132
Tableau 4-6 : résultats du test post-hoc pour l'efficacité avec la correction de Bonferroni. Seul le groupe « A+AR » est significativement différent des autres. _____	149
Tableau 4-7 : résultats du test post-hoc pour la satisfaction (score SUS). _____	150
Tableau 4-8 : résultats du test post-hoc pour la préférence utilisateur (classement). Notons que « A+AR » est significativement différent de « P+SR » _____	150
Tableau 4-9 : comparatif de quatre techniques en se basant sur les résultats de « SingleStimulus ». Les mots soulignés indiquent un effet négatif. _____	154
Tableau 4-10 : comparatif de trois techniques en se basant sur les résultats de « Slider. ». Les mots soulignés indiquent un effet négatif. _____	155
Tableau 4-11 : Synthèse de l'exp. 1.2. Pour chaque combinaison, son ordre selon cinq mesures est donné. La moyenne des ordres donne la puissance d'impact des combinaisons. Sachant que plus la « puissance d'impact » est grande, plus la technique sera efficace et satisfaisante comparée à une technique d'une puissance inférieure. _____	157
Tableau 4-12 : Classement des 12 techniques en fonction de leur « puissance d'impact ». Notons que les techniques de l'exp. 1.1 gardent leur hiérarchie. _____	157
Tableau 5-1 : synthèse de plusieurs combinaisons de techniques et aspects temporels pour transmettre un couple direction-distance. Les symboles entre parenthèses indiquent les différents aspects temporels de la transmission d'information : (E) = exclusif ; (A) = Alterné ; (C) = concurrent ; (S) = synergique. La technique « TextToSpeech » est monomodale et utilise exclusivement la parole pour transmettre le couple direction-distance, c'est notre technique contrôle. La modalité « sonification » repose sur l'utilisation des propriétés d'un son pour transmettre une information, telles que la durée (BipSolo) ou la fréquence (Mélodie) ou encore la spatialisation de la parole (TTS-Lat). Les tactons sont des messages abstraits vibrotactiles ou de mouvements saccadés (entrecoupés). _____	165
Tableau 7-1 : Classement des 12 techniques de la première étude, en fonction de leur « puissance d'impact » ; avec en rouge celles qui violent la règle. A+AR > P+AR > A+SR > P+SR. _____	225

Tableau 7-2 : Classement des 12 techniques de la première étude, en fonction de leur « puissance d'impact » ; avec les règles R1 et R2 qu'elles suivent. \_\_\_\_\_ 226

#### 9.4 Cahier des sujets de l'exp. 1.2

# **Cahier d'expérience du sujet N°**



## FORMULAIRE D'INFORMATION ET DE CONSENTEMENT

### EXPERIENCE DE MEMORISATION D'ITINERAIRE (MODALITE NON VISUELLE)

#### IDENTIFICATION

**Doctorant** responsable du projet : **Damien APPERT**

#### BUT GÉNÉRAL DU PROJET ET DIRECTION

Vous êtes invité(e) à prendre part à une expérience portant sur la mémorisation d'itinéraire. Vous allez utiliser quatre techniques d'interaction différentes pour mémoriser les itinéraires.

#### PROCÉDURE(S) OU TÂCHES DEMANDÉES AU PARTICIPANT

Votre participation consiste à tester quatre techniques différentes pour mémoriser des itinéraires. Vos réponses seront enregistrées et l'entrevue sera filmée (éventuellement. Nous vous le préciserons à l'oral) avec votre permission et prendra environ 2 x 1 heure de votre temps. Toutes les données sont anonymes.

#### AVANTAGES ET RISQUES

Votre participation contribuera à l'avancement des connaissances dans les domaines de l'Interaction-Homme-Machine. Il n'y a pas de risque d'inconfort important associé à votre participation à cette rencontre. Vous pouvez arrêter quand vous voulez sans donner de justifications.

#### ANONYMAT ET CONFIDENTIALITÉ

Il est entendu que les renseignements recueillis lors de l'entrevue sont confidentiels et que seuls, le responsable du projet et ses directeurs de recherche, **C. JOUFFRAIS** et **E. DUBOIS**, auront accès à votre enregistrement et au contenu de sa transcription. Le matériel de recherche (enregistrement numérique et transcription codés) ainsi que votre formulaire de consentement seront conservés séparément sous clé par le responsable du projet pour la durée totale du projet. Les enregistrements ainsi que les formulaires de consentement seront détruits 5 ans après les dernières publications.

#### PARTICIPATION VOLONTAIRE

Votre participation à ce projet est volontaire. Cela signifie que vous acceptez de participer au projet sans aucune contrainte ou pression extérieure, et que par ailleurs vous être libre de mettre fin à votre participation en tout temps au cours de cette recherche. Dans ce cas, les renseignements vous concernant seront détruits. Votre accord à participer implique également que vous acceptez que le responsable du projet puisse utiliser aux fins de la présente recherche (articles, mémoire, essai ou thèse, conférences et communications scientifiques) les renseignements recueillis à la condition qu'aucune information permettant de vous identifier ne soit divulguée publiquement à moins d'un consentement explicite de votre part.

#### COMPENSATION FINANCIÈRE

Votre participation à ce projet est offerte gratuitement.

### **DES QUESTIONS SUR LE PROJET OU SUR VOS DROITS ?**

Vous pouvez contacter le responsable du projet Damien Appert, au numéro 05 61 55 7404 pour des questions additionnelles sur le projet.

### **REMERCIEMENTS**

Votre collaboration est importante à la réalisation de ce projet et nous tenons à vous en remercier.

### **SIGNATURES :**

Je reconnais avoir lu le présent formulaire de consentement et consens volontairement à participer à ce projet de recherche. Je reconnais aussi que le responsable du projet a répondu à mes questions de manière satisfaisante et que j'ai disposé suffisamment de temps pour réfléchir à ma décision de participer. Je comprends que ma participation à cette recherche est totalement volontaire et que je peux y mettre fin en tout temps, sans pénalité d'aucune forme, ni justification à donner. Il me suffit d'en informer le responsable du projet.

Signature du participant :

Date :

Nom (lettres moulées) et coordonnées (**@mail**) :

Je déclare avoir expliqué le but, la nature, les avantages, les risques du projet et avoir répondu au meilleur de ma connaissance aux questions posées.

Signature du responsable du projet :

Date :

Nom (lettres moulées) et coordonnées (**@mail**) :

**Damien APPERT, [damien.appert@irit.fr](mailto:damien.appert@irit.fr)**



**TTS-TTS** : Donne la direction puis la distance avec une synthèse vocale.

COMMENTAIRE :

**TTS-Latéralisée** : Donne une distance avec une synthèse vocale. Le son est à gauche pour la direction gauche, à droite pour la direction droite et avec les deux oreilles pour En Face.

COMMENTAIRE :

**Mélodie** : Une séquence de Sons va être jouée. Il y a un son différent pour chaque direction (Gauche/Droite/En Face). Le son est à gauche pour la direction gauche, à droite pour la direction droite et avec les deux oreilles pour En Face. **Attention : la durée d'un son correspond à la distance. Nous allons faire quelques essais avec des distances uniquement pour nous assurer que vous avez bien compris la consigne.**

**Essais à 3 puis 5. Plusieurs fois avec feedback. Leur rappeler qu'ils peuvent chanter la mélodie avant de répondre.**

COMMENTAIRE :

**PG-TTS** : La direction est transmise par l'angle d'une aiguille dans la paume de la main et la distance est exprimée par une synthèse vocale.

COMMENTAIRE :

**Ordre de préférence des techniques d'interaction :**

**PROTOCOLE**

Cette expérience vise à évaluer quatre techniques permettant de mémoriser un itinéraire les yeux bandés.

Un itinéraire peut être composé de 5 ou 7 segments avec pour chacun, une direction et une distance. Il y a trois possibilités pour la direction : Gauche, Droite et En Face. **La distance peut varier de 1 à 100 mètres.**

Après une phase de familiarisation pour chaque technique, nous ferons une série d'itinéraires avec 5 segments (donc 5 directions et 5 distances à mémoriser), puis une série d'itinéraires avec 7 segments (donc 7 directions et 7 distances à mémoriser).

Nous vous demanderons de restituer cet itinéraire **deux fois**. Une première fois, juste après l'avoir reçu, puis une seconde fois quelques secondes plus tard. Cependant, entre la première et la seconde fois, nous vous demanderons de réaliser une tâche de comptage à rebours.

A la fin de chaque bloc, vous répondrez à un questionnaire puis vous pourrez donner votre avis sur la technique que vous venez d'utiliser. A la fin des quatre blocs, vous pourrez classer les 4 techniques par ordre de préférence.

**Je vous rappelle que se sont les techniques d'interaction qui sont évaluées et pas vous.**

Vous pourrez poser toutes les questions que vous voulez entre les tests. Par contre, je vous demanderai de vous concentrer, d'éteindre votre téléphone et de ne pas discuter pendant les tests.

Afin de vous familiariser avec le mode de restitution, je vais vous présenter un exemple d'itinéraire et le tableau correspondant.

**FEUILLE DE TEST ... 1er PASSAGE**

	I	II	III	IV	V	
	DIR	DIS	DIR	DIS	DIR	DIS
<b>A</b>						
<b>B</b>						
<b>C</b>						
<b>D</b>						
<b>E</b>						
<b>F</b>						



**FEUILLE DE TEST --- 2ieme PASSAGE**

	I	II	III	IV	V	
	DIR	DIS	DIR	DIS	DIR	DIS
<b>A</b>						
<b>B</b>						
<b>C</b>						
<b>D</b>						
<b>E</b>						
<b>F</b>						



TTS\_5.4


TTS\_5.3


TTS\_5.2


TTS\_5.1


TTS\_7.4


TTS\_7.3


TTS\_7.2


TTS\_7.1


MELODIE\_5.4


MELODIE\_5.3


MELODIE\_5.2


MELODIE\_5.1


MELODIE\_7.4


MELODIE\_7.3


MELODIE\_7.2


MELODIE\_7.1


TTS\_Lat.\_5.4


TTS\_Lat.\_5.3


TTS\_Lat.\_5.2


TTS\_Lat.\_5.1


TTS\_Lat.\_7.4


TTS\_Lat.\_7.3


TTS\_Lat.\_7.2


TTS\_Lat.\_7.1


PG-TTS\_5.4


PG-TTS\_5.3


PG-TTS\_5.2


PG-TTS\_5.1


PG-TTS\_7.4


PG-TTS\_7.3


PG-TTS\_7.2


PG-TTS\_7.1


TTS\_5\_4\_2ieme\_passage



TTS\_5\_3\_2ieme\_passage

TTS\_5\_2\_2ieme\_passage


TTS\_5\_1\_2ieme\_passage


TTS\_7\_4\_2ieme\_passage


TTS\_7\_3\_2ieme\_passage


TTS\_7\_2\_2ieme\_passage


TTS\_7\_1\_2ieme\_passage


MELODIE\_5\_4\_2ieme\_passage


MELODIE\_5\_3\_2ieme\_passage


MELODIE\_5\_2\_2ieme\_passage


MELODIE\_5\_1\_2ieme\_passage


MELODIE\_7\_4\_2ieme\_passage


MELODIE\_7\_3\_2ieme\_passage


MELODIE\_7\_2\_2ieme\_passage


MELODIE\_7\_1\_2ieme\_passage


+ NASA\_TLX

+ NASA\_TLX

TTS\_Lat. \_ 5 . 4\_2ieme\_passage


+ NASA\_TLX

TTS\_Lat. \_ 5 . 3\_2ieme\_passage


TTS\_Lat. \_ 5 . 2\_2ieme\_passage


TTS\_Lat. \_ 5 . 1\_2ieme\_passage


TTS\_Lat. \_ 7 . 4\_2ieme\_passage


TTS\_Lat. \_ 7 . 3\_2ieme\_passage


TTS\_Lat. \_ 7 . 2\_2ieme\_passage


TTS\_Lat. \_ 7 . 1\_2ieme\_passage


PG-TTS \_ 5 . 4\_2ieme\_passage


+ NASA\_TLX

PG-TTS \_ 5 . 3\_2ieme\_passage


PG-TTS \_ 5 . 2\_2ieme\_passage


PG-TTS \_ 5 . 1\_2ieme\_passage


PG-TTS \_ 7 . 4\_2ieme\_passage


PG-TTS \_ 7 . 3\_2ieme\_passage


PG-TTS \_ 7 . 2\_2ieme\_passage


PG-TTS \_ 7 . 1\_2ieme\_passage


## 9.5 Articles et posters

### Sommaire :

#### En soumission :

- Appert, D., Dubois, E., Truillet, P., and Jouffrais, C. Eyes-free interaction techniques for distance estimation. En soumission.

#### Liste des publications : (<https://cloud.irit.fr/index.php/s/V7Sj1V5ffcg9q1F>)

1. Appert, D., Camors, D., Durand, J.-B., and Jouffrais, C. Tactile Cues for Improving Target Localization in Subjects with Tunnel Vision. *Proceedings of the 27th Conference on L'Interaction Homme-Machine*, ACM (2015), 6:1–6:10.
2. Appert, D., Jouffrais, C., and Dubois, E. Non-visual Interactions to Support Information Sequence Transmission. *Proceedings of the 27th Conference on L'Interaction Homme-Machine*, ACM (2015), 32:1–32:6. **AVEC POSTER**
3. Appert, D. and Truillet, P. Influence of the Presentation of Words for Dyslexic Readers. *Proceedings of the 27th Conference on L'Interaction Homme-Machine*, ACM (2015), 33:1–33:6. **AVEC POSTER**
4. Dubois, E., Bortolaso, C., Appert, D., and Gauffre, G. An MDE-based Framework to Support the Development of Mixed Interactive Systems. *Science of Computer Programming*. 89, PB (2014), 199–221.

## 9.5.1 Eyes-free interaction techniques for distance estimation

Damien Appert, Emmanuel Dubois, Philippe Truillet, Christophe Jouffrais

Corresponding author: C. Jouffrais

### ABSTRACT

Distance perception is a very important function in daily living, which is mainly achieved by vision. In many situations - including visual impairment but also eyes-free perception and interaction - it is important to provide distance cues. In this paper, we designed eight affordable 'passive' and 'active' non-visual interaction techniques that provide the user with absolute distance cues. Passive techniques relied on auditory or tactile stimulation without any movement required from the user. Active techniques were designed so that the user had to move an artefact in order to receive an auditory or tactile cue indicating the distance value. Results show that auditory and tactile cues are equivalent in both passive and active conditions. Altogether, active techniques were more efficient than passive techniques in distance estimation, although two passive techniques were particularly efficient. Another factor explaining the better usability of the designed techniques was the presence or absence of a sensory reference cue in order to estimate the target distance. Overall, these results show that it is possible to design low-cost interaction techniques, including mass market devices, which provide the user with non-visual distance perception. These techniques are usable in many situations, including visual impairment but also eyes-free perception for sighted people, both in real and virtual environments.

Index Terms—Auditory, Haptic, Non-visual perception, HCI, Assistive Technology, Eyes-free, visually impaired

### Introduction

Distance perception is a pervasive task in everyday life. Although the mapping between physical space and visual space (visually perceived space) is still a major research question in psychology and neuroscience, we know that vision is the more efficient sensory modality to perceive distance. In absence of vision, perception of distance is more difficult but not impossible. For instance, many auditory cues (intensity, reverberation, frequency, spectrum, binaural differences) allow perceiving the egocentric distance (defined with the observer as origin) of one or many sound sources (Coleman, 1963). However, perceiving an absolute distance between two objects (allo- or exo-centric distance) without vision is more complicated. Auditory interaction techniques can be used to convey many types of information, including spatial cues. Auditory displays may facilitate user learning (Demarey and Plénacoste, 2001), and then are often designed around metaphors that support analogical transfer between previous knowledge and interaction with the display. The sonification metaphor is used to present data that cannot be linguistically summarized by a computer. Dimensions of the source information (e.g. distance between two objects) are encoded in the dimensions of non-speech sounds by the auditory display. The listener must then decode the source information when the sound is played.

Although research and design on tactile displays is less prominent, it is similarly possible to map information onto tactile stimuli (Gemperle *et al.*, 2001; Tan and Pentland, 1997). Many tactile technologies – relying on skin vibration, pressure, stretching, etc. - are available and are particularly appropriate for wearable and private applications. In addition, they can alleviate perceptual load on other sensory modalities (vision and audition).

Of course, the mapping between the source data on one side, and the sound or tactile dimensions on the other side, is important and may benefit from previous experimental research. For instance, in an auditory display, ordinal rankings are more clearly conveyed by earcon register, an ordinal property, than by timbre, a nominal property (Parente, 2008).

In this paper we focused on non-visual distance perception. We specifically aimed at comparing a set of interaction techniques based on auditory or tactile modalities, which all provide the user with a perception of distance. We designed eight different active (user is actively actuated the artefact in order to receive the stimulus) and passive (user is passively receiving the stimulus) techniques that were tested with two different final feedback cues: Auditory (pure tone) and Tactile (vibration). The experiments that we performed revealed interesting and significant impact of the interaction techniques: although the distinction between vibrotactile and auditory final feedback is not significant, we highlighted significant effect of user agency, as well as presence of a reference cue within the designed interaction techniques.

Before reporting the detailed results of the experiment, we review existing works on tactile and auditory perception in psychology, but also on tactile and auditory displays. Based on this overview, we then present the eight interaction techniques that we designed. Finally we introduce the experimental protocol and results.

## **Distance perception: state of the art**

Distance perception has been widely studied, especially in the field of 3D real and virtual environments. Of course, the impact of many visual cues (e.g. monocular depth cue and motion parallax) has been addressed in numerous experiments (Renner *et al.*, 2013). However, because our research work is oriented toward the design of non-visual interaction techniques, we will not further detail the role of visual cues in distance perception. Here, we first focus our literature survey on the experimental methods that were used in experimental psychology research to assess distance perception. In the second section, we focus on papers addressing the non-visual cues that allow distance perception, and then on interaction design papers that address the rendering of spatial features (and hence distance) via tactile or auditory displays in real or virtual environments.

### **Characteristics of distance perception experiments**

Distance perception – also called distance judgment in some papers (Richardson and Waller, 2007) – has been extensively studied in psychology. In distance perception protocols, two different types of tasks have generally been adopted (Loomis and Knapp, 2003). The first one provides the user with a cue related to an egocentric distance: the goal for the participant is to estimate the distance to a target object. The second one provides the user with an allocentric (exocentric) distance cue: the goal is to estimate the distance separating two points of interest or objects placed in the participant's environment (Easton and Sholl, 1995).

Beyond the type of task itself, reporting the measures of perceived distance is wide ranging (Allan, 1979; Altenhoff *et al.*, 2012; Loomis *et al.*, 1999). The first possibility consists in reporting a response that reflects the judgment of a perceived distance. Most common are procedures based on numerical estimates. Here the subject estimates the distance either in familiar distance units (e.g., meters) or as a multiple of previously perceived distance. The subject then reports the estimate by speech (verbal report) or keyboard input. Matching protocol is another possibility that involves the use of indicative movements: here, the participant performs a movement of the body, e.g. the arm, which is proportional to the estimated distance.

Finally, experimental protocols can rely on two different types of judgments: relative or absolute (Silva, 1985). In delayed matching-to-sample tasks, the answer consists in comparing two distance cues. The estimated distance is then equal, shorter or longer than a reference. In production tasks (Allan, 1979), subjects must estimate the target distance without any sample but according to previous knowledge.

Independently of the method used to express the distance to estimate, action performed by the user is one of the aspects to be considered in distance perception experiments. Indeed, Rieser and Pick showed, in many different tasks, that motor and perceptual systems are mutually calibrated (Rieser *et al.*, 1995). This calibration is dynamic and is probably happening even during normal everyday activities (Mon-Williams and Bingham, 2007). It has also been suggested that distance perception may also depend on the physiological state or the intention to act of the observer. For instance Witt *et al.* observed that objects in personal space appeared closer when the observer was holding a tool and thus able to reach the object (Witt *et al.*, 2007). Balci and Dunning (Balci and Dunning, 2010) also showed that desired targets appeared closer.

### **Distance coding in auditory and tactile displays**

Auditory and haptic sensory modalities can be used to map distance cues. Each sensory modality has specific characteristics in terms of perceptual cues potentially useful to render a distance. In the following paragraphs, we made a non-exhaustive review about psychophysics research on non-visual distance estimation, as well as auditory and tactile displays.

The auditory sensory system is able to perceive distance. Four parameters of static sounds have been identified to render distance cues: intensity, reverberation, spectral contents and binaural differences (Loomis *et al.*, 1999). In an early study Mershon and King (Mershon and King, 1975) have shown that intensity and reverberation are sufficient to estimate perceived distance. More precisely they showed that auditory intensity differences may not serve as an absolute cue to auditory distance, but as a cue to changes in distance. On the contrary, reverberation could serve as an absolute cue, with greater reverberation being associated with greater perceived distances. Change in spectral content may also serve to indicate a relative distance (Little *et al.*, 1992). Finally, a sound is not arriving at the two ears with the same spectrum and intensity. This is a consequence of the filtering effect of the head. Such "head shadow" effects provide directional spatial information but are essentially invariant with source distance (Mills, 1958).

Accordingly, the parameters of the rendered sound can be used to map distance cues in auditory displays (Shinn-Cunningham, 2000). Incorporating overall level effects is straightforward and requires little computational power, but only gives relative distance cues. On the contrary, incorporating realistic reverberation into a display requires extensive computation (Shinn-Cunningham B. *et al.*,

1997). In the field of assistive technology for visually impaired user, many abstract mappings between distance and rendered sounds have also been used. Talbot and Cowan (Talbot and Cowan, 2009), for instance, observed that pitch variation of the sound was not an effective encoding, which is a surprise because many electronic travel aids encode distance as pitch. Temporal variation of the sound, as well as high-frequencies filtering, was found to be a better mapping. Combination of cues has also been used to generate rich signals such as the Doppler Effect and the sound effect anechoic environmental (Zahorik *et al.*, 2005).

The somatosensory system is not, by essence, providing any egocentric or allocentric distance cue. However vibratory or electrotactile devices have been designed that could be used to map distances – obviously with a scale factor – on different parts of the skin (Asamura *et al.*, 1998; Lu *et al.*, 2007; Pasquero and Hayward, 2003; Patel and Abowd, 2004). Actually, most of these devices have been designed for assisting blind users. In their simpler form, they have mainly been used to provide tactile feedback after completion of a task, alert signal or directional cues during navigation (Brunet *et al.*, 2013; Luk *et al.*, 2006). Similar but more complex devices of this type include an embedded camera and a larger array. Here a matrix of tens of vibrators is actuated according to the brightness of the camera image. The vibrotactile array is then used to provide the user with a 2D “vibratory image” of the visual scene (Bach-y-Rita *et al.*, 1998; Kaczmarek *et al.*, 1991). However, although different studies acknowledged the functional principle of the device, to date no practical systems are in use (Bach-y-Rita, 1983). Another type of devices that include force feedback has been used to map distance and size cues (Tzovaras *et al.*, 2004). Evaluations with blind users showed that such devices can be used for educational purposes, mobility and orientation training and navigation in 3-D spaces. The caveat concerning this type of devices is that they are cumbersome and quite expensive.

Based on these different results, and considering the spread of low-cost interactive technologies in mass market, we aimed to design eyes-free interaction techniques to provide distance perception in many different contexts.

## **IT Design for non-visual distance perception**

The different studies on distance perception, auditory, and tactile displays that we identified in the literature helped us designing non-visual interaction techniques, but also the experimental procedure of the study.

### **Design choices**

Our aim was to design low-cost and simple non-visual interaction techniques that allow estimation of either egocentric or allocentric distance. The techniques should then provide the user with an absolute distance value (not in comparison with) that was accurate and reliable. We considered two different possibilities corresponding to the user’s involvement in the perception task. Indeed, in line with the research on the link between perception and action that we previously discussed, it has been shown that the use of explicit command is also an immediacy of feedback, and that natural translation of intentions to actions make some tasks easier (Hutchins *et al.*, 1985). Then, we hypothesized that if the user is actively manipulating an artefact, he would get a better estimate of the target distance than when he is passively receiving a stimulus representing the distance to estimate. We then designed what we called “active” and “passive” techniques.

Obviously, for passive techniques, we focused our design on the audio and tactile media. The most simple design choice was to map the duration of single auditory or tactile stimuli on the distance to estimate. Fraise (Fraise, 1984) established a distinction between short stimuli (psychological present) with upper limit which hardly exceeds 5 sec, and an average value of 2 to 3 sec, and longer stimuli that refer to sets of objects involving short-term memory. In order to respect this observation and avoid having too long stimulation time, we fixed the maximum duration of single stimuli to 1500 ms. The minimum duration was set to 100 ms. We also designed a dual stimulus technique where a reference value (i.e. 750 ms, which is half the maximal duration) was always provided before the target value. In order to optimize sensory perception, it has been demonstrated that a discontinuity between two stimuli would require at least 10 ms (Fraise, 1984). We used gaps of at least 100 ms between different stimuli.

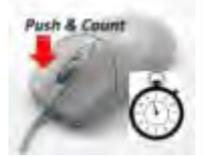
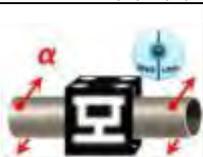
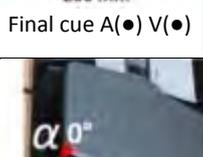
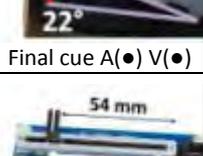
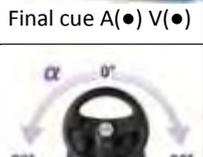
For auditory stimulation, we chose regular headset (Plantronics audio 650 USB). For tactile stimulation, we used a simple and low-cost vibrating motor (VPM2 Vibrating Disk Motor) calibrated at 250 Hz and inserted in home-made bracelet. This frequency corresponds to the optimal frequency for human perception (Makous *et al.*, 1995). These two technologies are available in mass market. Vibrators, for instance, are now systematically inserted in mobile phones, smartwatches, etc.

We also designed active techniques, which required the user to actively manipulate the artefact in order to estimate the target distance. Once again, we chose technologies that are available in mass market (mouse, cursor, pedal, and wheel) and imply different types of movement (rotation, shift, push and click). The designed techniques rely on the actual movement of the user in order to “reach” the target distance. A short auditory or tactile cue (100 ms) was used at the end of the movement to indicate that the target value was reached.

### **Overview of the Interaction Techniques implemented**

We designed 8 interaction techniques that were either passive (Single Stimulus, Dual Stimulus, Palm Gauge) or active (Button, Pipe, Pedal, Slider, and Wheel). Table 1 describes each of the 8 techniques with name, illustrative picture, technical and functional details. For instance, with the audio single stimulus technique (passive), users passively received a pure sound whose duration corresponds to the target value: the longer the tone, the larger the distance to estimate. As a second example, with the palm gauge (passive), the hand of the user was resting on the gauge. The angle of the gauge was proportional to the target distance. With the slider technique (active), the user was moving the slider until he received a short auditory or tactile cue which indicated that the target distance has been reached. The length between the starting point and the actual position was proportional to the target distance to estimate. Before the next trial, the user had to move back the slider towards the starting point.

**Table 1: Description of the 8 non-visual interaction techniques that were designed. The three first ones are passive techniques: user is passively receiving a stimulus. The five last ones are active techniques: user is actively moving the artefact until the final cue indicates that the target value has been reached. The final cue was either tactile or auditory.**

Passive techniques	Single Stimulus		<ul style="list-style-type: none"> <li>- Material: Headset or wrist vibrator.</li> <li>- Stimulation: auditory or vibrotactile</li> <li>- Distance (1-100) = duration of a single stimulus (15-1500 ms)</li> <li>- Reset: automatic</li> </ul>
	Dual Stimulus		<ul style="list-style-type: none"> <li>- Material: Headset or wrist vibrator</li> <li>- First stimulus is a reference (distance 50=750 ms)</li> <li>- Stimulation: auditory or vibrotactile</li> <li>- Distance (1-100) = duration of a single stimulus (15-1500 ms)</li> <li>- Reset: automatic</li> </ul>
	Palm Gauge		<ul style="list-style-type: none"> <li>- Material: Rotating needle driven by an Arduino card [1061_1 PhidgetAdvancedServo 8-motor] and servomotor [HS-322HS Servo]</li> <li>- Stimulation: tactile</li> <li>- Distance (1-100) = latest position of the needle within the palm (1=1.8° ; 100=180°)</li> <li>- Reset: automatic</li> </ul>
Active techniques	Button	 Final cue A(●) V(●)	<ul style="list-style-type: none"> <li>- Material: Mouse with one button</li> <li>- Distance (1-100) = duration of button press until final cue (1=15ms and 100=1500ms)</li> <li>- Final cue: auditory or vibrotactile</li> <li>- Reset: when button is released</li> </ul>
	Pipe	 Final cue A(●) V(●)	<ul style="list-style-type: none"> <li>- Material: 4 sides with 4 markers mounted on a steel pipe (200 mm, 20mm Ø); video recognition of the markers orientation with a webcam [Logitech Quickcam pro 9000] and library Nyartoolkits-4.1.1]</li> <li>- Distance (1-100) = rotation angle of the pipe until final cue (1=1/400<sup>th</sup> turn and 100=¾ turn)</li> <li>- Final cue: auditory or vibrotactile</li> <li>- Reset: manual (line up the pipe with zero position)</li> </ul>
	Pedal	 Final cue A(●) V(●)	<ul style="list-style-type: none"> <li>- Material: game pedal [Logitech precision vibration feedback wheel]</li> <li>- Distance (1-100) = angle of the pedal pushed with the hand until final cue (1=0.2° and 100=22°)</li> <li>- Final cue: auditory or vibrotactile</li> <li>- Reset: when pedal is released</li> </ul>
	Slider	 Final cue A(●) V(●)	<ul style="list-style-type: none"> <li>- Material: Phidget Slider [1112_1 slider 60]; Arduino card [1018_2 PhidgetInterfacekit 8/8/8]</li> <li>- Distance (1-100) = slider translation (1=0.54 and 100=54 mm)</li> <li>- Final cue: auditory or vibrotactile</li> <li>- Reset: active (bring back the slider to position 0)</li> </ul>
	Wheel	 Final cue A(●) V(●)	<ul style="list-style-type: none"> <li>- Material: Game Wheel [Logitech precision vibration feedback wheel]</li> <li>- Distance (1-100) = wheel rotation angle until final cue (1=0.90° and 100=90°)</li> <li>- Reset: automatic</li> <li>- Final cue: auditory or vibrotactile</li> <li>- Remark: wheel can be turn to the left or to the right</li> </ul>

## Experimental Procedure

A total of 12 blindfolded subjects participated in the experiment. None of the subjects previously participated in any related study. The group consisted of nine males and three females, with a mean age of  $28.7 \pm 7.2$ . The experimental setup was installed in a silent room and included eight prototypes (see Table 1) connected to a PC via a USB2 hub (Mobility Lab Hub USB 2.0 10 ports). The experiment was controlled by home-made software written in C# language.

Subjects completed two experimental sessions, one with auditory cues and the other with tactile cues. This dichotomy was valid for passive techniques where the stimulus is either auditory or vibrotactile (except for the palm gauge that is always tactile). It is also valid for active techniques where the cue at the end is either tactile (short vibration) or auditory (short pure sound). The two sessions were at least one day from each other and the order was counterbalanced across subjects. During a session, the subjects were sitting in front of the table with the different devices. They were told that the aim of the task was to estimate an absolute value – between 1 and 100 – based on eight different interaction techniques. Each session included eight blocks corresponding to the eight interaction techniques. At the beginning of each block, the aim of the task was repeated. Then, subjects completed a familiarization phase including ten trials with pseudo-random values between 1 and 100. They received a verbal feedback about their own response and target value after each trial. They were allowed to repeat familiarization trials on demand. The test *per se* included 30 trials with pseudo-random values between 1 and 100. After each trial, subjects indicated the estimated value but did not receive any feedback concerning the target value or the correctness of their response. After each block, subjects answered a satisfaction questionnaire concerning the technique that they just used. The questionnaire included five questions (is the technique Enjoyable / Accurate / Frustrating / Useful / Easy to understand?) with a 5-degree Likert scale (1: strongly disagree and 5: strongly agree) and a SUS questionnaire. At the end of the eight blocks, the subjects used a set of cards to classify the eight techniques according to their preference. Within a session, the different subjects fulfilled the eight blocks with a different order. The block order was modified with the following rules: 1/ a technique never had the same position in a series; 2/ a passive technique (palm gauge, continuous stimulus and dual stimulus) never followed another passive technique. In addition, a given subject did not receive the same block order in the second experimental session.

The main dependent variable was the error between the estimated and target values (accuracy). We also used the standard deviation of the error as a measure of precision. The independent variables were the eight techniques, and the sensory modality – auditory vs. vibrotactile – when relevant (i.e. for all the techniques but the palm gauge).

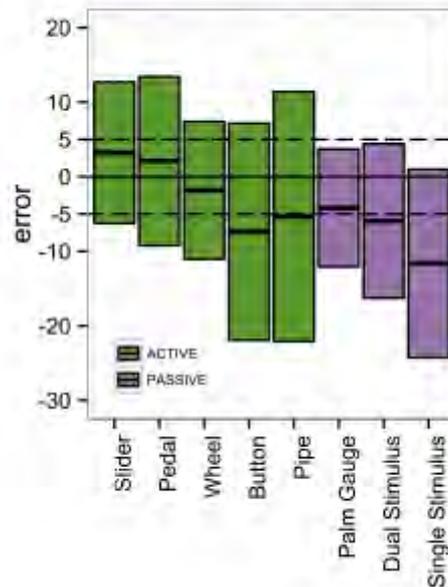
## Results

When not specifically mentioned, the statistical test was a Friedman test performed with the R software. For all the statistical tests, the significance threshold was set to .05. Bonferroni adjustment for multiple comparisons was performed when needed.

### Accuracy and precision of the estimate

We first checked that the session number has no effect on the accuracy of the response ( $X^2=1.33$ ,  $df=1$ ,  $p\text{-value}=n.s.$ ), which rules out a potential global effect of the experimental design. In our experimental design, we included “active” and “passive” techniques. A Friedman test showed a

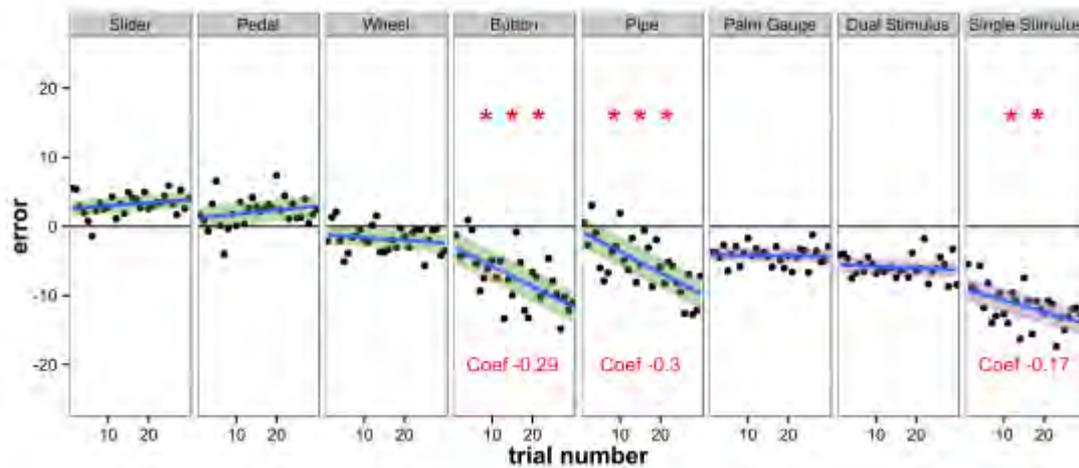
significant difference of accuracy (error) between both categories of techniques ( $X^2= 12$ ,  $df=1$ ,  $p < .001$ ). The active techniques induce a very small average error of  $-1.9 \pm 13.3$ . The accuracy for the passive techniques was lower with  $-7.2 \pm 10.9$ . We then performed a Friedman test with techniques (8 levels) and sensory modality (2 levels) as factors. As expected, we observed a highly significant main effect of the technique on the accuracy ( $X^2=58.9$ ,  $df=7$ ,  $p < .001$ ). (Fig. 1) shows the mean error and standard deviation observed for each technique, which indicate accuracy and precision of the techniques respectively. As illustrated on the plots, the Palm Gauge, Pedal, Slider and Wheel were significantly more accurate and more precise than the other techniques. We also observed that there was no main effect of the sensory modality on the error ( $X^2=0$ ,  $df=1$ ,  $p=n.s.$ ).



**Fig. 1. Error (accuracy) and standard deviation (precision) according to the interaction technique. The palm gauge, the Pedal, the Slider and the wheel are significantly more accurate (smaller error) and more precise (smaller SD) than the other techniques. Dotted lines show +5 and -5 limits.**

### Evolution of the accuracy for each technique

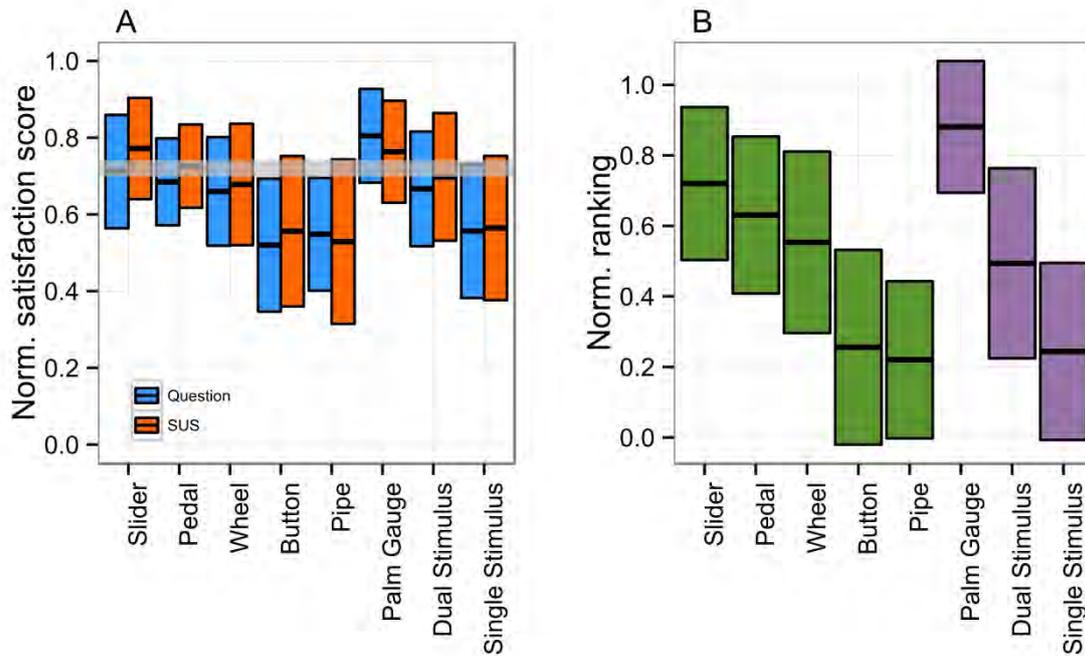
We made a linear regression on the mean error for each technique (Fig. 2). We observed two groups of techniques according to the evolution of the accuracy across trials. In one group - including Palm gauge, Slider, Pedal, Wheel, and Dual stimulus - the accuracy remained stable. In the other group, including the Single stimulus ( $F=-0.17X - 9.0$ ,  $r^2=27.6\%$ , adjusted  $r^2=25.0\%$ ,  $p < .01$ ), Button ( $F=-0.29X - 2.85$ ,  $r^2= 40.9\%$ , adjusted  $r^2=38.8\%$ ,  $p < .001$ ) and Pipe ( $F=-0.30X - 0.63$ ,  $r^2= 44.5\%$ , adjusted  $r^2=42.5\%$ ,  $p < .001$ ), the accuracy significantly and continuously decreased between the first and the last trial.



**Fig. 2.** Linear regression of the mean accuracy (12 subjects, auditory and tactile modality) across trial for each technique. The last three techniques on the graph induce significant and increasing underestimation of the target value (\*\*  $p < .01$ ; \*\*\*,  $p < .001$ ). Slopes of the regression lines are indicated in red for significant regressions only.

### User' Satisfaction

After each block of trials, subjects answered a SUS questionnaire and a series of five independent questions (is the technique Enjoyable / Accurate / Frustrating / Useful / Easy to understand?) The mean normalized SUS and questions scores for the 12 subjects are shown in (Fig. 3A). These two scores are perfectly coherent and show that three techniques were considered as good or acceptable (Bangor *et al.*, 2008): Palm Gauge (76.4, good), Slider (77.2, good) and Pedal (72.6). The remaining techniques (Dual Stimulus (69.8), Wheel (67.8), Single stimulus (56.5), Button (55.6) and Pipe (52.9)) were considered as not acceptable. Finally, the users ranked the eight techniques in preference order. The mean rank was then computed for each technique and normalized. (Fig. 3B) shows that the top ranked techniques were Palm Gauge, Slider and Pedal, whereas the last ranked techniques were Single Stimulus, Pipe and Button.



**Fig. 3. (A) Satisfaction scores for each technique. The blue and orange boxes represent the mean and SD of questions scores and SUS from 12 subjects. The Y axis represents normalized scores. The grey box indicates 3 SUS categories: good (above the box), acceptable (within the box), and not acceptable (below the box). Three techniques only are considered as good or acceptable. (B) Technique ranking averaged over the 12 subjects.**

### Active techniques or presence of a reference

The previous results showed that the usability (including performance and satisfaction) of some techniques, namely Palm Gauge, Slider and Pedal, was excellent as compared to other ones. Our hypothesis for enhanced usability was user involvement in the task (active vs passive techniques). Fig. 4A shows that, indeed, the type of technique plays a significant role in accuracy. Active techniques are significantly more accurate than passive ones. However the type of technique is not statistically discriminant for accuracy stability (Fig. 5, first and second plots) neither for user satisfaction (Fig. 6, first and second plots). Indeed, individual results concerning the techniques showed that the preferred and most usable technique was passive (palm gauge), followed by three active techniques (Slider, Pedal, and Wheel). Specifically, the palm gauge provided the user with excellent and stable distance estimates. In addition, it was ranked as the most satisfactory of the eight techniques. Then, the three active techniques (Slider, Pedal, and Wheel) and another passive technique (dual stimulus) provided good and reliable distance estimates, and almost acceptable satisfaction. Two active techniques (Button and Pipe) were at the same time not efficient, not stable and not satisfying.

Qualitative analysis of these different active and passive techniques showed that another hypothesis may explain the observed differences. Indeed, when using these techniques, the user was able to use a persistent reference cue in order to estimate target distances. The Palm Gauge technique was considered as passive because the user receives the instruction (target distance) without performing any action. The gauge just turns over the resting palm. During informal discussion with the users, it appeared that a comparison method was frequently used to estimate target values. In fact, the

gauge stimulates the palm close to the digits. Many subjects were then using the digits as landmarks for estimating the target value (e.g. around 25 units for each digit). The Slider, Pedal, and Wheel were based on a movement executed by the user between two limits (abutments) that were imposed by the device. These mechanical limits provided landmarks for target estimation when the movement stopped on it. During the experiments we observed users that were discreetly looking for the abutment of the device. For instance, using the slider, subjects were looking for the end of the device with the little finger after receiving the final distance cue, and before estimating the target value. This last action was not necessary to succeed in the task and emerged as a recurrent behavior in some subjects. Similarly, using the Wheel or Pedal, they were sometimes completing the movement with the device after receiving the final cue. Therefore, one common observation for this group of techniques is the presence of landmarks that help to estimate the target value.

We made a subdivision of the techniques according to the presence/absence of a perceptible reference cue. Palm Gauge, Slider, Pedal, and Wheel were included in the group “with reference”. The Dual Stimulus technique was also included in this group because the first stimulus in this technique was a reference cue by definition (half the maximal distance to estimate). The three other techniques (Single Stimulus, Button, and Pipe) were included in the group ‘without reference’. The result of a Friedman comparison test on the accuracy and precision of the two groups is illustrated in (Fig. 4B). The results confirm the importance of the user agency (Fig. 4A), but also show a very significant effect of the presence of the reference cue ( $X^2=12$ ,  $df=1$ ,  $p < .001$ ; Fig. 4B). Techniques with a reference cue provided an accuracy of  $-1.3 \pm 10.3$ , whereas techniques without reference provided an accuracy of  $-8.1 \pm 15.0$ . Interestingly, temporal stability and user satisfaction were also significantly different between these two groups while no clear tendency was established when opposing active and passive techniques. (Fig. 5A) clearly shows that the drift of the estimate is similar for active and passive techniques, with a negative slope of .11 and .07. This negative slope indicates that the target value is systematically underestimated, and that underestimation is increasing along trials. On the contrary, techniques with a reference cue provide a stable estimate along trials (Fig. 5B, left plot). For the techniques without reference cue (Fig. 5B, right plot), we observed a very pronounced drift with a slope of .26, which indicates that the subdivision according to the reference cue is more pertinent than the active/passive subdivision. Finally, this subdivision also provided a better explanation to user satisfaction scores (Fig. 6). Indeed, independently of the observed variable (SUS score, open questions, rank), Mann-Whitney-Wilcoxon tests did not reveal any difference according to the passive/active categorization. In contrast, the presence of a reference cue led to a significantly greater satisfaction (SUS,  $W=27$ ,  $p < .01$ ; Questions,  $W=19$ ,  $p < .01$ ; and Ranking,  $W=0$ ,  $p < .001$ ).

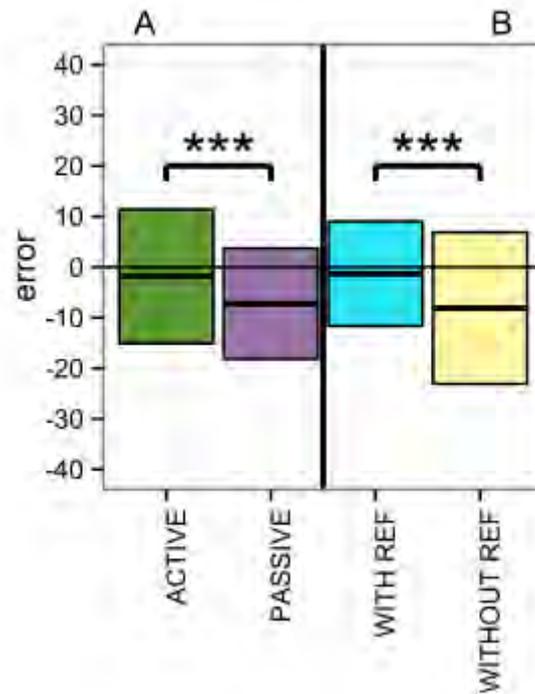
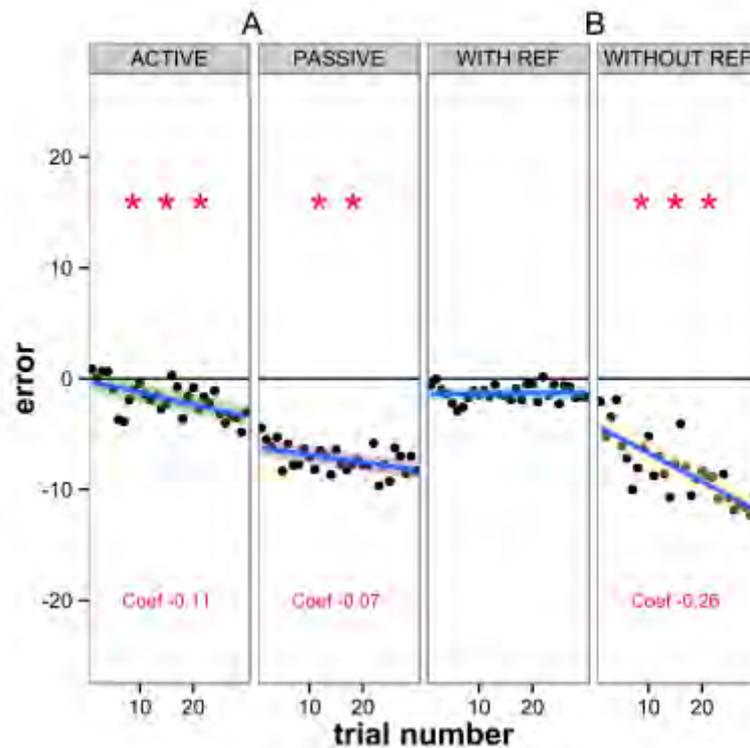
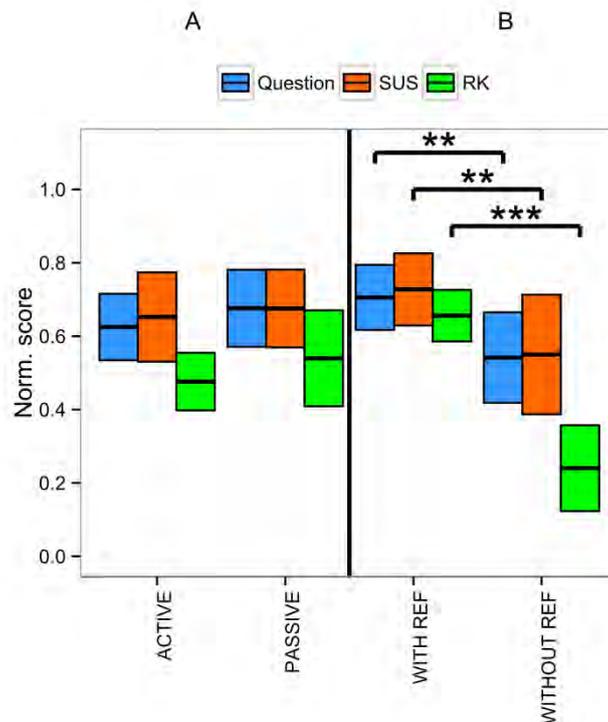


Fig. 4. Error (accuracy) and standard deviation (precision) of the techniques according to: A) user agency (passive vs. active), or B) presence or absence of a reference cue. Active techniques are significantly more efficient than passive techniques. Techniques with a reference cue are also more efficient than techniques without reference cue (\*\*\*,  $p < 0.001$ ). Note that the two plots in A) and B) present the same data subdivided with a different criteria (passive/active or presence/absence of a cue, respectively).



**Fig. 5. Linear regression of the mean accuracy (12 subjects, auditory and tactile modalities grouped) across trials according to: A) user agency (passive vs. active), or B) presence or absence of a reference cue. Only techniques with a reference (third plot) were perfectly stable during the experiment. Note that the two plots in A) and B) present the same data subdivided with a different criteria (passive/active or presence/absence of a cue, respectively).**



**Fig. 6. Satisfaction scores (Questions, SUS, and Ranking) according to: A) passive/active feature; or B) presence/absence of a reference cue. Y axis represents normalized scores with SD. Significant differences are indicated with brackets (\*\*, p<0.01,\*\*\*, p<0.001).**

## Discussion

In this experiment, we designed eight different non-visual techniques that were used to estimate an absolute value between 1 and 100 (it used to be a distance in our protocol). The first main result showed that there is no difference according to the sensory modality (auditory vs. tactile) that was used to provide the instruction (passive techniques) or to render the final cue (active techniques). Thus, according to the CARE properties (Coutaz and Nigay, 1994), these two sensory modalities are functionally equivalent in this task. This result is important in that it allows designers to choose the more adequate sensory modality according to the context. For instance, it is known that blind users are using natural sounds from the environment when traveling. Tactile techniques would allow providing them with non-overlapping, discrete and reliable cues about distances of objects, paths, etc. This is also true for designing interfaces devoted to sighted users that need to release the visual and auditory sensory channels (e.g. pilots, firemen, etc.)

Putting apart the sensory modality, the Table 2 recaps the different results obtained in the experiment. It clearly shows that three techniques only are definitely usable, with good accuracy and high satisfaction rate. When considering these results altogether, the three most usable techniques were the Palm Gauge, the Slider and the Pedal.

**Table 2: Recap of the results for the eight techniques that were designed and evaluated**

TI	Accuracy (Mean)	Precision (SD)	Stability	SUS	Rank	Agency	Refer.
Palm gauge	<5	<12	OK	Good	1	Passive	W
Slider	<5	<12	OK	Good	2	Active	W
Pedal	<5	<12	OK	Acceptable	2	Active	W
Wheel	<5	<12	OK	Not Accept.	3	Active	W
Dual Stimulus	>5	<12	OK	Not Accept.	3	Passive	W
Single Stimulus	>5	>12	Not OK	Not Accept.	4	Passive	WO
Button	>5	less	Not OK	Not Accept.	4	Active	WO
Pipe	>5	less	Not OK	Not Accept.	4	Active	WO

### Agency of the user or presence of a reference cue?

Our first hypothesis was that user agency (active vs. passive) would explain most of the usability of the technique, which is partly true. As mentioned by (Hutchins *et al.*, 1985), it is likely that direct manipulation of the artefacts help the user in target value estimation. The immediacy of feedback and natural translation of intentions into actions make the task easier. However, it appears that the most usable technique (Palm Gauge) was passive and that the two following ones were active (Slider and Pedal). Even more intriguing, we observed that some active techniques (Button and Pipe) were absolutely not efficient and raised usability issues. Then although we noticed a global significant effect of the user agency on the accuracy of the responses, it seems that this factor is not sufficient to explain the technique usability.

The presence or absence of a perceptible reference while manipulating the technique was a better factor to explain the dichotomy between usable and not usable techniques. For the Palm Gauge, the immediacy of the feedback is preserved as the mapping between the stimulus (rotation of the needle within the palm) and the estimated value is quite easy. In addition, the tactile receptor density in the hand is very high (Johansson and Vallbo, 1979) which probably provides the user with a high confidence. This very easy mapping probably explains the excellent accuracy of the method (SD is 7.9 only), and why the Palm Gauge has been ranked in the first position, with a very high SUS. Furthermore, although we did not specifically check this point in the study, we guess that the stability of the estimate was provided via a comparison method between the position of the needle in the palm and the digits. This method is potentially very powerful as it provides five points of comparison corresponding to the different digits and not only the maximum reachable point as with the Slider.

More generally, the results very clearly showed that the techniques with a reference were providing the user with a very stable estimate of the distance across time (Fig. 5). Interestingly, the reference

was of different type according to the techniques. In the only passive technique with a reference (dual stimulus), the reference was a half-distance stimulus (750 ms duration) that was systematically played prior to the target stimulus. Because it was identical in all the trials, this stimulus was non-significant for the task, but it provided the user with the possibility to match the target with the initial sample whose duration was known. This observation is probably very close to the results of (Fraisse, 1984) who observed that durations of stimuli must be close to learned duration values if they are to be perceived with a certain amount of precision and to be clearly distinguished.

The other techniques with a reference cue were active (pedal, slider and wheel). These techniques were designed according to three types of movement (push, shift, and rotate) that had to be performed by the user in order to receive the final cue. By essence, these techniques rely on artefacts that have a course limited by two mechanical abutments. We observed that many subjects used these abutments in order to compare the target value with the minimal or maximal values. Indeed they, for instance, moved the slider or the wheel to the maximal abutment after receiving the target cue, which is not essential to complete the task. Hence, here again what we called the reference was very useful to estimate the target, and with a great stability over time (Fig. 5). As it was not addressed in the current study, we now need to know if this comparison is conscious or not.

### **Accuracy, precision, stability and satisfaction are coherent**

Interestingly, the more accurate, precise and stable techniques were given the best satisfaction scores and ranking (Table 2). This observation is not trivial in that the subjects did not receive any feedback during the experiments, and were then not able to know if their successive estimates were correct or not. It has previously been shown that, independently of the technique used, the role of feedback is very important (Altenhoff *et al.*, 2012). We already advocated why the more accurate techniques were also the most reliable and satisfactory. By opposition, the Pipe, Button and Single stimulus were the worse usable techniques. Button and Single stimulus techniques were similar in that the only way to estimate the target value was to integrate elapsed time. It happens that without any reference cue, time integration is very difficult. Our results showed that it does not make a big difference when the subject has to estimate the duration of a single auditory or vibratory stimulus, or the duration of a button press.

The Pipe was the worse technique according to precision (SD = 16.8), stability (slope across trials = 0.30), satisfaction (SUS=52.9), and ranking (eight position). One explanation for this observation relies on the design of the technique itself. Estimating a value with the Pipe is based on the integration of the own movement only, without any material landmark. The Pipe can rotate indefinitely. Although this design may provide enhanced control possibilities, there are no starting or ending positions which may help the user to take a reference. Furthermore, using the Pipe implies a bimanual manipulation: one hand controls the rotation while the other one is maintaining the Pipe. The user's action onto the Pipe is thus a combination of a rotation angle applied to the Pipe, and several repositioning of the user's hand manipulating the Pipe (the "clutching" mechanism). This might increase the difficulty of the judgment.

### **Usage scenario with eyes-free interaction techniques for distance perception**

Obviously, there are many situations that visually impaired people are facing in which such non visual interaction techniques would be very relevant. It might be useful in the context of mobility, to provide the user with distance cues to landmarks or obstacles (Fujimoto and Turk, 2014). It might

also be useful in the context of map exploration at home. It may, for instance, be included in tactile interactive maps (Brock, 2013; Brock *et al.*, 2015) and serve as an advanced function to provide the user with inconspicuous and private distance cues between different landmarks.

But sighted person might also take advantage of such techniques in contexts where vision is not usable. It has been shown, for instance, that firemen and blind people may require similar non-visual guidance techniques (Renaudin and Gilliéron, 2011). Tasks in which vision is already very highly engaged (driving, piloting, etc.) represent another very appropriate context. Indeed, eyes-free techniques may provide the user with very useful distance cues concerning the environment without having to change gaze orientation. More subtle interactive tasks may also benefit from these eyes-free techniques. Map exploration, for instance, is a relevant example for sighted people. With the increasing amount of available data, maps are increasingly used as a structured support to visualize contextualized information: pictures taken by visitors, restaurant, itineraries, on-going constructions, damage or accident reports, etc. Therefore, visual displaying of all this data simultaneously turns out to be impossible or results in overloaded rendering (Keim *et al.*, 2008). Enriching the map interaction with non-visual modalities would provide the user with different kind of information in a synergistic way (Trocka-Leszczynska and Jablonska, 2013).

### **Towards an integrated version of the Palm Gauge IT, Slider and Pedal**

We strongly believe that such eyes-free interaction techniques could be easily integrated into many different personal devices. Obviously, it does not represent a big engineering challenge to design desktop devices that integrate such techniques. A keyboard or a mouse may easily integrate an artefact similar to the Palm Gauge technique for instance. However, we especially believe that smartphones represent a promising platform for offering such eyes-free interaction techniques. Concretely, the recent Lumia 830 now provides a rounded screen for a rapid access to specific menus. Clearly the rounded part, covering the right edge of the device, could serve as a support for implementing the Slider technique. Finally, shape-changing handheld devices are currently addressed at the academic level (Hemmert, 2010; Kim *et al.*, 2008; Pedersen *et al.*, 2014) but are also slightly emerging as industrial concepts and commercial products (Kildal *et al.*, 2012). This is of course still exploratory but the Palm Gauge technique that we designed could easily be mapped on this type of devices.

### **Conclusion**

In this paper we have explored the design and usability of eyes-free interaction techniques that provide the user with an estimation of an absolute value. In our scenario, the target value was a distance but could have been anything else. We have designed and optimized eight different techniques based on previous knowledge on human distance judgment from a dual point of view: psychological research on distance perception, and research on auditory and tactile displays. Out of these eight solutions, three clearly emerged in terms of performance (accuracy and stability of the estimates) as well as user's satisfaction (SUS, ranking): the Palm Gauge (a needle rotating in the palm), the Slider and the Pedal (regular slider and pedal but with a sensory cue indicating the target value when it is reached). We introduced user's agency in our design factors, which resulted in passive versus active interaction techniques. Although the results indeed revealed a significant effect of agency on distance estimation, it appeared that the presence of a sensory reference in order to estimate the target value had a more significant impact. Future work will focus on the *in-situ*

deployment of such techniques. Indeed, they could be used in much more demanding tasks than evaluating just one distance value. For instance, it may be worth exploring the usability of these techniques in the context of an itinerary memorization, i.e. when they carry out a set of successive distances. Furthermore, it may be of interest to check if they alleviate perceptual load when used during very demanding visual task.

## References

1. Allan, L.G., 1979. The perception of time. *Perception & Psychophysics* 26, 340–354.
2. Altenhoff, B.M., Napieralski, P.E., Long, L.O., Bertrand, J.W., Pagano, C.C., Babu, S. V., Davis, T.A., 2012. Effects of calibration to visual and haptic feedback on near-field depth perception in an immersive virtual environment, in: *Proceedings of the ACM Symposium on Applied Perception - SAP '12*. ACM Press, New York, USA, pp. 71–78.
3. Asamura, N., Yokoyama, N., Shinoda, H., 1998. Selectively stimulating skin receptors for tactile display. *IEEE Computer Graphics and Applications* 18, 32–37.
4. Bach-y-Rita, P., 1983. Tactile vision substitution: past and future. *The International journal of neuroscience* 19, 29–36.
5. Bach-y-Rita, P., Kaczmarek, K.A., Tyler, M.E., Garcia-Lara, J., 1998. Form perception with a 49-point electrotactile stimulus array on the tongue: a technical note. *Journal of rehabilitation research and development* 35, 427–430.
6. Balcetis, E., Dunning, D., 2010. Wishful seeing: more desired objects are seen as closer. *Psychological science* 21, 147–152.
7. Bangor, A., Kortum, P.T., Miller, J.T., 2008. An Empirical Evaluation of the System Usability Scale, in: *International Journal of Human-Computer Interaction*. pp. 574–594.
8. Brock, A.M., 2013. Touch the map!: designing interactive maps for visually impaired people. *SIGACCESS Access. Comput.* 105 9–14.
9. Brock, A.M., Truillet, P., Oriola, B., Picard, D., Jouffrais, C., 2015. Interactivity Improves Usability of Geographic Maps for Visually Impaired People. *Human-Computer Interaction* 30, 156–194.
10. Brunet, L., Megard, C., Paneels, S., Changeon, G., Lozada, J., Daniel, M.P., Darses, F., 2013. "Invitation to the voyage": The design of tactile metaphors to fulfill occasional travelers' needs in transportation networks, in: *2013 World Haptics Conference, WHC 2013*. IEEE, pp. 259–264.
11. Coleman, P.D., 1963. An analysis of cues to auditory depth perception in free space. *Psychological Bulletin* 60, 302–315.
12. Coutaz, J., Nigay, L., 1994. Les propriétés "CARE" dans les interfaces multimodales, in: *IHM'94 Sixièmes Journées Sur l'Ingénierie Des Interfaces Homme-Machine*. Lille, pp. 7–14.
13. Demarey, C., Plénacoste, P., 2001. User, sound context and use context: What are their roles in 3D sound metaphors design?, in: Hiipakka, J., Zacharov, N., Takala, T. (Eds.), *International Community for Auditory Display*. Georgia Institute of Technology International Community on Auditory Display, Espoo, Finland, pp. 136–140.
14. Easton, R.D., Sholl, M.J., 1995. Object-array structure, frames of reference, and retrieval of spatial knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 21, 483–500.
15. Fraisse, P., 1984. Perception and estimation of time. *Annual review of psychology* 35, 1–36.
16. Fujimoto, E., Turk, M., 2014. Non-Visual Navigation Using Combined Audio Music and Haptic Cues, in: *Proceedings of the 16th International Conference on Multimodal Interaction - ICMI '14*. ACM Press, New York, USA, pp. 411–418.
17. Gemperle, F., Ota, N., Siewiorek, D., 2001. Design of a wearable tactile display, in: *Proceedings Fifth International Symposium on Wearable Computers*. IEEE, Zurich, pp. 5–12.
18. Hemmert, F., 2010. Interfaces beyond the surface: a structural approach to embodiment, in: *CHI '10 Extended Abstracts on Human Factors in Computing Systems (CHI EA '10)*. ACM Press, New York, USA, pp. 2903–2906.

19. Hutchins, E., Hollan, J., Norman, D., 1985. Direct Manipulation Interfaces. *Human-Computer Interaction* 1, 311–338.
20. Johansson, R.S., Vallbo, a B., 1979. Tactile sensibility in the human hand: relative and absolute densities of four types of mechanoreceptive units in glabrous skin. *The Journal of physiology* 286, 283–300.
21. Kaczmarek, K. a., Webster, J.G., Bach-y-Rita, P., Tompkins, W.J., 1991. Electrotactile and vibrotactile displays for sensory substitution systems. *IEEE Transactions on Biomedical Engineering* 38, 1–16.
22. Keim, D., Andrienko, G., Fekete, J.-D., Görg, C., Kohlhammer, J., Melançon, G., 2008. Visual Analytics: Definition, Process, and Challenges, in: Kerren, A., Stasko, J., Fekete, J.-D., North, C. (Eds.), *Information Visualization: Human-Centered Issues and Perspectives*, Lecture Notes in Computer Science. Springer, pp. 154–175.
23. Kildal, J., Paasovaara, S., Aaltonen, V., 2012. Kinetic device: designing interactions with a deformable mobile interface., in: *CHI '12 Extended Abstracts on Human Factors in Computing Systems (CHI EA '12)*. ACM Press, New York, USA, pp. 1871–1876.
24. Kim, S., Kim, H., Lee, B., Nam, T.-J., Lee, W., 2008. Inflatable mouse: volume-adjustable mouse with air-pressure-sensitive input and haptic feedback., in: *Proceeding of the Twenty-Sixth Annual CHI Conference on Human Factors in Computing Systems - CHI '08*. ACM Press, New York, USA, pp. 211–224.
25. Little, a. D., Mershon, D.H., Cox, P.H., 1992. Spectral content as a cue to perceived auditory distance. *Perception* 21, 405–416.
26. Loomis, J., Knapp, J., 2003. *Virtual and Adaptive Environments*, Mahwah, NJ. ed, Virtual and adaptive environments. CRC Press.
27. Loomis, J.M., Klatzky, R.L., Gollidge, R.G., 1999. Auditory Distance Perception in Real, Virtual, and Mixed Environments, in: Ohta, Y., Tamura, H. (Eds.), *Mixed Reality : Merging Real And Virtual Worlds*. Tokyo, pp. 201–214.
28. Lu, Y.C., Cooke, M., Christensen, H., 2007. Active binaural distance estimation for dynamic sources, in: *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*. pp. 933–936.
29. Luk, J., Pasquero, J., Little, S., MacLean, K., Levesque, V., Hayward, V., 2006. A role for haptics in mobile interaction: initial design using a handheld tactile display prototype, in: Rebecca Grinter, Rodden, T., Aoki, P., Cutrell, E., Jeffries, R., Olson, G. (Eds.), *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, (CHI '06)*. ACM Press, New York, USA, pp. 171–180.
30. Makous, J.C., Friedman, R.M., Vierck, C.J., 1995. A critical band filter in touch. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 15, 2808–2818.
31. Mershon, D.H., King, L.E., 1975. Intensity and reverberation as factors in the auditory perception of egocentric distance. *Perception & Psychophysics* 18, 409–415.
32. Mills, a. W., 1958. On the Minimum Audible Angle. *The Journal of the Acoustical Society of America* 30, 237.
33. Mon-Williams, M., Bingham, G.P., 2007. Calibrating reach distance to visual targets. *Journal of experimental psychology. Human perception and performance* 33, 645–656.
34. Parente, P., 2008. *Clique: perceptually based, task oriented auditory display for gui applications*. University of North Carolina.
35. Pasquero, J., Hayward, V., 2003. STReSS : A Practical Tactile Display System with One Millimeter Spatial Resolution and 700 Hz Refresh Rate, in: *Proc Eurohaptics*. Dublin, Ireland, pp. 94–110.
36. Patel, S.N., Abowd, G.D., 2004. The ContextCam: Automated Point of Capture Video Annotation, *Proceedings of The International Conference on Ubiquitous Computing (UbiComp '04)*, Lecture Notes in Computer Science. Springer Berlin Heidelberg, Berlin, Heidelberg.
37. Pedersen, E.W., Subramanian, S., Hornbæk, K., 2014. Is my phone alive?: a large-scale study of shape change in handheld devices using videos., in: *Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems - CHI '14*. ACM, New York, pp. 2579–2588.
38. Renaudin, V., Gilliéron, P.-Y., 2011. Personal Robust Navigation in Challenging Applications. *Journal of Navigation* 64, 235–249.

39. Renner, R.S., Velichkovsky, B.M., Helmert, J.R., 2013. The perception of egocentric distances in virtual environments - A review. *ACM Computing Surveys* 46, Article 23 (December 2013), 40 pages.
40. Richardson, A.R., Waller, D., 2007. Interaction with an immersive virtual environment corrects users' distance estimates. *Human factors* 49, 507–517.
41. Rieser, J.J., Pick, H.L., Ashmead, D.H., Garing, A.E., 1995. Calibration of human locomotion and models of perceptual-motor organization. *Journal of Experimental Psychology: Human Perception and Performance* 21, 480–497.
42. Shinn-Cunningham B., Lehnert, H., Kramer, G., Wenzel, E.M., Durlach, N.I., 1997. Auditory Displays, in: Gilkey, R., Anderson, T. (Eds.), *Binaural and Spatial Hearing in Real and Virtual Environments*. New York, pp. 611–663.
43. Shinn-Cunningham, B.G., 2000. Distance Cues For Virtual Auditory Space, in: *Proceedings of the IEEE-PCM*. pp. 227–230.
44. Silva, J.A. Da, 1985. Scales for Perceived Egocentric Distance in a Large Open Field: Comparison of Three Psychophysical Methods. *The American Journal of Psychology* 98, 119.
45. Talbot, M., Cowan, W., 2009. On the audio representation of distance for blind users, in: *Proceedings of the 27th International Conference on Human Factors in Computing Systems - CHI 09*. ACM Press, New York, USA, pp. 1839–1848.
46. Tan, H.Z., Pentland, A., 1997. Tactual displays for wearable computing. *Personal Technologies* 1, 225–230.
47. Trocka-Leszczynska, E., Jablonska, J., 2013. Design of Modern Hotels – Humanization of the Residential Environment, in: Stephanidis, C., Antona, M. (Eds.), *Universal Access in Human-Computer Interaction. Design Methods, Tools, and Interaction Techniques for inclusion*. Springer Berlin Heidelberg, pp. 358–367.
48. Tzovaras, D., Nikolakis, G., Fergadis, G., Malasiotis, S., Stavrakis, M., 2004. Design and implementation of haptic virtual environments for the training of the visually impaired. *IEEE transactions on neural systems and rehabilitation engineering : a publication of the IEEE Engineering in Medicine and Biology Society* 12, 266–278.
49. Witt, J.K., Stefanucci, J.K., Riener, C.R., Proffitt, D.R., 2007. Seeing beyond the target: Environmental context affects distance perception. *Perception* 36, 1752–1768.
50. Zahorik, P., Brungart, D.S., Bronkhorst, A.W., 2005. Auditory distance perception in humans: A summary of past and present research. *Acta Acustica united with Acustica* 91, 409–420.



## 9.5.2 Assistance tactile à la localisation de cibles périphériques pour des personnes à vision tubulaire

<b>Damien Appert *</b> Université de Toulouse & CNRS ; UPS ; IRIT ; Toulouse, France Damien.Appert@irit.fr * <b>co-premiers auteurs</b>	<b>Damien Camors *</b> Université de Toulouse & CNRS ; UPS ; IRIT Toulouse, France Damien.Camors@irit.fr	<b>Jean-Baptiste Durand</b> CNRS; CerCo; Toulouse, France jbdurand@cerco.ups-tlse.fr	<b>Christophe Jouffrais</b> CNRS; Université de Toulouse; UPS; IRIT; France Christophe.Jouffrais@irit.fr
--	--	---	--

### Résumé

La vision tubulaire est une déficience visuelle dans laquelle seul le champ central de la vision est préservé. Elle génère de grandes difficultés dans la vie quotidienne, notamment lorsqu'il s'agit de repérer un objet d'intérêt dans l'environnement. Des informations tactiles, considérées comme discrètes, personnelles et peu invasives, peuvent permettre d'améliorer le comportement de recherche visuelle. Dans cette étude, nous avons conçu quatre techniques tactiles permettant de localiser un point spécifique dans l'espace. Les stimulations tactiles étaient soit un seul stimulus soit un train de stimuli transmis dans un système de coordonnées cartésien ou polaire. Les quatre techniques ont été comparées dans une tâche d'orientation de la tête. La plus efficace des techniques a été évaluée avec une tâche de recherche visuelle dans un environnement virtuel complexe. L'évaluation impliquait dix sujets avec un champ visuel artificiellement restreint à 10°, et un sujet avec une vision tubulaire due à un glaucome. Notre dispositif d'assistance a significativement amélioré l'efficacité de la recherche visuelle d'un facteur trois. Le dispositif pourrait être facilement intégré dans des lunettes intelligentes et détecter des cibles d'intérêt, soit sur demande, soit de façon automatique (par ex. les obstacles potentiels), facilitant ainsi, la recherche visuelle et la perception spatiale de l'environnement.

### Mots Clés :

Interaction tactile ; Technologies d'assistance ; déficience visuelle ; perception spatiale ; interaction en mobilité

### ACM Classification Keywords

H.5.2 [Information Interfaces and Presentation]: User Interfaces — *Input devices and strategies*

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [Permissions@acm.org](mailto:Permissions@acm.org). IHM'15, October 27-30, 2015, Toulouse, France © 2015 ACM. ISBN 978-1-4503- 3844-8/15/10...\$15.00 DOI: <http://dx.doi.org/10.1145/2820619.2820625>

## INTRODUCTION

La rétinite pigmentaire et le glaucome sont des affections dégénératives de la rétine. Ces deux affections conduisent à une perte progressive du champ visuel et éventuellement à la cécité [22]. Ces affections, actuellement incurables, commencent par détériorer la vision périphérique et causent ce

que l'on appelle la vision « tubulaire » ou « en tunnel ». Pour comprendre ce qu'est une vision tubulaire, il faut imaginer regarder à travers un tuyau, le champ visuel est alors réduit à quelques degrés d'angle. Il peut y avoir d'autres symptômes associés comme une gêne visuelle importante dans la pénombre (héméralopie) ou une adaptation plus lente aux changements de luminosité. La vision tubulaire a d'importantes conséquences sur la recherche visuelle [38,46], et par conséquent l'autonomie des patients [25]. Par exemple des tâches usuelles comme trouver un produit sur une étagère ou rechercher un ami dans une foule, deviennent très difficiles et chronophages quand la complexité visuelle de l'environnement s'accroît [49] et que le champ visuel de l'utilisateur est dégradé [34].

## ETAT DE L'ART

De nombreuses technologies ont été conçues pour améliorer la recherche visuelle et la perception de l'espace chez des personnes avec un champ visuel limité. Certaines sont basées sur les capacités visuelles résiduelles, mais d'autres sont dépendantes des modalités auditives et tactiles. Des dispositifs spécifiques à la vision tubulaire ont été élaborés afin d'assister la localisation de cibles qui sont hors du champ visuel. Certains dispositifs s'appuient sur le champ visuel résiduel et reposent sur une projection de l'information présente dans la partie aveugle dans ce champ résiduel. Une première possibilité consiste à compresser l'information visuelle avec une « minification » optique [32,33]. D'autres dispositifs utilisent un prisme qui dévie l'information visuelle dans le champ visuel résiduel [50]. Dans plusieurs dispositifs basés sur la réalité augmentée, une information visuelle spécifique (par ex. la détection de contours) peut être superposée dans la partie intacte du champ visuel [48]. Dans tous ces cas, le champ visuel résiduel est surchargé avec plus d'informations visuelles, ce qui détériore l'acuité visuelle, perturbe l'exploration visuelle, et déforme la relation spatiale entre les objets [19]. Tous ces inconvénients peuvent avoir un impact important sur la recherche visuelle [7].

Des technologies d'assistance basées sur l'audition ont aussi été proposées. Les systèmes de substitution sensorielle convertissent toute une scène visuelle en signaux auditifs (ex. « the VOICe »). Ils peuvent fournir à l'utilisateur suffisamment d'informations pour une tâche de locomotion, de localisation ou de pointage [1]. Cependant, même une tâche simple de pointage requiert beaucoup de temps et d'attention. De plus, les sons naturels de l'environnement, ont un rôle important dans de nombreuses tâches de la vie quotidienne. Les personnes déficientes visuelles sont donc réticentes à encombrer leur système auditif avec des informations additionnelles.

La substitution sensorielle a aussi été proposée en utilisant la modalité tactile. Dans ces systèmes, le principe consiste à transformer l'image provenant d'une caméra en un ensemble de signaux tactiles appliqués sur la langue ou l'abdomen [2,3]. Plusieurs expériences réalisées en condition de laboratoire montrent qu'il est par exemple possible de reconnaître des formes géométriques. Cependant, après quarante ans de recherches sur la substitution sensorielle, aucun dispositif n'est utilisable dans des conditions écologiques [36]. De plus, ils ne sont pas utilisables pour localiser un objet spécifique dans des scènes naturelles pouvant être surchargées.

Plutôt que de substituer toute une scène visuelle, les dispositifs d'assistance peuvent être utilisés pour localiser précisément une cible ou un point d'intérêt dans la scène [5]. Dans le cas d'un utilisateur avec une vision tubulaire, il est possible de guider le regard ou la main vers une région spécifique dans l'environnement. Par exemple, il est possible de faciliter le pointage vers une cible grâce à des sons binauraux en fournissant à l'utilisateur l'azimut et l'élévation de la cible [12,31].

Cependant, comme mentionné précédemment, les informations auditives additionnelles interfèrent avec les sons de l'environnement, ce qui réduit leur utilisabilité et peut masquer un danger potentiel [27].

Dans cette étude, nous avons conçu un dispositif tactile, privé et peu intrusif, permettant de localiser des cibles spécifiques dans l'espace. Après un état de l'art sur les dispositifs tactiles, nous décrivons nos choix de conception, ainsi que les différentes expériences que nous avons menées dans le but d'améliorer la localisation de cibles périphériques. Ces expériences ont été réalisées avec un simulateur de vision tubulaire dans lequel le champ visuel a été restreint artificiellement. Dans la dernière expérience, le dispositif a été testé avec un sujet atteint d'un glaucome.

## **Dispositifs tactiles pour la perception spatiale**

Les stimulations tactiles ne sont pas en compétition avec le traitement de stimuli auditifs ou visuels écologiques. De plus, elles permettent de réaliser des discriminations spatiales et temporelles avec des performances excellentes [21]. Les dispositifs tactiles ont été utilisés dans un grand nombre d'applications concernant le guidage, la navigation, la notification, les alertes, etc. Par exemple, Boll et ses collaborateurs ont utilisé un dispositif tactile sur le poignet ou le torse dans le but d'assister la navigation et le guidage [24,26]. Les dispositifs tactiles peuvent aussi servir de détecteur d'obstacle afin de prévenir toute collision pendant un déplacement [9]. Des travaux antérieurs ont également porté sur la localisation d'une cible spécifique dans l'environnement. Dans ce domaine, deux types d'indices ont été décrits par Lehtinen *et al.* ([35]) : les indices que nous appellerons de *guidage* (« guiding » cues) et d'*anticipation* (« preview » cues). Les indices de guidage fournissent une information continue de l'erreur entre la cible et la partie mobile du corps, généralement la main. Les indices de guidage peuvent inciter l'utilisateur à porter attention à une position précise dans l'espace, ce qui est particulièrement utile quand la cible est cachée dans une scène très fortement encombrée [15,16,35,45]. Par exemple, Lehtinen *et al.* ont placé deux vibreurs dans la paume et deux sur le dos de la main des utilisateurs [35]. La main des utilisateurs était alors continuellement guidée vers une lettre B parmi un ensemble de lettres P distractives. Avec le guidage tactile, ils ont observé une réduction significative du temps nécessaire pour trouver la lettre B. Cependant, les sujets admettaient « se concentrer d'abord sur le signal tactile » puis « sauter avec le regard vers la zone indiquée ». Ce faisant, les sujets ont ignoré le guidage tactile continu. Ils ont plutôt utilisé le signal tactile comme un indice d'anticipation avant de commencer la recherche visuelle dans une région déterminée de l'espace. De plus, ils ont mentionné ne pas être convaincus que le guidage soit utile après avoir commencé la recherche. Par conséquent, il semblerait que les indices de guidage soient moins pertinents que les indices d'anticipation. Les indices d'anticipation indiquent la direction ou la partie de l'espace à atteindre avant qu'une quelconque partie du corps ne bouge. Les indices d'anticipation n'accroissent pas la charge cognitive, notamment pendant la phase d'exploration. Ceci est particulièrement important durant une recherche visuelle complexe, et d'autant plus pour les personnes avec une vision tubulaire.

Les indices d'anticipation ont déjà été utilisés dans une tâche de recherche visuelle [29,37,44]. Les expériences ont montré une réduction significative du temps de recherche. Malheureusement, elles ont été utilisées pour fournir des informations de directions imprécises (par ex. « dans le coin en haut à gauche »), et de surcroît dans un espace non surchargé. Ainsi, la question demeure : les dispositifs tactiles peuvent-ils assister des personnes déficientes visuelles, et plus particulièrement celles avec une vision tubulaire, afin de trouver une cible spécifique dans l'environnement ? Dans ce

cas, les indices d'anticipation doivent indiquer une portion précise de l'espace, suffisamment proche de la cible d'intérêt pour que celle-ci se retrouve directement dans le champ visuel résiduel.

Dans cette étude, nous avons conçu et comparé plusieurs types d'indices tactiles d'anticipation (exp. 1). Nous avons ensuite utilisé le plus efficace et le plus satisfaisant dans une tâche de recherche visuelle avec des sujets voyants sur lesquels le champ visuel était artificiellement réduit (exp. 2). Nous avons finalement validé ce résultat avec un patient atteint de glaucome et présentant un déficit visuel important (exp. 3).

## **EXP 1 : CONCEPTION ET COMPARAISON D'INDICES TACTILES PERMETTANT DE LOCALISER UNE CIBLE PÉRIPHÉRIQUE**

### **Conception du code tactile**

Dans cette première expérience, nous avons conçu et comparé quatre codes tactiles (tacton [5] et référentiel) permettant de localiser une cible dans l'environnement. La conception du dispositif tactile permettant d'accélérer une tâche de recherche visuelle est détaillée dans la seconde expérience.

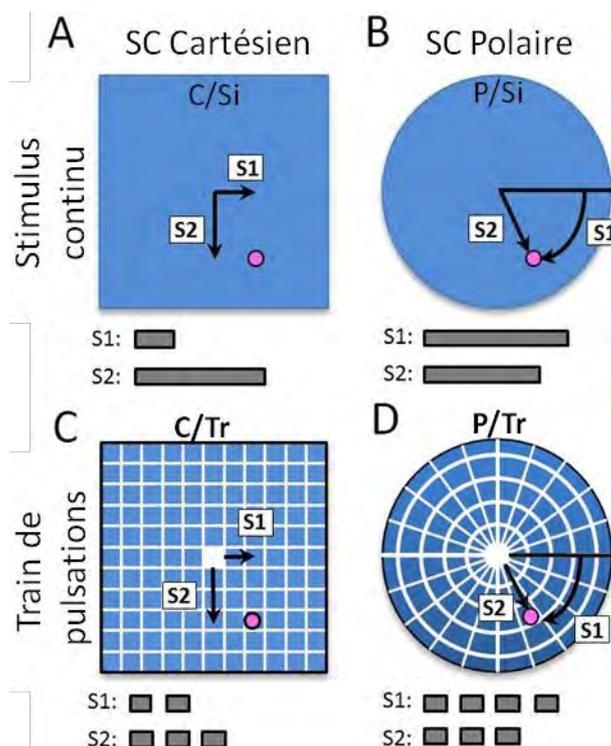
### **Référentiel et système de coordonnées**

Dans un dispositif d'assistance à la localisation de cibles visuelles, les caméras sont indispensables. Elles peuvent être facilement montées sur des lunettes, et peuvent être utilisées pour localiser des cibles spécifiques dans une scène visuelle. Le système peut aussi intégrer un accéléromètre pour compenser une faible résolution temporelle de la caméra et procurer la position des cibles malgré les déplacements de la tête (voir [30] par ex).

Le plus logique des référentiels pour guider le regard vers la cible est celui des yeux avec comme origine, la fovéa. En supposant que la cible ait été localisée par la caméra attachée à la tête, et que la position des yeux dans l'orbite soit connue grâce à un oculomètre portable, l'indice tactile peut alors directement indiquer l'erreur entre la fovéa et la cible. Cependant, les oculomètres portables sont chers et doivent être fréquemment calibrés. Comme les yeux et la tête sont synchronisés pendant les mouvements du regard [42], la tête peut être un référentiel plus pertinent pour assister la recherche visuelle. En effet, cela repose uniquement sur la présence de la caméra montée sur les lunettes. Les indices de guidage peuvent alors directement indiquer l'erreur entre la position de la tête et la position de la cible dans l'image sans aucune transformation de coordonnées.

La localisation d'une cible dans l'espace repose sur trois coordonnées, à savoir l'azimut, l'élévation et la profondeur (ou distance). Cependant la profondeur entre l'utilisateur et la cible n'est pas obligatoire pour la recherche visuelle. En effet, quand la cible est localisée dans le plan de l'image de la caméra, seules deux coordonnées (azimut et élévation) suffisent pour amener le champ visuel en face de la cible. Dans le but de transmettre ces deux coordonnées, nous avons évalué deux systèmes de coordonnées (SC) différents, cartésien et polaire, ancrés sur la tête.

Dans le système cartésien, la position de la cible était décomposée en deux indices tactiles successifs dépendant de l'orientation de la tête. L'azimut (axe horizontal) était présenté en premier, suivi de l'élévation (axe vertical). Dans le système polaire, l'azimut et l'élévation de la cible étaient convertis en deux stimulations successives représentant un angle et une distance radiale (cf. Figure 1).



**Figure 1 : Codage tactile d'une même cible (point rose) avec quatre techniques différentes. Deux codes tactiles ont été utilisés : le stimulus continu (A,B) et le train de pulsations (C,D). Les coordonnées de la cible sont transmises dans les SC cartésien (A,C) ou polaire (B,D). Les deux coordonnées sont rendues par deux stimulations successives (S1, S2) représentées avec des barres grises. Dans cette expérience, la position de la cible était choisie de façon pseudo-aléatoire dans toute la zone bleue.**

### Codage des tactons

Comme mentionné précédemment, les indices d'anticipation ne guident pas progressivement vers une cible. Après une brève stimulation, l'utilisateur doit être capable d'estimer la position de la cible avec une bonne précision. Si la stimulation n'est pas suffisamment explicite, l'utilisateur devra redemander les indices, ce qui n'est pas utilisable. Par conséquent, l'indice tactile doit être aussi bref que possible et permettre d'accélérer la phase suivante de recherche visuelle. Nous avons évalué deux codes tactiles différents (cf. Figure 1).

Le premier est basé sur une seule vibration de durée variable correspondant à la grandeur de la coordonnée transmise. La durée maximale est de 1,5 secondes. Le choix de cette durée était basé sur des pré-tests succincts, non-détaillés ici, qui indiquaient que la discriminable était possible avec cette durée, et qu'une durée plus longue semblait gênante pour les utilisateurs. La durée maximale de 1,5 secondes correspondait à un azimut ou une élévation maximums de 45°.

Dans le second code, une coordonnée est définie comme un train (une séquence) de courtes pulsations de 200 ms séparées entre elles par 100 ms. Le nombre de pulsations dans le train correspond à la zone cible dans l'espace frontal, avec chaque pulsation indiquant une zone de 9°. Le train comporte cinq pulsations au maximum, correspondant au même maximum de 45°.

Ces deux codes tactiles ont été évalués dans les deux systèmes de coordonnées décrits précédemment. Par conséquent, nous avons évalué quatre techniques notées: stimulus continu dans un SC Cartésien (C/Si), train de pulsations dans un SC Cartésien (C/Tr), stimulus continu dans un SC Polaire (P/Si) et train de pulsations dans un SC Polaire (P/Tr).

## Position du dispositif tactile

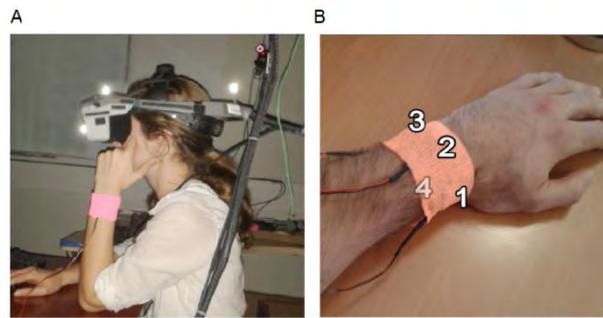
Les dispositifs tactiles peuvent être placés sur différentes parties du corps humain. Afin d'assister la recherche visuelle, une solution serait de placer un dispositif tactile autour des yeux. La correspondance entre la stimulation et la direction souhaitée regard serait probablement très intuitive. Cependant, la sensibilité tactile sur le visage est très importante. Le dispositif tactile pourrait rapidement provoquer une irritation. De plus, tous les sujets interrogés étaient réticents à porter un dispositif sur le visage.

Les indices directionnels sont le plus souvent transmis par des dispositifs tactiles attachés à une ceinture, sur le dos ou sur les mains. La ceinture est une solution à la fois élégante et non-intrusive pour transmettre l'azimut [14,26], mais pas pour l'élévation. Le dos et le torse présentent une importante surface tactile, ce qui permet d'y placer un nombre important de vibreurs. Les premières études ont utilisé des vibreurs placés sur une chaise dans le but de transmettre des indices d'anticipation sur le dos [29,37]. Ce dispositif n'était pas mobile ce qui en limitait l'usage. Plus récemment, les vibreurs ont été placés dans des gilets [13], ce qui peut freiner l'acceptabilité du dispositif d'assistance.

La main est souvent utilisée [15,16,35,45], ce qui est évidemment un avantage quand un mouvement de celle-ci est requis. En effet, cela évite de transformer les coordonnées entre le capteur et l'effecteur, et par conséquent facilite les tâches de pointage. Cependant, les dispositifs qui sont placés dans la paume de la main posent des problèmes de préhension et de manipulation. Le dos de la main ou le poignet sont de bonnes alternatives. Le poignet peut offrir une discrimination efficace des stimuli [10,40]. Contrairement au dos et à la hanche, le poignet est impliqué dans le processus visuo-moteur permettant la coordination œil-main [23]. Un autre argument pour fixer le dispositif tactile au poignet est l'arrivée des « smartwatches » dans la communauté des déficients visuels. Ces dispositifs possèdent des vibreurs pouvant être activés à distance (par ex. en Bluetooth). Si nécessaire, des vibreurs additionnels peuvent être attachés au bracelet de la montre.

## Évaluation expérimentale

Nous avons évalué l'utilisabilité des quatre techniques dans une tâche d'orientation de la tête. Les sujets portaient un casque de réalité virtuelle fournissant un champ visuel restreint à dix degrés et un vibreur attaché au poignet (cf. Figure 2A). Pour chaque condition, ils devaient estimer le plus précisément possible la position d'une cible après avoir reçu les deux indices tactiles correspondant à sa coordonnée. Nous avons utilisé un design expérimental à 2x2 facteurs, avec le système de coordonnées (Cartésien vs. Polaire) et le code tactile (continu vs. train) comme facteurs principaux. Chaque sujet a passé les quatre conditions dans un ordre pseudo-aléatoire. Pour chaque condition, les vingt premiers essais étaient utilisés comme une phase de familiarisation et n'étaient pas inclus dans les résultats. Ils étaient suivis par quarante essais où les mouvements de la tête étaient enregistrés.



**Figure 2 : Photos d'un sujet voyant équipé avec l'interface tactile au poignet et le casque de réalité virtuelle permettant de simuler la vision tubulaire. Les mouvements de la tête sont enregistrés par un système infrarouge, visible en A. Dans l'expérience 1, les sujets ne disposent que d'un seul vibreur, placé sur le poignet par une bande cohésive (A). Dans les expériences 2 et 3, quatre vibreurs sont placés sur le poignet gauche (B). Le quatrième vibreur, placé sur la partie antérieure du poignet n'est pas visible sur la photo.**

Dans cette première expérience, un quadrant du référentiel était sélectionné aléatoirement pour chaque essai et était indiqué par un indice visuel. Dans le repère cartésien, deux flèches rouges étaient affichées sur les deux demi-axes du quadrant choisi. Dans le repère polaire, une barre rouge indiquait un demi-axe parmi les quatre possibles. Les deux coordonnées correspondant à la position de la cible dans ce quadrant étaient alors transmises par un seul vibreur, séparées par un silence de 0,5s. Ce couple de modalités contraignait la stimulation sur un seul vibreur et permettait aux sujets de se concentrer uniquement sur le décodage de la stimulation tactile des deux coordonnées.

À la fin de chaque condition, les sujets devaient remplir un questionnaire d'utilisabilité (SUS) sur la technique qu'ils venaient d'utiliser. Pour chaque sujet, l'expérience a duré environ une heure et demie.

### **Participants**

Six sujets (quatre hommes et deux femmes, âgés de  $27,2 \pm 2,6$  ans) ont réalisé l'expérience. Conformément à la Déclaration d'Helsinki, ils ont donné leur consentement éclairé par écrit, avant de participer à l'étude. L'étude a été approuvée par un comité local d'éthique (CLERIT).

### Dispositif expérimental et stimuli

Les sujets étaient placés dans une scène visuelle artificielle contenant seulement un fond bruité (ressemblant à un ciel étoilé). Le fond était utilisé pour fournir un feedback visuel pendant les mouvements de la tête, mais ne fournissait aucune information sémantique ou symbolique pouvant servir de repère. Pendant l'expérience, le champ visuel du sujet dans le casque de réalité virtuelle était réduit à dix degrés par un masque noir couvrant toute la périphérie du champ visuel.

Les stimuli visuels étaient présentés dans un casque de réalité virtuelle NVisor® SX-60 HMD (NVIS Inc., USA), avec une résolution de 1280x1024 pixels, équivalant à 44x33 degrés d'angle visuel. L'orientation de la caméra (virtuelle) était directement liée à celle de la tête grâce à un système de capture de mouvement cadencé à 60 Hz (OptiTrack®, Natural Point, USA). Les sujets étaient donc libres de regarder autour d'eux (virtuellement) en bougeant la tête. Les sujets étaient assis sur une chaise fixe qui limitait les mouvements de leur tronc. Ils étaient libres de bouger les yeux mais une croix rouge présentée au centre du champ visuel servait de point d'ancrage. L'environnement virtuel (EV) était généré et animé par le moteur 3D open source Irrlicht (v1.8). Le dispositif tactile, spécialement conçu pour l'expérience, utilisait une carte Arduino® ATmega1280 et des vibreurs

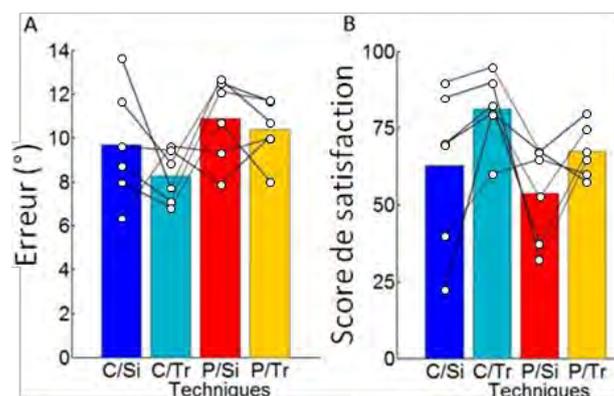
VPM2. Les stimuli tactiles étaient transmis via un seul vibreur placé sur le dessus du poignet (cf. Figure 2A). Le vibreur était calibré pour une fréquence de 250 Hz, optimale pour une stimulation tactile sur la peau [39].

### Tâche comportementale

Les sujets ont été testés dans une tâche d'orientation de la tête en présence de l'expérimentateur. Le sujet devait estimer le plus précisément possible, la position d'une cible dans un environnement virtuel. A chaque nouvel essai, une cible était générée au hasard dans un espace de 90°x90° d'angle visuel, centré sur la position de la tête. En début d'expérience, chaque sujet devait définir une position de repos, avec la tête bien droite. Cette position était symbolisée par une sphère verte (2° de diamètre) et allumée quand nécessaire. A chaque début d'essai, le sujet devait aligner le centre de son champ visuel (indiqué par une croix rouge) et la sphère verte. Après une durée de 1,4 s à 2 s, les coordonnées de la cible étaient transmises avec le dispositif tactile. A la fin de la transmission, la sphère verte s'effaçait. Le sujet devait alors orienter sa tête vers la position de la cible le plus précisément possible, puis valider cette orientation en appuyant sur la barre d'espace. La validation était toujours suivie par l'affichage de la position de la cible (sphère rouge) et de la position estimée (sphère blanche). Cette phase permettait au sujet d'avoir un feedback de sa performance à chaque essai. Un nouvel essai commençait lorsque le sujet était revenu à sa position de repos.

### Résultats

L'erreur de pointage était calculée pour chaque essai, puis moyennée par sujet et par technique. Elle correspond à la distance entre la position de la cible et celle pointée par le sujet. Une analyse de variance à mesures répétées (ANOVA-RM) a montré un effet significatif du référentiel (Cartésien vs. Polaire ;  $p < 0,05$ ) mais pas d'effet du code tactile (Continu vs. Train), ni aucune interaction entre ces deux facteurs. La moyenne de l'erreur de pointage liée au référentiel cartésien ( $9,0 \pm 2,1$ ) était plus réduite que celle liée au référentiel polaire ( $10,6 \pm 1,6$ ) (cf. Figure 3A).



**Figure 3 : efficacité (A) et satisfaction (B) pour les quatre techniques conçues. A) L'erreur moyenne de pointage pour chacune des techniques est représentée par des barres de couleur. La moyenne de l'erreur de pointage par sujet est représentée avec des points blancs. La technique C/Tr montre l'erreur moyenne la plus faible ( $8,2^\circ$ ) avec l'écart-type le plus faible ( $1,2^\circ$ ). B) Le score moyen de satisfaction (SUS) pour chacune de ces techniques est représenté par des barres de couleur. Le score de satisfaction par sujet est représenté avec des points blancs. La technique C/Tr reçoit le score le plus élevé avec  $81,3 \pm 12,0$ .**

Afin de révéler l'existence d'un biais cognitif ou moteur systématique, comme une sous-estimation ou une surestimation des deux coordonnées, nous avons calculé la moyenne de l'erreur de pointage dans chaque référentiel. Nous n'avons pas observé de biais pour le référentiel cartésien. Toutefois

dans la condition P/Tr, les sujets sous-estimaient considérablement et systématiquement la distance angulaire (test-t,  $p < 0,05$ , après correction de Bonferroni). De même, dans la condition P/Si, les sujets surestimaient la distance angulaire (test-t,  $p < 0,05$ , après correction de Bonferroni). De plus, indépendamment des techniques, les résultats montrent un effet significatif des positions spatiales des cibles (proches/éloignées) sur la précision de pointage. Les sujets sont moins bons à pointer une cible éloignée, avec un ratio moyen de 1,7.

Nous avons finalement comparé les scores d'utilisabilité pour chaque technique par une ANOVA à deux facteurs. Nous avons observé un effet significatif du code tactile sur l'utilisabilité ( $p < 0,05$ ). Le « train de pulsations » a été considéré comme plus satisfaisant que le « stimulus continu » (test-t,  $p < 0,05$ ) (cf. Figure 3B). Pendant les interviews finales, quatre sujets ont considéré la technique C/Tr comme étant leur préférée. Les deux autres sujets ont considéré la technique C/Tr comme leur deuxième choix.

### **Conclusions sur la première expérience**

Les réponses de pointage sont plus précises dans le référentiel cartésien. Nous n'avons pas observé de biais systématique et la moyenne de l'erreur est de 9.0 degrés d'angle visuel seulement. Au contraire, les stimuli tactiles dans le référentiel polaire ont conduit à des résultats significativement moins précis, avec des biais systématiques. Par conséquent, il apparaît que le référentiel cartésien est plus précis pour communiquer la position d'une cible à l'aide du dispositif tactile.

La satisfaction liée à ces deux systèmes n'est pas significativement différente. Cependant, le score est meilleur lorsque les coordonnées sont communiquées avec le « train de pulsations » plutôt qu'avec le « stimulus continu ». En combinant précision et satisfaction, la technique reposant sur un train de pulsations dans un référentiel cartésien (C/Tr) est la meilleure parmi les quatre que nous avons conçues. De plus, cette technique a montré la plus faible variabilité inter sujets et la meilleure moyenne au SUS. Le score SUS de cette technique a atteint  $81 \pm 12$ , ce qui correspond à un « bon » dispositif [4].

## **EXP 2 : RECHERCHE VISUELLE D'UNE PETITE CIBLE DANS UN SALON VIRTUEL ENCOMBRÉ**

Dans cette deuxième expérience, nous avons évalué la technique la plus efficace et préférée (C/Tr) dans des conditions proches d'une situation écologique. Nous avons simulé un champ visuel sévèrement dégradé ( $10^\circ$  d'angle visuel) avec un casque de réalité virtuelle. Les sujets étaient immergés dans un salon et avaient pour tâche de localiser le plus rapidement possible une cible dans cette salle.

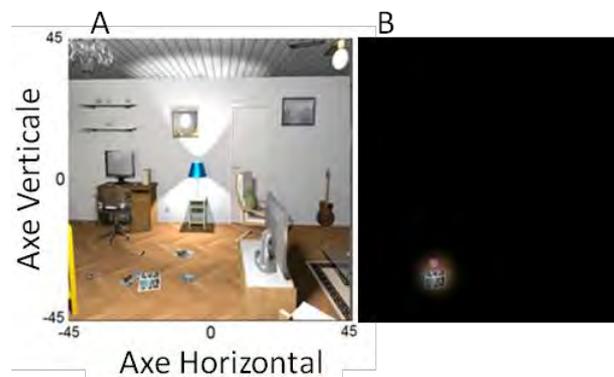
### **Participants**

Dix sujets (six hommes et quatre femmes,  $25,4 \pm 1,6$  ans) ont réalisé l'expérience. Ils ont donné par écrit leur consentement éclairé avant de participer à l'étude, conformément à la Déclaration d'Helsinki. L'étude a été approuvée par un comité d'éthique local (CLERIT).

### **Environnement virtuel, dispositif tactile et stimuli**

La configuration expérimentale utilisée dans cette expérience était identique à celle utilisée dans l'expérience précédente.

Une scène virtuelle représentant un salon a été créée avec « Sweet Home 3D<sup>®</sup> ». Une grande quantité d'objets (lampe, chaise, livre, etc.) a été placée de façon homogène dans l'ensemble de la scène visuelle (cf. Figure 4A). La cible (sphère rose) était placée au hasard dans la scène virtuelle, dans une partie de 90° x 90° d'angle visuel. La scène virtuelle était centrée sur la position de la tête avant chaque essai. Nous nous sommes assurés que la cible et les éléments de la scène visuelle ne pouvaient pas être confondus. Par conséquent, les teintes roses n'étaient pas utilisées dans la scène visuelle. Toutefois, afin d'éviter un effet « pop out » de la cible, celle-ci était légèrement transparente. Contrairement à une situation écologique, la cible était disposée de façon aléatoire dans la scène et non sur des objets. Ce choix permet de mesurer des temps de recherche non biaisés par le contexte, dépendant uniquement de l'assistance tactile. Pendant toute l'expérience, le champ visuel à l'intérieur du casque virtuel était limité à 10°. Un masque noir était appliqué sur le pourtour de la scène, ce qui simulait une vision tubulaire sévère (cf. Figure 4B).



**Figure 4 : A – Portion de l'environnement visuel. La cible rose apparaît en bas à gauche. B – Exemple de la simulation artificielle de la vision tubulaire de 10°.**

Comme résultat de la première expérience, les deux coordonnées de la cible étaient transmises avec deux « trains de pulsations » tactiles encodés dans un référentiel cartésien. Cependant, afin d'être dans des conditions plus réalistes, les indices tactiles étaient transmis dans cette expérience par quatre vibreurs disposés autour du poignet gauche comme s'ils étaient répartis sur un bracelet (cf. Figure 2B). Les vibreurs étaient attachés avec un bandage. Nous avons choisi les axes et leurs directions de façon arbitraire, en fonction d'une position fréquente de la main gauche au repos, la paume vers le bas. Les coordonnées horizontales étaient transmises par les vibreurs placés vers le pouce (direction droite) et le petit doigt (dir. gauche). Les coordonnées verticales étaient transmises par les vibreurs de la partie dorsale (dir. haut) et ventrale (dir. bas). Comme dans la première expérience, les pulsations tactiles avaient une durée fixe de 200 ms séparées par 100 ms. Le nombre de pulsations dans le train correspondait à la zone cible dans l'espace, avec chaque pulsation indiquant une zone de 9°. Le maximum de stimuli était de cinq et correspondait à l'angle maximal de 45°.

### **Procédure expérimentale**

Les sujets étaient évalués individuellement dans une tâche de recherche visuelle en présence de l'expérimentateur. Le sujet devait estimer le plus rapidement possible la position d'une cible dans un environnement virtuel. Comme dans la tâche précédente, la position de repos de la tête a été définie

au début de l'expérience, et était marquée dans l'EV par une sphère verte (2° diamètre) quand nécessaire. Le sujet devait maintenir sa position de repos pendant une durée aléatoire entre 700 ms et 1400 ms. Puis, les coordonnées de la cible étaient transmises au sujet. À la fin de la stimulation, la sphère verte disparaissait, et les lumières de la scène s'allumaient. Le sujet devait alors rechercher et trouver la cible le plus rapidement possible, puis valider sa réponse en appuyant sur la barre d'espace.

Pour chaque technique, le sujet commençait par une phase de familiarisation incluant quinze essais qui n'étaient pas enregistrés. Ensuite, la phase de test était répartie en deux blocs de vingt essais où le temps de recherche et l'ensemble des mouvements de tête étaient enregistrés.

Dans cette expérience, la technique d'assistance était comparée à une condition contrôle absolument identique, sauf qu'elle n'incluait aucun indice tactile concernant l'emplacement de la cible. Chaque sujet a passé quarante essais répartis en deux blocs correspondant aux deux techniques. Cinquante pour cent des sujets ont commencé par la condition contrôle. Les autres ont commencé par la technique d'assistance.

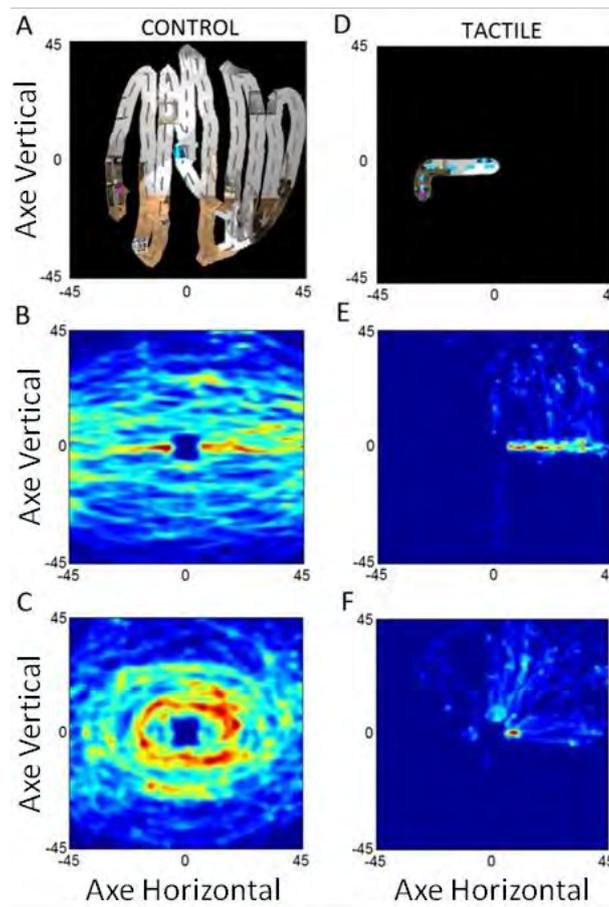
L'ensemble de l'expérience a duré environ soixante minutes par sujet. À la fin de l'expérience, nous avons demandé aux sujets de remplir un questionnaire de satisfaction (SUS) concernant la technique d'assistance.

### **L'analyse des données**

Nous avons mesuré le temps de recherche pour chaque essai. Il correspond à la durée écoulée entre l'envoi de la première stimulation et la validation du sujet. Cela correspond au temps nécessaire pour effectuer la recherche visuelle. La précision temporelle de 60 Hz nous a permis d'analyser finement, pour chaque essai, la trajectoire de la tête au cours de la recherche visuelle. La distance totale parcourue a été calculée pour chaque essai. De plus, nous avons calculé un ratio exprimant la distance totale sur la distance optimale. La distance optimale est la plus courte distance entre le centre du référentiel et la cible. Ce ratio est complémentaire au temps de recherche.

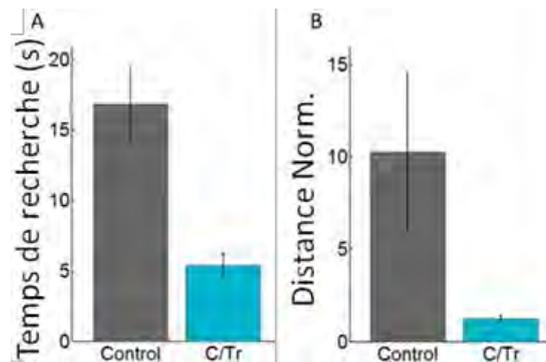
### **Résultats**

Dans la condition contrôle, nous avons observé différentes stratégies de recherche visuelle en fonction des sujets (cf. Figure 5). Certains sujets ont effectué un motif de balayage sur l'axe vertical ou sur l'axe horizontal (Fig. 5A et B resp.). D'autres sujets ont effectué des cercles concentriques autour du centre de la scène visuelle (Fig. 5C). Ces motifs particuliers n'ont pas été observés en présence d'indices tactiles. Dans ces conditions, les sujets déplaçaient tout d'abord leur tête sur l'axe horizontal, puis sur l'axe vertical (Fig. 5D et E). Nous avons également observé chez certains sujets, des mouvements en diagonale. Cela suggère une bonne intégration des deux coordonnées (horizontales et verticales) avant le mouvement de la tête (Fig. 5F).



**Figure 5 : Exemples de mouvements de tête observés pendant une recherche visuelle dans les conditions contrôle (A, B, C) et tactile (D, E, F). A - Balayage sur l'axe vertical. B – Carte de densité d'une recherche avec un balayage principalement sur l'axe horizontal. C – Carte de densité d'une exploration des cercles concentriques. D – Mouvements ordonnés le long de l'axe horizontal puis vertical suivant l'ordre de transmission des deux indices tactiles. E – Carte de densité d'un mouvement commençant essentiellement par l'axe horizontal. F- Carte de densité d'un mouvement intégré. Ici, le sujet a été capable d'intégrer les deux coordonnées avant de bouger, puis de déplacer la tête en faisant des mouvements obliques vers les cibles. Chaque carte de densité représente les mouvements de tête d'un sujet pendant un essai. La position de la cible et les mouvements de tête ont été normalisés avec une transformation symétrique sur l'axe horizontal puis vertical, en plaçant les cibles dans le coin supérieur droit. Le rouge indique des densités élevées et le bleu foncé l'absence de position de tête.**

Nous avons calculé les temps de recherche médians par sujet et par technique. Les moyennes de ces temps de recherche médians sont présentées dans la Figure 6A. Les sujets passaient en moyenne  $16,8 \pm 2,5$  secondes pour trouver une cible dans la condition contrôle ; et  $5,4 \pm 0,8$  secondes avec des indices tactiles. Nous avons observé une différence similaire pour chaque sujet. La différence est très significative (test-t,  $p < 0,001$ ). Nous avons également calculé la distance médiane parcourue par sujet et par condition. Chaque distance a été comparée à la distance optimale (distance la plus courte entre la position de la cible et la position initiale du sujet). En condition contrôle, la distance médiane était en moyenne dix fois plus importante que la distance optimale (cf. Figure 6B). Avec des indices tactiles, ce ratio est en moyenne de  $1,2 \pm 0,2$ . Nous avons observé un ratio similaire pour tous les sujets. Par conséquent, l'analyse a montré une différence très significative (test-t  $p < 0,001$ ). Nous n'avons observé aucune amélioration de leur temps de recherche, ni de la distance parcourue, que ce soit dans les conditions contrôle ou tactile.



**Figure 6 : Performance de la recherche visuelle (A – temps de recherche et B- distance parcourue) avec et sans indices tactiles. En gris, la condition contrôle sans indices ; en cyan, la technique d’assistance (C/Tr, voir expérience 1).**

Une des erreurs observées avec la technique d’assistance était liée à l’inversion de la direction sur les axes. Nous avons calculé les erreurs d’inversion selon une règle simple. Une erreur est détectée lorsque la tête se déplace dans la direction opposée, et atteint le seuil de cinq degrés d’angle visuel. En moyenne, les sujets ont fait  $4,5 \pm 3,3$  % (resp.  $8,8 \pm 4,0$  %) d’erreurs d’inversion sur l’axe horizontal (resp. vertical). Le score de satisfaction (SUS) pour la technique d’assistance était en moyenne de  $82 \pm 16$ , ce qui est considéré comme « bon » [4].

### Conclusions de la deuxième expérience

L’expérience a montré que la recherche visuelle avec une assistance tactile était en moyenne trois fois plus courte. De plus, l’assistance tactile était fiable et satisfaisante. L’expérience a donc clairement démontré l’utilisabilité de la technique pour assister une recherche visuelle dans un environnement visuellement encombré.

### EXP 3 : ETUDE DE CAS SUR UN SUJET AVEC UNE VISION TUBULAIRE

Dans cette dernière expérience, nous avons souhaité vérifier que les résultats observés avec une simulation sont vrais pour des utilisateurs déficients visuels. Nous avons donc répété l’expérience avec un sujet ayant un champ visuel restreint dû à un glaucome.

#### Participant

Un homme de 47 ans a réalisé cette expérience complémentaire. Il s’est porté volontaire suite à une annonce passée dans un centre d’éducation spécialisée pour déficients visuels (IA). Il a donné son consentement éclairé par écrit avant de participer à l’étude conformément à la Déclaration d’Helsinki. Le sujet présentait un champ visuel bilatéral de  $10^\circ \times 15^\circ$ . Ce sujet avait une perception altérée des couleurs et une adaptation lente à la luminosité, ce qui est souvent observé chez ces patients [43].

#### Environnement virtuel, dispositif tactile et cibles

Le sujet était placé dans les mêmes conditions que précédemment, excepté les points suivants. La couleur de la cible a été changée en un dégradé de couleur (du jaune au blanc), et la transparence a été supprimée. Dans le but de réduire l’inconfort dû aux changements de luminosité, nous l’avons gardée constante pendant toute l’expérience. Évidemment, le champ visuel n’était pas artificiellement réduit par un masque noir.

## Procédure expérimentale

La procédure était la même que dans l'expérience précédente, à l'exception que la lumière n'était jamais éteinte pendant les essais. Le sujet a tout d'abord réalisé une série de trente essais (dont dix pour se familiariser) sans assistance tactile. Puis, il a réalisé une série de trente-cinq essais (dont quinze pour se familiariser) avec assistance tactile.

## Résultats

Nous avons calculé la durée médiane pour trouver la cible avec et sans indices tactiles. Elle était significativement réduite avec les indices tactiles (9,4s) comparée à la condition contrôle (27,3 s) (Wilcoxon  $p < 0,001$ ). Le score de satisfaction (SUS) pour la technique d'assistance était de 72,5, ce qui est considéré comme « bon » [4].

Enfin, la discussion ouverte avec l'utilisateur nous a permis de recueillir des résultats qualitatifs. Celui-ci a notamment remarqué que « c'est vraiment une technique prometteuse » qui va « l'aider à localiser un objet ». Si elle était disponible, il s'en « servirait régulièrement ». Il a aussi précisé qu'il a « eu parfois quelques difficultés à détecter les vibrations », et que, par conséquent, « il faudrait ralentir la fréquence des trains de stimulations ».

## Discussion

Dans la première expérience, nous avons conçu quatre codes tactiles différents transmis par un seul vibreur fixé au poignet. Les indices d'anticipation étaient transmis avant tout mouvement de tête, et visaient à fournir l'assistance pour localiser une cible dans un espace frontal. La première observation est que tous ces codes permettent à l'utilisateur d'estimer correctement la position d'une cible dans un repère égocentrique lié à l'orientation de la tête, ce qui est similaire à la façon dont la mémoire peut stocker des informations spatiales [41]. Toutefois, la comparaison directe de ces quatre techniques dans une tâche d'orientation de la tête avec une vision tubulaire artificielle a montré qu'une succession de pulsations transmise dans un système de coordonnées cartésien était la plus efficace et la plus satisfaisante. Dans cette technique, deux trains de pulsations correspondent à l'abscisse et l'ordonnée de la cible relatives à l'orientation courante de la tête. Et, un nombre croissant de pulsations dans le train indique que la portion recherchée de l'espace est plus éloignée de la position actuelle.

Dans l'expérience suivante, nous avons évalué cette technique, considérée comme la plus efficace et la plus satisfaisante, dans une situation la plus écologique possible. Ainsi, les sujets avaient un champ visuel équivalent à une vision tubulaire de 10°. Ils devaient trouver une petite cible dans un salon encombré. Les sujets portaient un dispositif tactile conçu par nos soins, composé de quatre vibreurs répartis autour du poignet. Nous avons choisi d'associer une direction par vibreur (Haut, Bas, Gauche, Droite) et un axe par paire de vibreurs (vertical vs. horizontal). L'évaluation de la technique a montré l'efficacité de ce design puisqu'il a permis une estimation très efficace de la localisation de la cible dans un environnement surchargé. Encore une fois, il est important de noter que ce n'est pas une technique de guidage mais d'anticipation. L'information est fournie de façon brève, avant tout mouvement, et permet d'estimer la position de la cible dans un repère centré sur la tête. La recherche visuelle, réalisée dans un environnement encombré, était trois fois plus rapide avec les indices d'anticipation qu'en leur absence. La distance parcourue pour trouver la cible était aussi dix fois plus courte avec les indices. De plus, tous les sujets testés ont montré une réduction significative

du temps de recherche et de la distance parcourue. Une autre observation importante est que les directions étaient dans la grande majorité des cas (>90%) correctement interprétée, ce qui suggère d'une part que le bracelet permet une bonne discrimination des vibreurs autour du poignet, mais aussi que le code que nous avons utilisé est facile à décoder.

Comme nous l'avons mentionné, le choix de la durée des stimulations tactiles a été basé sur des pré-tests succincts permettant de se situer dans une zone de discriminabilité confortable. Au niveau physiologique, la discrimination temporelle entre deux stimulations d'une même partie du corps est en moyenne entre 25 ms et 40 ms [28]. Il serait donc possible d'envisager de raccourcir le temps de stimulation. Cependant nous avons noté que le ressenti des utilisateurs était très variable. Les sujets voyants (expériences 1 et 2) n'ont fait aucun commentaire sur une éventuelle optimisation des stimulations, alors que le sujet déficient visuel (expérience 3) aimerait les ralentir. Il est probable que de nombreux facteurs influencent la perception des stimuli tactiles (déficience visuelle éventuelle mais aussi fatigue, âge etc.). Il serait donc nécessaire de réaliser une étude systématique permettant d'optimiser la durée et la fréquence des stimuli. Nous pourrions aussi imaginer un système dans lequel l'utilisateur peut aisément choisir ces deux paramètres.

Au cours de l'expérience, nous n'avons pas observé une augmentation ou une diminution des performances de la recherche visuelle entre les essais. Cette observation signifie que l'utilisateur avait besoin de moins de quinze essais, correspondant à la phase de familiarisation, pour atteindre un bon niveau de performance. Cela suggère un apprentissage rapide du code tactile. Enfin, bien que notre prototype soit rustique et bon marché (quatre vibreurs attachés avec un bandage), il a été de façon surprenante très apprécié. Avec des scores de satisfaction et d'efficacité très bons, il apparaît que ce dispositif, destiné aux personnes ayant un champ visuel restreint, pourrait être très utilisable pour transmettre une connaissance a priori de cibles environnantes. Une première amélioration de notre approche serait d'augmenter sa portée, i.e. d'indiquer la position d'objets dans une zone plus large, voire de 360° autour de l'utilisateur. Afin d'accroître les performances de l'approche nous nous focaliserons sur la réduction du temps de transmission des coordonnées tout en conservant la bonne discriminabilité des vibreurs. La diminution de la durée des vibrations et des intervalles serait une bonne piste, mais il y en a d'autre comme par exemple la parallélisation du message avec de la multimodalité ou encore améliorer la version stimuli continu afin de gagner en précision.

Il était cependant important de valider notre approche sur un sujet déficient visuel. Différentes études comportementales ont déjà montré que cette même approche est fiable dans des tâches de cognition spatiale [17,18,47]. Néanmoins, les personnes avec une vision tubulaire peuvent présenter d'autres déficits tels que la faible détection des contrastes, la confusion des couleurs, ou la perte de l'acuité visuelle [20,52]. Or ces déficits ne sont pas simulés dans cette simulation. Ils peuvent également développer des stratégies particulières pour la recherche visuelle lors de la perte progressive de leur champ visuel. Ces stratégies peuvent entraver les résultats que nous avons observés dans cette expérience. Par conséquent, nous avons évalué le dispositif sur un sujet ayant un champ visuel restreint en raison d'un glaucome. Son champ visuel était cohérent avec la simulation, mais sa capacité à détecter les couleurs était aussi altérée. Nos résultats montrent qu'avec des indices d'anticipation, il était trois fois plus rapide pour trouver une cible. Le ratio est similaire à celui observé avec la simulation, bien que le sujet déficient visuel soit légèrement moins rapide pour trouver les cibles dans les deux conditions. Une capacité plus faible de détection des contrastes et des couleurs pourrait expliquer cette différence. Au cours du débriefing, le sujet déficient visuel mentionnait qu'une partie de son attention était portée sur la discrimination des vibreurs actifs.

Toutefois, le sujet a montré un grand intérêt pour le dispositif d'assistance et le score (72,5) qu'il a donné confirme sa satisfaction. Néanmoins, la suppression de la transparence a probablement augmenté l'effet «pop out» de la cible et faciliter sa détection. Cependant dans l'expérience 3, l'effet pop-out était identique dans les deux conditions. Et après une analyse supplémentaire, le pourcentage de réduction en termes de durée entre la condition contrôle et la technique d'assistance observé dans les exp. 2 et 3 était similaire (67% pour l'exp. 2 et 65% pour l'exp. 3), ce qui confirme l'hypothèse que la diminution du temps de recherche dans l'exp. 3 est bien due aux informations tactiles.

La prochaine étape consistera à tester l'utilisabilité du dispositif en situation réelle. Il peut facilement être intégré dans des lunettes intelligentes actuellement en développement (i.e. Google Glass®) ou même commercialisées (OrCam®). Tous ces dispositifs sont capables de reconnaître et localiser des cibles d'intérêts, et de calculer leur position relative à celle de la tête. Les algorithmes de vision par ordinateur sont en constante progression, et reconnaissent déjà des visages ou des objets avec précision. Il est possible de distinguer des milliers d'objets sur la base d'images en temps réel [11]. Ces cibles d'intérêt pourraient être détectées soit sur demande grâce à des commandes vocales, soit être suggérées par le dispositif de façon automatique (par ex. des obstacles potentiels ou des cibles récurrentes), ce qui améliorerait la cognition spatiale de l'utilisateur. Une version commerciale de notre bracelet tactile pourrait facilement être intégrée dans le bracelet d'une « smartwatch » et communiquer par Bluetooth. Une étude récente a notamment mis en évidence l'intérêt croissant que les personnes déficientes visuelles portent aux smartphones et autres appareils portables [51]. Il est également intéressant de noter qu'en plus de la recherche visuelle, un dispositif tactile sur le poignet pourrait être utile dans d'autres situations quotidiennes problématiques pour les déficients visuels, comme la navigation piétonne [6] ou le transport [8].

## REMERCIEMENTS

Les auteurs remercient Louis-Pierre Bergé, Gary Perelman et Victor Vergnieux pour l'aide précieuse lors de l'élaboration des tâches comportementales. Ce travail a été financé par des subventions du CNRS-INSB et UNAEV.

## BIBLIOGRAPHIE

1. Auvray, M., Hanneton, S., and O'Regan, J.K. Learning to perceive with a visuo-auditory substitution system: localisation and object recognition with 'the vOICe'. *Perception* 36, 3 (2007), 416–30.
2. Bach-Y-Rita, P., Collins, C.C., Saunders, F.A., White, B., and Scadden, L. Vision Substitution by Tactile Image Projection. *Nature* 221, 5184 (1969), 963–964.
3. Bach-y-Rita, P., Kaczmarek, K.A., Tyler, M.E., and Garcia-Lara, J. Form perception with a 49-point electrotactile stimulus array on the tongue: a technical note. *Journal of rehabilitation research and development* 35, 4 (1998), 427–430.
4. Bangor, A., Kortum, P.T., and Miller, J.T. An Empirical Evaluation of the System Usability Scale. *International Journal of Human-Computer Interaction*, (2008), 574–594.
5. Brewster, S.S. a. and Brown, L.M.L.M. Tactons: structured tactile messages for non-visual information display. *Proceedings of the fifth conference on Australasian user interface (AUIC '04)*, Australian Computer Society, Inc. (2004), 15–23.
6. Brock, A., Kammoun, S., Macé, M., and Jouffrais, C. Using wrist vibrations to guide hand movement and whole body navigation. *I-Com* 13, 3 (2014), 19–28.
7. Brooks, D.I., Rasmussen, I.P., and Hollingworth, A. The nesting of search contexts within natural scenes: evidence from contextual cuing. *Journal of experimental psychology. Human perception and performance* 36, 6 (2010), 1406–18.

8. Brunet, L., Megard, C., Panëels, S.A., *et al.* 'Invitation to the voyage': The design of tactile metaphors to fulfill occasional travelers' needs in transportation networks. *2013 World Haptics Conference, WHC 2013*, IEEE (2013), 259–264.
9. Cassinelli, A., Reynolds, C., and Ishikawa, M. Augmenting spatial awareness with Haptic Radar. *2006 10th IEEE International Symposium on Wearable Computers*, IEEE (2006), 61–64.
10. Chen, H., Santos, J., Graves, M., Kim, K., and Tan, H.Z. Tactor localization at the wrist. *Proceedings of the EuroHaptics 2008 Conference*, (2008), 209–218.
11. Dean, T., Ruzon, M.A., Segal, M., Shlens, J., Vijayanarasimhan, S., and Yagnik, J. Fast, Accurate Detection of 100,000 Object Classes on a Single Machine. *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*, (2013), 1814–1821.
12. Dramas, F., Oriola, B., Katz, B.F.G., Thorpe, S.J., and Jouffrais, C. Designing an assistive device for the blind based on object localization and augmented auditory reality. *Proceedings of the 10th international ACM SIGACCESS conference on Computers and accessibility - Assets '08*, (2008), 263–264.
13. Van Erp, J. Presenting directions with a vibrotactile torso display. *Ergonomics* 48, 3 (2005), 302–313.
14. Van Erp, J.B.F., Van Veen, H.A.H.C., Jansen, C., and Dobbins, T. Waypoint navigation with a vibrotactile waist belt. *ACM Transactions on Applied Perception* 2, 2 (2005), 106–117.
15. Fiannaca, A., Morelli, T., and Folmer, E. Haptic target acquisition to enable spatial gestures in nonvisual displays. *Proceedings of Graphics Interface 2013*, (2013), 213–219.
16. Folmer, E. and Morelli, T. Spatial gestures using a tactile-proprioceptive display. *Proceedings of the Sixth International Conference on Tangible, Embedded and Embodied Interaction 1*, (2012), 139–142.
17. Fortenbaugh, F.C., Hicks, J.C., Hao, L., and Turano, K. a. Losing sight of the bigger picture: peripheral field loss compresses representations of space. *Vision research* 47, 19 (2007), 2506–20.
18. Fortenbaugh, F.C., Hicks, J.C., and Turano, K. a. The effect of peripheral visual field loss on representations of space: evidence for distortion and adaptation. *Investigative ophthalmology & visual science* 49, 6 (2008), 2765–72.
19. Fox, C.R. Low vision rehabilitation for patients with severe glaucoma. In *Principles and Practice of Ophthalmology*. 2000, 3031–3043.
20. Gawande, a a, Donovan, W.J., Ginsburg, a P., and Marmor, M.F. Photoaversion in retinitis pigmentosa. *The British journal of ophthalmology* 73, 2 (1989), 115–120.
21. Geldard, F. a. Adventures in tactile literacy. *American Psychologist* 12, 3 (1957), 115–124.
22. Grover, S., Fishman, G.A., and Brown, J. Patterns of visual field progression in patients with retinitis pigmentosa. *Ophthalmology* 105, (1998), 1069–1075.
23. Henriques, D.Y.P., Medendorp, W.P., Khan, A.Z., and Crawford, J.D. Visuomotor transformations for eye-hand coordination. *Progress in brain research* 140, (2002), 329–40.
24. Henze, N., Heuten, W., and Boll, S. Non-intrusive somatosensory navigation support for blind pedestrians. *Proc of Eurohaptics*, (2006), 459–464.
25. Herse, P. Retinitis pigmentosa: visual function and multidisciplinary management. *Clinical & experimental optometry : journal of the Australian Optometrical Association* 88, 5 (2005), 335–50.
26. Heuten, W., Henze, N., Boll, S., and Pielot, M. Tactile wayfinder: a non-visual support system for wayfinding. *Proceedings of the 5th Nordic conference on Human-computer interaction building bridges - NordiCHI '08*, ACM, New York, NY, USA (2008), 172–181.
27. Hoggan, E., Crossan, a., Brewster, S. a., and Kaaresoja, T. Audio or tactile feedback: which modality when? (2009), 2–5.
28. Hoshiyama, M., Kakigi, R., and Tamura, Y. Temporal discrimination threshold on various parts of the body. *Muscle and Nerve* 29, 2 (2004), 243–247.
29. Jones, C.M., Gray, R., Spence, C., and Tan, H.Z. Directing visual attention with spatially informative and spatially noninformative tactile cues. *Experimental brain research* 186, 4 (2008), 659–69.
30. Katz, B.F.G., Dramas, F., Parsehian, G., *et al.* NAVIG: Guidance system for the visually impaired using virtual augmented reality. *Technology and Disability* 24, 2 (2012), 163–178.

31. Katz, B.F.G., Kammoun, S., Parsehian, G., *et al.* NAVIG: Augmented reality guidance system for the visually impaired: Combining object localization, GNSS, and spatial audio. *Virtual Reality* 16, 4 (2012), 253–269.
32. Kozlowski, J.M.D. and Jalkh, A.E. An Improved Negative-Lens Field Expander for Patients With Concentric Field Constriction. *Archives of Ophthalmology* 103, 3 (1985), 326–326.
33. Krefman, R.A. Reversed telescopes on visual efficiency scores in field-restricted patients. *American journal of optometry and physiological optics* 58, 2 (1981), 159–62.
34. Kuyk, T.K., Liu, L., and Fuhr, P.S.W. Feature search in persons with severe visual impairment. *Vision Research* 45, (2005), 3224–3234.
35. Lehtinen, V., Oulasvirta, A., Salovaara, A., and Nurmi, P. Dynamic tactile guidance for visual search tasks. *Proceedings of the 25th annual ACM symposium on User interface software and technology - UIST '12*, (2012), 445.
36. Lenay, C., Gapenne, O., Hanneton, S., Marque, C., and Genouëlle, C. Sensory substitution: limits and perspectives. In *Touching for knowing: cognitive psychology of haptic manual perception*. Yvette, Hatwell, 2003, 320.
37. Lindeman, R., Yanagida, Y., Sibert, J.L., and Lavine, R. Effective vibrotactile cueing in a visual search task. *Proceedings of the ninth IFIP TC13 international conference on Human-computer interaction (INTERACT 2003)*, (2003), 89–96.
38. Lowe, J. and Drasdo, N. Patient's Responses to Retinitis Pigmentosa. *Optometry & Vision Science* 69, 3 (1992), 182–185.
39. Makous, J.C., Friedman, R.M., and Vierck, C.J. A critical band filter in touch. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 15, April (1995), 2808–2818.
40. Matscheko, M., Ferscha, A., Riener, A., and Lehner, M. Tactor placement in wrist worn wearables. *International Symposium on Wearable Computers (ISWC) 2010*, (2010), 1–8.
41. Medendorp, W.P., Beurze, S.M., Van Pelt, S., and Van Der Werf, J. Behavioral and cortical mechanisms for spatial coding and action planning. *Cortex* 44, (2008), 587–597.
42. Nakashima, R. and Shioiri, S. Why do we move our head to look at an object in our peripheral region? Lateral viewing interferes with attentive search. *PLoS ONE* 9, 3 (2014).
43. Nelson, P., Aspinall, P., and O'Brien, C. Patients' perception of visual impairment in glaucoma: a pilot study. *The British journal of ophthalmology* 83, 5 (1999), 546–552.
44. Ngo, M.K. and Spence, C. Auditory, tactile, and multisensory cues facilitate search for dynamic visual stimuli. *Attention, perception & psychophysics* 72, 6 (2010), 1654–65.
45. Oron-Gilad, T., Downs, J.L., Gilson, R.D., and Hancock, P. a. Vibrotactile Guidance Cues for Target Acquisition. *IEEE Transactions on Systems, Man and Cybernetics, Part C (Applications and Reviews)* 37, 5 (2007), 993–1004.
46. Smith, N.D., Crabb, D.P., and Garway-Heath, D.F. An exploratory study of visual search performance in glaucoma. *Ophthalmic & physiological optics : the journal of the British College of Ophthalmic Opticians (Optometrists)* 31, 3 (2011), 225–32.
47. Turano, K. a., Yu, D., Hao, L., and Hicks, J.C. Optic-flow and egocentric-direction strategies in walking: Central vs peripheral visual field. *Vision Research* 45, 25-26 (2005), 3117–3132.
48. Vargas-Martin, F. and Peli, E. Augmented-view for restricted visual field: multiple device implementations. *Optometry & Vision Science* 79, 11 (2002), 715–723.
49. Wolfe, J.M. What can 1,000,000 trials tell us about visual search? *Psychological Science* 9, 1 (1998), 33–39.
50. Woods, R.L., Giorgi, R.G., Berson, E.L., and Peli, E. Extended wearing trial of Trifield lens device for 'tunnel vision.' *Ophthalmic and Physiological Optics* 30, 3 (2010), 240–252.
51. Ye, H., Malu, M., Oh, U., and Findlater, L. Current and Future Mobile and Wearable Device Use by People With Visual Impairments. *CHI 2014*, (2014), 3123–3132.
52. Young, R.S.L. and Fishman, G.A. Sensitivity losses in a long wavelength sensitive mechanism of patients with retinitis pigmentosa. *Vision Research* 22, 1 (1982), 163–172.





# Interactions non-visuelles pour faciliter le rendu d'une séquence d'informations

Damien APPERT  
Damien.Appert@irit.fr

Christophe JOUFFRAIS  
Christophe.Jouffrais@irit.fr

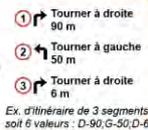
Emmanuel DUBOIS  
Emmanuel.Dubois@irit.fr

Université de Toulouse & CNRS, IRIT, Toulouse, France

Dans des situations où la perception visuelle est fortement contrainte ou déficiente, il est nécessaire de rendre perceptible l'information par une autre modalité sensorielle.

## Contexte ...

1. Percevoir des couples de valeur (ici, direction-distance)
2. Et s'en souvenir



## ... et limitations

- ❖ Multiples solutions monomodales [Zahorik 2005, Heuten 2008] explorées (son, parole, tactile, etc.)
- ❖ Mémoire à court terme limitée à  $7 \pm 2$  mnèmes, [Miller 1956]

## Notre approche : évaluation de quatre techniques non-visuelles

Explorer l'approche multimodale de rendu d'un itinéraire, i.e. une séquence de couples direction-distance

→ Ouverture vers un potentiel multi ( $7 \pm 2$  mnèmes)



angle tactile puis parole

- Modalités : tactile et parole
- Complémentaire en alternance

### 1 Technique témoin : la parole (TTS)

- 2 TTS-Lat. : parole spatialisée
- Modalités : parole et sa spatialisation
  - Complémentaire en synergie

### 3 Mélodie

- notes spatialisées
- Modalités : son et sa spatialisation
  - Complémentaire en synergie

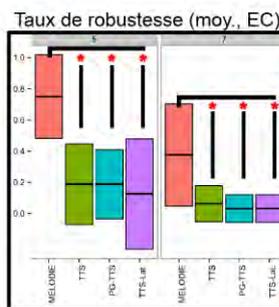
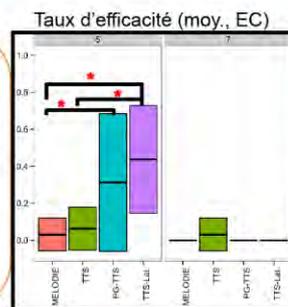


## Résultats et conclusions :

L'efficacité a été mesurée selon le nombre d'itinéraires restitués correctement lors du premier rappel. La robustesse a été mesurée selon le nombre d'itinéraires restitués de la même façon lors d'un premier rappel puis lors d'un second rappel différé.

Taux de données correctes par techniques pour des séquences de 5 et 7 couples pour 8 sujets.

- Avantage certain pour la TTS-Lat.
- Confirmation de l'intérêt de la multimodalité

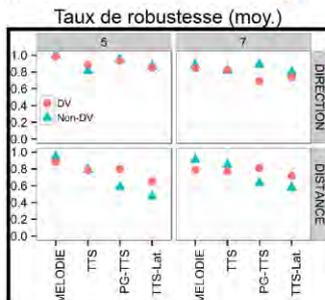
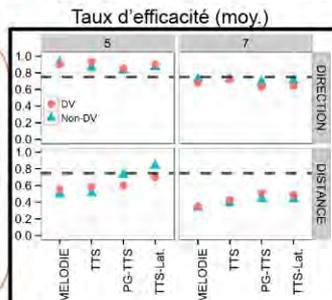


Taux de données correctes par techniques pour des séquences de 5 et 7 couples pour 8 sujets.

- Avantage significatif pour la Mélodie

Comparatif du taux de donnée correcte par techniques par type de valeur pour des séquences de 5 et 7 couples pour 8 sujets (Non-DV) et 3 sujets aveugles (DV).

- Confirmation de l'intérêt de la multimodalité



Comparatif du taux de donnée correcte par techniques par type de valeur pour des séquences de 5 et 7 couples pour 8 sujets (Non-DV) et 3 sujets aveugles (DV).

- Avantage significatif pour la Mélodie

## 9.5.3 Interactions non-visuelles pour faciliter le rendu d'une séquence d'informations

### Résumé

#### Damien Appert

Université de Toulouse & CNRS  
IRIT  
Toulouse, France  
damien.appert@irit.fr

#### Emmanuel Dubois

Université de Toulouse & CNRS  
IRIT  
Toulouse, France  
[Emmanuel.Dubois@irit.fr](mailto:Emmanuel.Dubois@irit.fr)

Cet article présente la conception et l'évaluation de techniques d'interaction multimodale non-visuelle pour la perception et la mémorisation d'une séquence d'informations. Cette tâche est représentative de multiples contextes dont notamment la restitution d'itinéraire représenté sous la forme d'une séquence

de couple direction-distance. Parmi une vingtaine de techniques que nous avons conçues, quatre ont fait l'objet d'une évaluation utilisateur. Les résultats de cette étude établissent l'intérêt de deux techniques en termes de perception, de mémorisation et de rappel des informations et une troisième s'avère plus résistante à une tâche d'interférence.

#### Mots clés choisis par les auteurs

Interactions non-visuelles, multimodalité, mémorisation

#### Mot clés de la classification ACM

H.5.2 [Information Interfaces and Presentation]: User Interfaces—auditory feedback, evaluation, haptic I/O, voice I/O

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the Owner/Author.

Copyright is held by the owner/author(s).

*IHM'15*, October 27-30, 2015, Toulouse, France

ACM 978-1-4503-3844-8/15/10.

<http://dx.doi.org/10.1145/2820619.2825013>

## INTRODUCTION

Dans des situations où la perception visuelle est fortement contrainte ou déficiente, il est nécessaire de rendre perceptible l'information par une autre modalité sensorielle. Dans ces travaux, nous nous intéressons plus particulièrement à la transmission non-visuelle de couples de valeurs de type propriété-mesure, qui peuvent avoir un rôle informationnel, de prévention ou de guidage. Cette étude préliminaire s'inspire du cas particulier de la transmission d'une feuille de route pour piéton non-voyant, générée à partir d'applications facilement disponibles et gratuites mais non dédiées aux déficients visuels (ex. Google itinéraire). La tâche générique consiste alors en la transmission sous forme non-visuelle d'une séquence de couples direction-distance (voir Figure 1).



**Figure 1: Exemple d'itinéraire composé de trois couples direction-distance : droite-90 ; gauche-50 ; droite-6. Exemple généré à partir d'applications comme Mappy ou Google Maps.**

Dans la majorité des cas, la lecture de la feuille de route s'effectue au moyen d'un lecteur d'écran. Celui-ci permet de parcourir les informations de façon séquentielle et de les restituer à l'utilisateur sous forme non-visuelle : il s'agit le plus fréquemment d'une forme audio, au moyen d'une synthèse vocale, ou d'une forme tactile, à l'aide d'une plage braille. Cependant, la mise en œuvre de ces solutions alternatives demeure confrontée aux limites de l'utilisateur en termes de mémorisation : un parcours est en effet constitué d'une séquence de couples de valeurs de plus de sept plus ou moins deux mnèmes, limite bien identifiée dans la littérature [10]. Or, les travaux de Dubois [2] ont établi l'intérêt de la multimodalité sur la mémoire de travail dans le cadre de la perception de données de différentes natures. Nous proposons donc d'explorer l'apport des techniques d'interaction multimodale non-visuelle pour optimiser la perception et la mémorisation d'une séquence de couples de données de type propriété-mesure.

A ces fins, nous avons conçu, dans cette étude préliminaire, différentes techniques d'interaction non-visuelle et multimodale avec l'objectif de favoriser la mémorisation de ce type d'information en répartissant la quantité d'information sur deux modalités utilisées de manière complémentaire. Cet article présente les techniques conçues et synthétise les résultats préliminaires obtenus suite à l'évaluation de ces différentes techniques.

## ÉTAT DE L'ART

Compenser un déficit de perception visuelle conduit à utiliser les modalités sensorielles auditive ou tactile pour transmettre les informations. Par exemple, il est possible pour un utilisateur d'exprimer une mesure avec l'angle de son coude [3] ou de percevoir une propriété comme une direction au travers d'une ceinture vibrotactile [7]. Il est aussi possible de concevoir des techniques d'interaction basées sur la sonification d'information [1,6] pour transmettre des informations de distance et de direction [15]. L'intensité ou la fréquence d'un son peuvent également être utilisées pour encoder et transmettre une mesure [9,16]. D'autres attributs du son ont également été exploités dans la littérature comme véhicules d'une information à transmettre en particulier à l'aide d'un *panning* (balance des haut-parleurs gauche et droit), d'une réverbération ou de la spatialisation du son [4,8,11]. Certaines études ont aussi cherché à combiner ces différentes techniques pour transmettre un couple de valeurs [12], notamment un couple direction-distance dans des contextes informationnels où l'information n'a pas besoin d'être mémorisée car elle doit être utilisée immédiatement, comme c'est le cas pour du guidage [14] ou une alerte [5].

Ces travaux issus de la littérature s'inscrivent donc dans un usage immédiat des informations perçues. Notre situation d'étude dépasse ce contexte : le piéton doit se souvenir de l'itinéraire à parcourir pour pouvoir, par exemple, choisir le plus optimal, limiter son stress face à un itinéraire inconnu ou transmettre l'itinéraire à une autre personne (elle aussi déficiente visuelle). Nos travaux

visent donc à explorer les apports et les limitations de la multimodalité pour la perception et la mémorisation à moyen termes d'une séquence d'informations de type propriété-mesure.

## PRÉ-ÉTUDES ET CHOIX DE DESIGN

### Conception des techniques

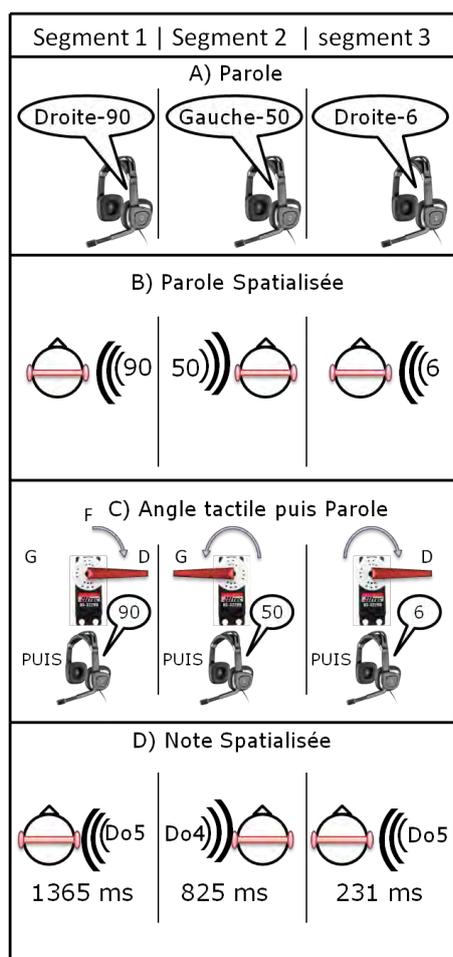
Dans le cadre d'une étude comparative, nous avons tout d'abord conçu une technique contrôle, monomodale, basée uniquement sur la synthèse vocale. Celle-ci correspond au rendu des informations grâce au lecteur d'écran et à la synthèse vocale. A partir d'une séance de brainstorming, impliquant des spécialistes en IHM, cognition spatiale et technologie de suppléances pour déficient visuel (DV), puis à la suite de plusieurs séances de réflexion aux solutions techniques, nous avons aussi conçu des techniques multimodales basées sur les modalités tactiles et sonores. Au cours de nombreuses itérations, nous avons testé informellement et avec trois sujets maximum, plusieurs techniques d'interaction multimodale, différant notamment par le rendu de l'information et les associations de modalité. Le tableau 1 synthétise les différentes options explorées et testées.

Modalité direction	Modalité distance
<b>Audio</b> Parole Sonification <b>Tactile</b> Tacton Mouvement saccadé Angle	<b>Audio</b> Parole Durée d'un son <b>Tactile</b> Angle Longueur
Présentation de l'information	
<b>Repère</b> Égocentré (G, D, F) Allocentré (O, N, E)	<b>Plage de valeur</b> Petite (<10) Grande (1 à 100)
<b>Multimodalité</b> : Fusion ; Séquentiel <b>1<sup>ère</sup> information (séquentiel)</b> : Direction-distance ; Distance-direction	

**Tableau 1: Synthèse des options envisagées lors des pré-tests. Une modalité de direction + une modalité de distance + choix de la présentation des informations = une technique multimodale.**

Les pré-tests nous ont amené à préciser les valeurs à transmettre : trois directions dans un repère égocentré (gauche, droite et en face) et la distance sur une plage de valeurs allant de 1 à 100, A l'issue de ces pré-tests, outre la technique contrôle monomodale, trois techniques multimodales ont finalement été retenues pour transmettre une séquence de couples de valeurs direction-distance. Nous décrivons ces quatre techniques ci-après.

### Description des techniques retenues



**Figure 2 : Illustration d'un même itinéraire de 3 segments, exprimé au travers des quatre techniques d'interaction ; A : TextToSpeech (TTS) ; B : TextToSpeech Latéralisée (TTS-Lat) ; C : PalmGauge-TextToSpeech (PG-TTS) ; D : Mélodie.**

La technique contrôle est la technique TextToSpeech (TTS). Elle repose sur une double synthèse de parole indiquant successivement la direction (parole) puis la distance (parole) pour chaque segment d'un parcours (Figure 2A).

La technique TextToSpeech Latéralisé (TTS-Lat) repose sur une modalité synthèse de parole et une modalité spatialisation. Seule la distance est vocalisée mais son rendu est spatialisé pour indiquer la direction : la source sonore se trouve virtuellement à gauche, en face ou à droite. TTS-Lat est une technique multimodale avec fusion d'informations : la direction et la distance sont deux informations fusionnées en une parole spatialisée (Figure 2B).

La technique PalmGauge-TextToSpeech (PG-TTS) repose sur la rotation d'un curseur placé sous la main qui se positionne à gauche (0°), en face (90°) ou à droite (180°). Puis, la distance est vocalisée (Figure 2C). PG-TTS est une technique multimodale séquentielle.

La technique Mélodie repose sur une note dont la spatialisation, la hauteur et le timbre correspondent de manière redondante à la direction et, dont la durée correspond à la distance (Figure 2D). L'enchaînement des notes correspond alors à une mélodie qui traduit une séquence de couples direction-distance. Mélodie est une technique multimodale avec fusion de données: la direction et la distance sont deux informations fusionnées en une note spatialisée. Elle a été conçue afin de permettre aux utilisateurs de « chanter la mélodie » pour s'en rappeler.

## Hypothèses de travail

Nous avons émis trois ensembles d'hypothèses basés sur les raisonnements suivants : TTS (condition contrôle) surcharge l'utilisateur d'une succession d'informations verbales et devrait être la moins utilisable. TTS-Lat, qui est la forme optimisée de TTS, devrait être plus utilisable. PG-TTS et Mélodie sont des techniques reposant sur deux principes différents permettant de mémoriser plus facilement les couples de valeurs. PG-TTS sépare les informations sur les modalités tactiles et vocales. Mélodie est conçue pour faciliter la mémorisation et le rappel en chantant. Ces deux techniques devraient donc être les plus utilisables dans ce contexte. Afin d'évaluer ces techniques selon ces ensembles d'hypothèses, nous avons conduit une expérimentation.

## PROTOCOLE EXPÉRIMENTAL

Les sujets étaient assis sur une chaise, les yeux bandés. Pour chaque sujet, l'expérimentation consistait en 4 blocs, chacun correspondant à une technique, dont l'ordre a été contrebalancé.

Un bloc commençait par quatre essais sur des itinéraires de familiarisation. L'évaluation comportait ensuite deux séries de quatre itinéraires. Les itinéraires de la première série étaient composés de cinq segments (soit 5 couples de direction et distance choisis de façon pseudo-aléatoire) ; les itinéraires de la seconde en contenaient sept. Après chaque itinéraire, le participant devait restituer oralement l'itinéraire perçu (l'expérimentateur le consignait par écrit). Cette tâche de restitution comportait deux rappels (R1 et R2) séparés par une tâche d'interférence [13] (dix décréments 3 par 3 à partir d'un nombre aléatoire) visant à établir l'apport de la technique en termes de mémorisation.

A l'issue d'un bloc, les sujets répondaient au questionnaire NASA-TLX qui permet d'évaluer la charge cognitive associée à la technique utilisée. Les deux séries étaient séparées par une courte pause (30 s). Les blocs étaient séparés par une pause plus longue (180 s).

Une fois tous les blocs passés, les sujets devaient ordonner les quatre techniques par ordre de préférence. La durée moyenne de l'expérimentation par sujet était de deux heures.

Cette expérimentation suit donc une conception intra-sujet 4x2x2 avec les techniques d'interaction, la taille (5 ou 7) et l'ordre du rappel comme facteurs. Compte tenu des ensembles d'hypothèses énoncés, nous avons analysé la satisfaction des utilisateurs vis-à-vis des techniques ainsi que la *robustesse* de la mémorisation et l'*efficacité* de chacune des techniques.

La *robustesse* a été mesurée selon le nombre d'itinéraires restitués de la même façon lors d'un premier rappel puis lors d'un second rappel différé. La robustesse était évaluée pour les directions seules, les distances seules puis pour les couples.

L'*efficacité* a été mesurée selon le nombre d'itinéraires restitués correctement lors du premier rappel. L'efficacité a été évaluée pour les directions seules, les distances seules puis pour les couples.

Onze participants ont pris part à l'évaluation : 8 sujets Non Déficiants Visuels (Non-DV, âge : 26,7±1,7) et 3 sujets Déficiants Visuels (DV, âge : 35,6±11,5). Nous relatons dans un premier temps les résultats relatifs à la population des Non-DV, puis conduisons une analyse succincte des résultats pour les DV.

## **ANALYSE DES RÉSULTATS DES SUJETS NON-DV**

Lors des comparaisons, nous avons utilisé les tests non-paramétriques de Friedman ou Wilcoxon avec un seuil de significativité à 0,05. Les graphiques représentent des moyennes (fig. 3 à 7) et des écarts-types (fig. 3 à 5).

Il ressort tout d'abord que lors de la première phase de restitution (R1) le nombre d'éléments corrects par itinéraire est en moyenne de 7,68 pour un nombre d'informations égal à 10 pour la taille 5 et à 14 pour la taille 7. La moyenne du nombre de réponses correctes du premier rappel pour chacune des conditions techniques x tailles, est de 6,88 pour la plus faible et de 8,53 pour la plus haute. C'est un résultat cohérent avec la littérature [10] (empan mnésique = 7±2).

**Satisfaction.** Les techniques préférées des Non-DV sont Mélodie et TTS-Lat (voir Figure 3 ; Friedman rank sum test  $X^2= 10,35$ ,  $df = 3$ ,  $p<0,02$ ). Aucune différence significative n'a été établie concernant le score obtenu avec le NASA-TLX (Friedman rank sum test :  $X^2= 4,95$ ,  $df = 3$ ,  $p>0,1$ ).

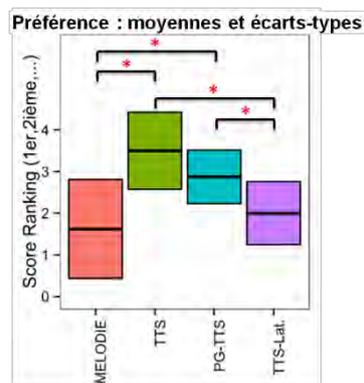


Figure 3 : Ordre moyen de préférence des techniques des sujets Non-DV; Mélodie et TTS-Lat. sont préférées à TTS et PG-TTS ; \* :  $p < 0,05$  (test post-hoc : Wilcoxon signed rank test)

**Robustesse.** La robustesse (taux) de la Mélodie était de 0,90 pour la direction seule et de 0,70 pour la distance seule quelle que soit la taille de l'itinéraire. Pour les couples, la robustesse de la Mélodie mesurée en taille 5 (resp. taille 7) atteint 0,75 (resp. 0,37) et était significativement supérieure à celle mesurée pour les autres techniques. En effet, les autres techniques ne dépassaient pas 0,18 en taille 5 et 0,07 en taille 7 (Taille 5 : Friedman rank sum test :  $X^2 = 11,7$ ,  $df = 3$ ,  $p < 0,01$  ; Taille 7 :  $X^2 = 12,9$ ,  $df = 3$ ,  $p < 0,01$  ; voir Figure 4 pour le post-hoc).

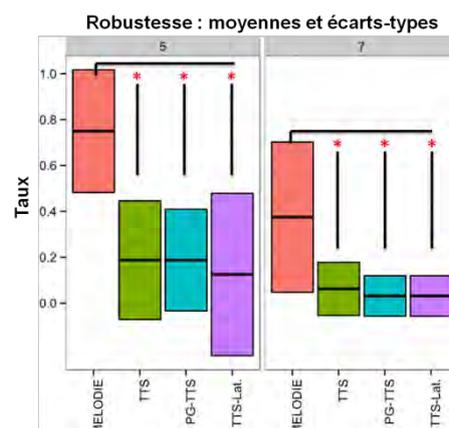


Figure 4 : Robustesse des techniques chez les non-DV (axe-X), selon deux tailles, avec le taux d'itinéraires identiques entre les deux restitutions (axe-Y). \* :  $p < 0,05$  ; test post-hoc : Wilcoxon signed rank test

**Efficacité.** Seuls les itinéraires de taille 5 font apparaître des différences notoires en termes d'efficacité (taux). Ainsi, en termes d'efficacité pour les couples, la TTS-Lat. (0,44) est significativement supérieure à celle des techniques Mélodie (0,03) et TTS (0,06). De même, l'efficacité pour les couples de la technique PG-TTS (0,31) est significativement supérieure à celle mesurée pour Mélodie (0,03) et TTS (0,06) (Taille 5 : Friedman rank sum test :  $X^2 = 12,7$ ,  $df = 3$ ,  $p < 0,01$  ; voir Figure 5 pour le post-hoc).

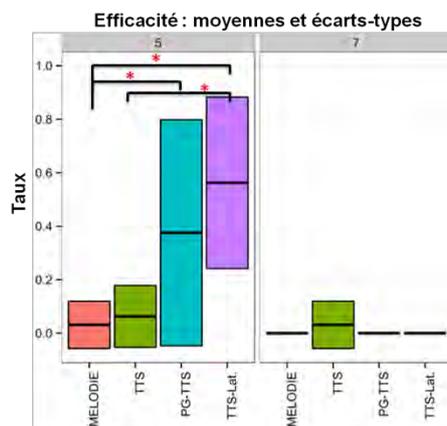


Figure 5 : Efficacité des techniques chez les non-DV (axe-X), selon deux tailles, avec le taux d'itinéraires corrects en première restitution (axes-Y) ; \* :  $p < 0,05$  ; test post-hoc : Wilcoxon signed rank test.

Concernant l'efficacité mesurée pour la direction seule, elle est excellente avec une moyenne pour toutes les techniques de 0,71, et en particulier 0,81 pour la mélodie et 0,72 pour la TTS-Lat. Notons en revanche, que l'efficacité mesurée pour la distance seule est très faible pour la Mélodie (<0,04).

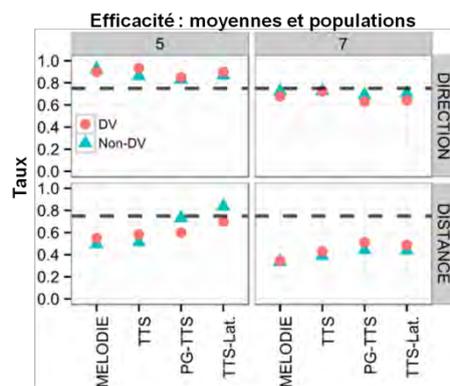
Au travers de l'analyse globale de l'efficacité, une tendance singulière semble se dégager : si pour une taille d'itinéraire et une technique données, l'efficacité pour la direction et la distance dépassent le taux de 0,75, alors l'efficacité pour les couples atteindra 0,40 (cf. Figure 6 ci-après).

## ANALYSE DES RÉSULTATS PRÉLIMINAIRES CONCERNANT LES SUJETS DV

Nous avons complété cette première analyse par une analyse spécifiquement liée aux participants DV. Compte tenu de l'échantillon très réduit, cette analyse est plus succincte que la précédente et ne peut faire appel à des tests statistiques.

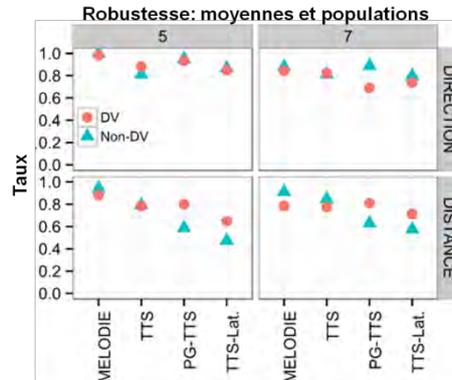
Il ressort ainsi que lors de la première phase de restitution (R1) le nombre d'éléments corrects par itinéraire est en moyenne de 7,65 (min: 7,17 ; max: 8,08). Comme pour la population des non-DV ce résultat est cohérent avec la limite bien connue de  $7 \pm 2$ .

Nous avons également observé le nombre moyen de directions et de distances correctes de la première restitution, pour chaque technique et selon les deux tailles et les deux populations. Il apparaît que les résultats des DV sont très similaires à ceux des Non-DV (points et triangles se chevauchant, voir Figure 6).



**Figure 6 : Efficacité des techniques (axe-X), selon deux tailles pour les deux populations, avec le taux de direction (partie haute) et de distance (partie basse) correctes en première restitution (axes-Y). Les lignes en pointillées indiquent le seuil de 0,75 d'efficacité.**

Il en va de même pour la robustesse (voir Figure 7), avec toutefois des différences (graphiques) plus marquées avec PG-TTS, notamment pour les directions et les distances.



**Figure 7 : Robustesse des techniques (axe-X), selon deux tailles pour les deux populations, avec le taux de direction (partie haute) et de distance (partie basse) identique entre les deux restitutions (axes-Y).**

Certaines formes de multimodalité pourraient donc faire apparaître une différence plus marquée chez les DV que chez les Non-DV.

## DISCUSSION & CONCLUSION

Nous avons conçu et évalué quatre techniques d'interaction non-visuelles dont trois étaient multimodales. Ces techniques ont été conçues afin d'évaluer leur impact sur la mémorisation et le rappel d'une séquence d'informations représentative, par exemple, d'un itinéraire pour piétons.

Nous avons étudié les résultats d'un groupe de sujets Non-DV et un groupe de sujets DV. Nos résultats ne montrent pas de différences (graphiques) entre les deux populations (Non-DV/DV) à part pour PG-TTS. Il est intéressant de noter que la limite théorique de l'empan mnésique ( $7 \pm 2$ ) aurait dû rendre impossible la mémorisation de tous les couples de valeurs pour des itinéraires courts (10 éléments), mais surtout pour des itinéraires longs (14 éléments). Néanmoins certaines techniques ont obtenu une efficacité correcte, ce qui d'après les résultats précédents de la littérature, semblent particulièrement prometteur. Par conséquent, cette étude établit que les techniques TTS-Lat. et PG-TTS semblent contribuer à optimiser la capacité de perception et de mémorisation d'une séquence de couples de valeurs.

Par ailleurs, la Mélodie a montré en taille 5, la meilleure efficacité pour les directions seules (0,81) et la pire pour les distances seules ( $<0,04$ ). Néanmoins sa robustesse pour ces mêmes itinéraires et les couples, dépasse le taux de 0,75. Il est possible qu'un mécanisme d'autosuggestion ait été utilisé par les participants pour restituer des distances (fausses) et que ce mécanisme optimise la mémorisation des valeurs. Compte tenu des retours utilisateurs et du fait qu'elle est la technique préférée des participants, nous pensons toutefois que c'est la technique elle-même qui joue ce rôle en permettant de chanter l'information. Mélodie bénéficiait aussi d'un triple encodage de la direction : la redondance dans la multimodalité pourrait donc jouer un rôle dans la perception et la mémorisation des données. Une expérimentation complémentaire est en cours pour étudier ces deux facteurs.

Notre technique de référence, la TTS ne s'est pas particulièrement démarquée. Cette technique était la moins préférée (avec PG-TTS), la moins efficace (avec Mélodie) et n'était pas robuste.

Pour compléter cette base de trois techniques d'interaction multimodale et non-visuelle pour la perception et la mémorisation de couples de valeurs, des améliorations sont requises afin d'accroître la robustesse de TTS-Lat et PG-TTS d'une part et d'autre part d'améliorer l'efficacité des trois techniques, notamment celle de la Mélodie. Cette étude a également permis de mettre en évidence un biais potentiel du protocole expérimental : il apparaît en effet nécessaire, d'adapter le mode de restitution en fonction de la technique d'interaction car la mélodie requiert une transformation (mélodie vers texte) susceptible d'affecter la restitution. De même pour une analyse plus réaliste des résultats, la notion de « distance correctement restituée » nécessite d'être affinée. Enfin, cette étude se limite à évaluer des techniques dans un cas d'utilisation stricte qui ne permet pas de généraliser sur l'efficacité des techniques *in situ*.

Nous avons également envisagé d'autres contextes basés sur l'utilisation de couples propriété-mesure, comme par exemple en géographie (région-superficie ou ville-habitant) pour lesquels de telles techniques pourraient s'avérer pertinentes et dans lesquels il conviendra d'étudier la répliquabilité des résultats observés.

## RÉFÉRENCES

- [1] Barrass, S. and Kramer, G. Using sonification. *Multimedia Systems - Special issue on audio and multimedia* 7, 1 (1999), 23–31.
- [2] Dubois, V., Gyselinck, V., Choplin, H. Multimodalité et mémoire de travail. *Environnements informatiques pour l'apprentissage humain*, (2003), 187– 198.
- [3] Enokibori, Y. and Mase, K. Human joint angle estimation with an e-textile sensor. *Proceedings of the 2014 ACM International Symposium on Wearable Computers - ISWC '14*, ACM Press (2014), 129–130.
- [4] Fontana, F. and Rocchesso, D. Auditory distance perception in an acoustic pipe. *ACM Transactions on Applied Perception (TAP)* 5, 3 (2008), 16:1–16:15.
- [5] Fujimoto, E. and Turk, M. Non-Visual Navigation Using Combined Audio Music and Haptic Cues. *ICMI '14*, ACM Press, 411–418.
- [6] Hermann, T., Hunt, A., Neuhoff, J.G., and Vickers, P. Chap 18-Sonification for Process Monitoring. In *The Sonification Handbook*. 2011, 455–491.
- [7] Heuten, W., Henze, N., Boll, S., and Pielot, M. Tactile wayfinder: a non-visual support system for wayfinding. *NordiCHI '08*, ACM Press, 172–181.
- [8] Larsen, C.H., Lauritsen, D.S., Larsen, J.J., Pilgaard, M., and Madsen, J.B. Differences in human audio localization performance between a HRTF- and a non-HRTF audio system. *AM '13*, (2013), 5:1–5:8.
- [9] Mershon, D.H. and King, L.E. Intensity and reverberation as factors in the auditory perception of egocentric distance. *Perception & Psychophysics* 18, 6 (1975), 409–415.

- [10] Miller, G.A. The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review* 63, 2 (1956), 81–97.
- [11] Mills, A.W. On the Minimum Audible Angle. *The Journal of the Acoustical Society of America* 30, 4 (1958), 237–246.
- [12] Nigay, L. and Coutaz, J. A design space for multimodal systems: concurrent processing and data fusion. *Proceedings of the INTERACT '93 and CHI '93*, ACM Press (1993), 172–178.
- [13] Peterson, L. and Peterson, M.J. Short-term retention of individual verbal items. *Journal of Experimental Psychology* 58, 3 (1959), 193–198.
- [14] Rümelin, S., Rukzio, E., and Hardy, R. NaviRadar: a novel tactile information display for pedestrian navigation. *UIST '11*, ACM Press, 293–302.
- [15] Yamano, S., Hamajo, T., Takahashi, S., and Higuchi, K. EyeSound: Single-modal Mobile Navigation Using Directionally Annotated Music. *AH'12*, 22:1–22:4.
- [16] Zahorik, P., Brungart, D.S., and Bronkhorst, A.W. Auditory distance perception in humans: A summary of past and present research. *Acta Acustica united with Acustica* 91, 3 (2005), 409–420.



# Influence de la présentation des mots pour des lecteurs dyslexiques

Damien Appert et Philippe Truillet  
 Université de Toulouse & CNRS, IRIT, Toulouse,  
 France  
[damien.appert@irit.fr](mailto:damien.appert@irit.fr) - [philippe.truillet@irit.fr](mailto:philippe.truillet@irit.fr)



## Problématique

- ❖ Quels indices visuels sont les plus intéressants ?
- ❖ Le processus de lecture peut-il être amélioré avec un code couleur ?

## Buts

- ❖ Améliorer les performances de lecture – mot bien lus
- ❖ Faciliter la lecture – nombre de fixation

## Nos idées

- ❖ Représentation colorée de phonème dans le langage écrit
- ❖ L'adapter dans le contexte de la dyslexie

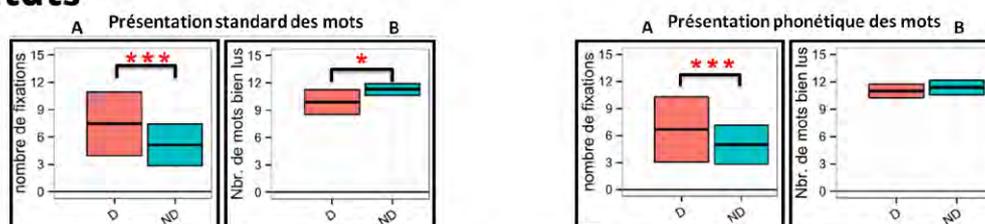
## Méthodologie

<p>+ mentonnière + microphone</p> <p>système eye-tracker</p>	<p>36 mots répartis :</p> <ul style="list-style-type: none"> <li>➤ 50% : 1<sup>ère</sup> lettre d</li> <li>➤ 50% : 1<sup>ère</sup> lettre t</li> <li>➤ 50% : 3 syllabes</li> <li>➤ 50% : 4 syllabes</li> </ul>	<p>document normal</p> <p>do   cu   ment syl   la   bi   que</p> <p>document phonétique</p>
	<p>❖ 1 plateforme expérimentale    ❖ 36 mots répartis deux critères    ❖ 3 présentations des mots</p>	

## Protocole expérimental

- ❖ 12 sujets non dyslexiques
- ❖ 8 sujets dyslexiques
- ❖ 1 heure par sujet
- ❖ Chaque sujet à lu à haute voix les 36 mots, répartis en 3 séries de présentations différentes de 12 mots. Une série commençait par une phase de familiarisation.

## Résultats



La présentation phonétique semble a priori, réduire le nombre de fixations et permettre d'approcher des performances de lecture des normo-lecteurs.

## Conclusion

- I. Le type de présentation « phonétique » a un impact à la fois sur le nombre de fixations des sujets dyslexiques et sur la facilité de lecture des mots.
- II. Ces premiers résultats peuvent guider l'analyse, la conception et l'évaluation d'interface de lecture pour les sujets dyslexiques.
- III. Nous sommes en train d'intégrer ce travail sur l'interface Ysilex afin de tester dans des conditions plus écologiques la lecture de textes.

Références : Appert, D., Outils d'analyse de mouvements oculaires, Master Thesis, June 2011, University of Toulouse ; Gattegno, C., Teaching Reading with Words in Color: A Scientific Study of the Problems of Reading, Educational Solutions, 1967-2011

## 9.5.4 Influence de la présentation des mots pour des lecteurs dyslexiques

### Damien Appert

Université de Toulouse & CNRS  
IRIT  
Toulouse, France  
damien.appert@irit.fr

### Philippe Truillet

Université de Toulouse & CNRS  
IRIT  
Toulouse, France  
philippe.truillet@irit.fr

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the Owner/Author.

Copyright is held by the owner/author(s).

IHM'15, October 27-30, 2015, Toulouse, France

ACM 978-1-4503-3844-8/15/10.

<http://dx.doi.org/10.1145/2820619.2825014>

### Résumé

Dans cet article, nous présentons une expérimentation de mesure de l'effet de trois formes de présentation des mots sur les performances de lecture. Douze sujets non dyslexiques et huit sujets dyslexiques ont participé à cette expérience. Nous avons pu montrer qu'une forme de présentation des mots, parmi les trois, amenait de meilleurs résultats de lecture pour les sujets dyslexiques.

### Mots clés choisis par les auteurs

Conception, expérimentation, dyslexie.

### Mot clés de la classification ACM

H.5.2. Information interfaces and presentation, K4.2 Assistive technologies for persons with disabilities

## INTRODUCTION

La dyslexie touche, dans tous les pays, à peu près 10% de la population. Elle correspond à un trouble de l'apprentissage du langage écrit et peut entraîner, au-delà de ces difficultés, des problèmes sociaux ou de comportement. Des études scientifiques [4, 18, 19] rapportent notamment des difficultés avec le traitement phonologique, la dénomination rapide ou un déficit de la vision, de la mémoire de travail, de la vitesse de traitement, etc. Ces difficultés de lecture et celles liées à l'orthographe continuent de causer des problèmes, en particulier pour les enfants dans le système scolaire. Le travail présenté dans ce papier est lié à la présentation du texte. Le principal résultat qui sera montré par la suite, est que les indices visuels ont un impact significatif sur les performances de lecture pour des lecteurs dyslexiques.

La section suivante relate les travaux connexes à cette étude et nos motivations. Nous présentons ensuite notre méthodologie expérimentale et les résultats principaux de cette étude. Nous concluons enfin par quelques perspectives.

## TRAVAUX CONNEXES

Si on peut considérer que les déficits phonologiques semblent être établis, la présence et la nature de déficiences visuelles sont encore très débattues. Certains auteurs ont montré que les mouvements oculaires des sujets dyslexiques sont différents de ceux des normo-lecteurs : il semble qu'ils aient des mouvements arrière plus fréquents afin de scanner les lettres plusieurs fois [11]. De nombreux travaux ont d'ailleurs testé l'impact de la perception visuelle avec des systèmes de suivi oculaire [1, 13, 14, 15].

Des études récentes affirment que la taille et l'épaisseur des lettres ont une influence sur les résultats de lecture en termes de vitesse et d'erreurs [7, 8, 13, 15]. En conséquence, il est fréquemment recommandé une présentation et une adaptation du matériel de lecture pour les personnes dyslexiques. Des polices de caractères spécifiques pour les dyslexiques ont été conçues en ce but. Par exemple, on peut citer les travaux de Boer<sup>55</sup>, Lexia Readable<sup>56</sup> Gill Dyslexic<sup>57</sup> ou de Gonzalez<sup>58</sup>. Dans ce cadre, Rello a mesuré l'impact du type de fonte de caractères sur la performance de lecture [16]. Il a pu montrer que certains types de fontes améliorent de manière significative, la lisibilité des textes et donc des performances générales de lecture. Il est d'ailleurs à noter que la plupart des interfaces de lecture de documents pour des personnes dyslexiques permettent une personnalisation de l'affichage.

Dans un autre ordre d'idée, Gattegno a proposé, il y a plusieurs décennies une méthode dénommée « *Words in Color* » pour l'apprentissage de la lecture [6]. En résumé, cette méthode consiste à associer des phonèmes d'une langue à un code couleur. Ceci amène à une « conscience phonologique » des sons qui seront produits. Ce travail est à la base de l'étude que nous avons menée à l'aide d'un oculomètre qui nous a servi d'outil de contrôle de la lecture.

## DESIGN EXPÉRIMENTAL

Le principal objectif de notre étude préliminaire est de tester l'impact de trois présentations visuelles différentes des mots sur les mouvements des yeux pour des sujets atteints de dyslexie. Le processus de lecture peut-il être amélioré avec un code couleur ou un code sémantique ? Quels indices visuels sont les plus intéressants ?

Nous avons défini trois types de présentations des mots (voir Figure 1) :

- une présentation de référence dite « présentation standard » : les mots sont écrits en blanc (code hexadécimal de la couleur #F2F2F2) sur un fond sombre (couleur #4F4F4F) ;
- une présentation dite « syllabique » pour laquelle chaque syllabe est séparée par une barre verticale qui est utilisée comme indice visuel de séparation ;
- une présentation dite « phonétique » dans laquelle la lettre « d » est colorée en couleur cyan et la lettre « t » en orange (selon la méthode de [6]). Les phonèmes associés à ces lettres, sont proches et causent de fréquentes erreurs de décodage pour les personnes dyslexiques.

<sup>55</sup> <http://www.studiostudio.nl>

<sup>56</sup> <http://www.k-type.com/?p=520>

<sup>57</sup> <http://www.pixelscript.net/gilldyslexic>

<sup>58</sup> <http://opendyslexic.org>

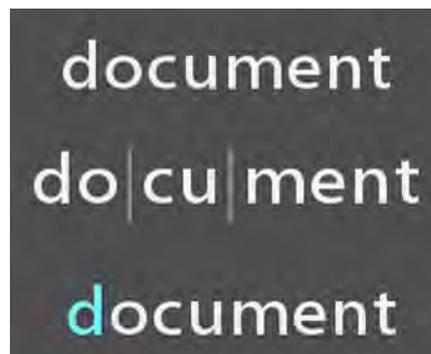


Figure 1. Illustration des trois présentations des mots : standard, syllabique et phonétique pour le mot

Nous avons utilisé la fonte « Sassoon Sans Bold » (cf. Figure 2). Cette fonte est dérivée de la fonte « Sassoon Primary », conçue avec et pour les enfants et connue pour être très lisible [17].



Figure 2. Fonte utilisée durant l'expérimentation (Sassoon Sans Bold<sup>59</sup>).

Nous avons défini un ensemble de 24 mots et de 12 non mots [10] commençant par les lettres « d » et « t ». Tous les mots contiennent 3 ou 4 syllabes. Afin de contrôler l'impact des indices visuels, les mots ont été placés à la même hauteur mais décalés vers la droite de l'écran (entre 1 et 4 centimètres aléatoirement). Ceci oblige l'œil à effectuer une saccade non automatique pour lire le mot.

## MÉTHODOLOGIE

### Conception

Les données ont été récoltées via un système eye-tracker [2] et un enregistrement sonore horodaté.

Le type de présentation et le type de sujet (Dyslexiques – D – et Non Dyslexiques – ND –) sont nos variables indépendantes et nous utilisons deux variables dépendantes : le nombre de fixations (depuis la première saccade jusqu'à la fin de l'enregistrement sonore) et le nombre d'erreurs de lecture.

*Nombre de fixations* : nous utilisons le nombre de fixations comme un indice objectif de facilité de lecture. En effet, selon Hyönä et Olson, le nombre de fixations peut être le reflet de problèmes d'identification de mots [9]. Nous corrélons cette variable avec les erreurs de lecture.

*Erreurs de lecture* : avec l'enregistrement sonore, nous avons été capables de corréler la lecture avec les mouvements oculaires.

Nous avons formulé les hypothèses suivantes :

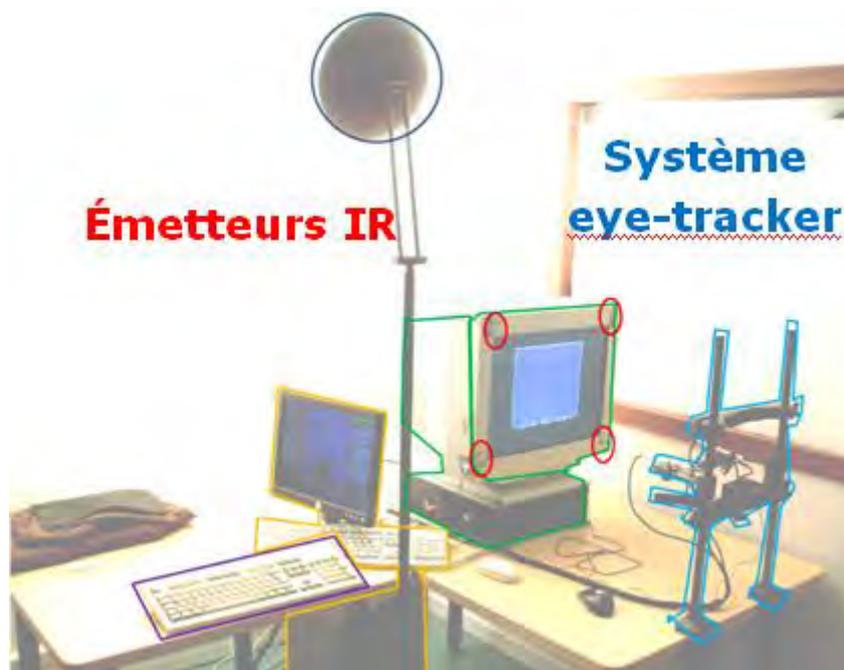
<sup>59</sup> <http://www.sassoonfont.co.uk>

*H1* : le type de présentation a un impact sur le nombre de fixations pour les sujets dyslexiques ;

*H2* : les sujets dyslexiques font moins d'erreurs avec la présentation phonétique qu'avec la présentation standard.

### Équipement

La plateforme expérimentale (voir Figure 3) est composée d'un système eye tracker (SMI Eyalink II), de caméras PTZ et d'un microphone afin de collecter la réponse des sujets. Nous avons utilisé un écran TFT 21 pouces avec une résolution de 1024x768 pixels. La précision temporelle de la mesure de l'eye-tracker est de 0,004 seconde. Les sujets étaient placés à 570 mm de l'écran, de sorte que n'importe quel mouvement oculaire d'un degré d'angle corresponde à exactement 10 mm (i.e. 26 pixels) sur l'écran.



**Figure 3. Plateforme expérimentale. Le système eye-tracker est composé d'émetteurs infrarouges (cercles rouges aux coins de l'écran) qui envoient des rayons IR dans les yeux du sujet dont le menton est posé sur la mentonnière ainsi que deux caméras infrarouges qui captent les rayons IR réfléchis par la cornée. Un logiciel d'analyse transforme ces données en données oculométriques. Un microphone était placé sous la mentonnière**

### Tâche expérimentale

Chaque sujet était assis sur une chaise, le menton posé sur une mentonnière, le front apposé à une barre prévue à cet effet. Le système eye-tracker fut adapté à la morphologie de chaque sujet. Après la lecture des consignes et la calibration de l'eye-tracker, les sujets commençaient la phase de test. Cette phase était composée de trois séries de présentations différentes. Chaque série comportait douze mots, choisis pseudo-aléatoirement parmi l'ensemble des trente-six mots afin d'équilibrer les mots et non-mots de 3 et 4 syllabes. Au final, tous les sujets ont lu l'ensemble des trente-six mots à travers trois séries de présentations différentes. L'ordre de présentation était contrebalancé entre les sujets. Les sujets devaient à la fois, lire et répéter le mot affiché à voix haute.

L'expérience s'est déroulée à l'Université Toulouse III - Paul Sabatier et a duré une heure environ par sujet. Toutes les données récoltées ont été postsynchronisées avant analyse.

## RÉSULTATS

### Participants

L'étude a été menée avec 12 sujets non dyslexiques et 8 sujets dyslexiques (8 femmes et 12 hommes). L'âge variait de 19 à 50 ans avec une moyenne de 27,3 ans. Personne n'avait de déficience mentale et tous les participants avaient une bonne vue (pas de lunettes). Les sujets devaient d'abord passer un test préliminaire (« L'alouette » [12]) afin de déterminer leur niveau de lecture. Notons que tous les sujets dyslexiques étaient reconnus comme handicapés par la MDPH<sup>60</sup>.

Nous avons utilisé le test de Kruskal-Wallis par rangs et des tests de la médiane pour l'ensemble des données en utilisant le logiciel de statistique R. Afin d'étudier les résultats significatifs, nous avons utilisé le test non-paramétrique de Friedman et le test de Wilcoxon pour des échantillons appariés. La correction de Bonferroni a été appliquée quand nécessaire. Les graphiques représentent des moyennes (fig. 4 à 6) et des écarts-types (fig. 4 à 6).

### Nombre de fixations

Nous avons calculé le nombre de fixations pour les trois présentations par catégorie de sujets. Nous pouvons constater graphiquement qu'il y a plus de fixations quand le sujet est dyslexique (voir Figure 4).

Pour les utilisateurs dyslexiques, nous avons trouvé un effet significatif entre les types de présentation (Kruskal-Wallis  $\chi^2=12,90$ ,  $df=2$ ,  $p < 0,01$ ). De plus, les présentations phonétique et standard sont significativement différentes (Pairwise Wilcoxon rank sum post-hoc  $p < 0,02$ ). En comparaison avec la forme de présentation standard, nos résultats montrent une réduction moyenne du nombre de fixations de 10,8% avec la forme phonétique (Figure 4A). Par contre, nous n'avons pas trouvé d'effets significatifs entre les présentations phonétiques et standards pour les sujets non-dyslexiques (Figure 4B ; Kruskal-Wallis  $\chi^2=14,1$ ,  $df=2$ ,  $p = n.s.$ ).

C'est pourquoi, nous nous sommes focalisés par la suite sur les résultats entre présentation standard et phonétique. La question suivante a été de savoir si ce résultat pouvait être corrélé avec de meilleures performances de lecture à haute voix.

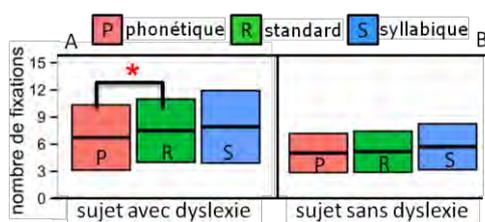
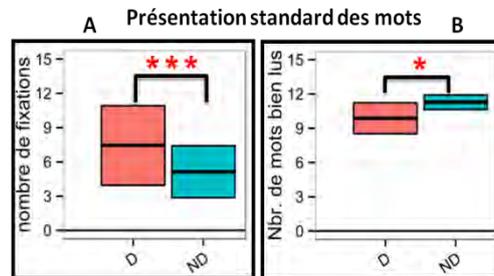


Figure 4. Nombre de fixations par type de présentation (R pour présentation standard de référence) – A pour les sujets dyslexiques et B, non dyslexiques (\*,  $p < 0,05$ ).

### Forme de présentation et mots bien lus

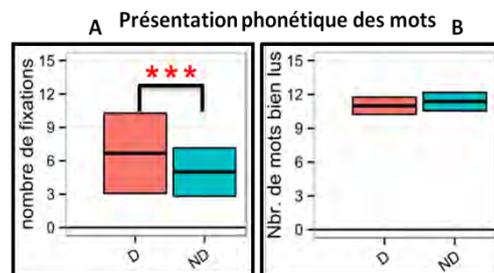
<sup>60</sup> Maison départementale des personnes handicapées

Comme nous avons enregistré la lecture de mots, nous avons pu comparer les mots bien énoncés avec le type de présentation (voir Figure 5 et 6). Pour la présentation standard, il y a un effet significatif entre le nombre de fixations entre sujets dyslexiques et non-dyslexiques (voir Figure 5A ; Kruskal-Wallis  $\chi^2=59,95$ ,  $df=1$ ,  $p < 0,001$ ) tout comme le nombre de mots correctement lus (voir Figure 5B ; Kruskal-Wallis  $\chi^2=5,29$ ,  $df=1$ ,  $p < 0,03$ ).



**Figure 5. A : nombre et de fixations et B : mots bien lus (présentation standard pour les sujets non dyslexiques et dyslexiques (\*, p<0,05 ; \*\*\*, p<0,001))**

Concernant la présentation phonétique, nous avons trouvé un effet significatif concernant le nombre de fixations entre sujets dyslexiques et non-dyslexiques (voir Figure 6A ; Kruskal-Wallis  $\chi^2=33,69$ ,  $df=1$ ,  $p < 0,001$ ). Néanmoins, il n'y a pas d'effet significatif concernant les mots correctement lus (voir Figure 6B ; Kruskal-Wallis  $\chi^2=1,13$ ,  $df=1$ ,  $p = n.s.$ ). En fait, il semble que la présentation phonétique permet de réduire les erreurs de lecture. La performance des sujets dyslexiques s'approche alors de celle des normo-lecteurs. Enfin, il est à noter que cette présentation n'affecte pas les performances des normo-lecteurs.



**Figure 6. A : nombre et de fixations et B : mots bien lus (Présentation phonétique pour les sujets non dyslexiques et dyslexiques (\*\*\*, p<0,001))**

## DISCUSSION

Comme nous venons de le voir, la présentation phonétique semble a priori, réduire le nombre de fixations et permettre d'approcher des performances de lecture des normo-lecteurs. Cependant, les résultats obtenus ne sont pas nécessairement généralisables à une lecture silencieuse de textes ou de segments de phrases.

Néanmoins, [9] a montré que dans la lecture à voix haute, les mouvements oculaires sont fortement corrélés aux processus de reconnaissance de mots.

Ces premiers résultats sur la lecture de mots et de la prononciation montrent que le type de présentation des mots a un effet sur la performance de lecture uniquement pour les sujets dyslexiques. Ceci est cohérent avec d'autres recherches et avec les recommandations de présentation de textes pour sujets dyslexiques [5, 16, 20]. Toutefois, ces études se sont

principalement focalisées sur la structure de documents ou sur les polices utilisées et finalement peu souvent sur les mots eux-mêmes et les indices visuels.

Il pourrait être intéressant de lier ces résultats aux études en neurosciences. La « saillance » des mots permet-elle de « voir » mieux et de mieux comprendre ce qui est écrit ? Est-il plus efficace de couper les mots en syllabes afin d'avoir une meilleure lecture (décodage) et une meilleure compréhension ?

## CONCLUSION ET PERSPECTIVES

La principale conclusion de ce travail préliminaire est que le type de présentation « phonétique » a un impact à la fois sur le nombre de fixations des sujets dyslexiques et sur la facilité de lecture des mots. Ces premiers résultats peuvent guider l'analyse, la conception et l'évaluation d'interface de lecture pour les sujets dyslexiques. Nous sommes en train d'intégrer ce travail sur l'interface Ysilex [3] afin de tester dans des conditions plus écologiques la lecture de textes.

## REMERCIEMENTS

Nous remercions Sylvie Athènes et Jean-Luc Vinot sans qui le projet n'aurait jamais pu se réaliser tout comme les utilisateurs qui ont participé aux tests.

## RÉFÉRENCES

- [1] Ahissar M., Lubin Y., Putter-Katz H., Banai K., Dyslexia and the failure to form a perceptual anchor, *Journal: Nature Neuroscience*, vol. 9, pp. 1558–1564, 2009
- [2] Appert, D., Outils d'analyse de mouvements oculaires, Master Thesis, June 2011, University of Toulouse
- [3] Athènes S., Raynal M., Truillet Ph., Vinot J-L. Ysilex: a Friendly reading Interface for Dyslexics, ICTA 2009, 2nd International Conference on Information and Communication Technologies and Accessibility, May 7-9 2009, Hammamet (Tunisia)
- [4] Expertise Collective de l'INSERM, (2007). Dyslexie, dysorthographe, dyscalculie: bilan des données scientifiques. Editions INSERM, Paris.
- [5] French M. M. J., Blood A., Bright N. D., Futak D.; Grohmann M. J., Hasthorpe A., Heritage J., Poland R., Reece S., Tabor J. Changing Font in Education: How the benefits vary with ability and dyslexia, *The Journal of Educational Research*, 106:301-304, 2013
- [6] Gattegno, C., Teaching Reading with Words in Color: A Scientific Study of the Problems of Reading, Educational Solutions, 1967-2011
- [7] Gregor P., Newell A., An empirical investigation of ways in which some of the problems encountered by some dyslexics may be alleviated using computer techniques, ASSETS'00, November 13-15, pp. 85-91
- [8] Hughes, L.E. & Wilkins, A.J. 2000, Typography in children's reading schemes may be suboptimal: Evidence from measures of reading rate. *Journal of Research in Reading*, 23(3), 314-324.
- [9] Hyönä J., Olson R. K., Eye Fixation Patterns AMons Dyslexic and Normal Readers: Effects of Word Length and Word Frequency, *Journal of Experimental Psychology, Learning, Memory and Cognition*, 1995, vol.21 n°6, 1430-1440,
- [10] Jacquier-Roux M., Valdois S., Zorman M., ODEDYS - Outil de dépistage des dyslexies, Académie de Grenoble, 2005, 55 pages, available at:<http://www.pedagonet.com/other/dyslexie.pdf>
- [11] Kapoula Z, Bucci MP, Vernet M, Gerard C. Normal speed and accuracy of saccade and vergence eye movements in dyslexic children. All-European Dyslexia Conference of the European Dyslexia Association, Bruges, Belgique, 22 to 24, April, 2010,
- [12] Lefavrais P., Description, définition et mesure de la dyslexie, Utilisation du test « L'Alouette ». *Revue de Psychologie Appliquée*, 15 (1), 33-44, 1965
- [13] MacKeben M, Trauzettel-Klosinski S, Reinhard J, Dürrwächter U, Adler M, Klosinski G., Eye movement control during single-word reading in dyslexics, *Journal of Vision*, 2004 May 14;4(5):388-402.
- [14] Prado, C., Dubois, M., Valdois, S. Eye movements in reading aloud and visual search in developmental dyslexia: impact of the VA span. *Vision Research*, 47, 2521-2530,
- [15] Rello L., Kanvinde G., Baeza-Yates R., Layout guidelines for web and a web service to improve accessibility for dyslexics, International Cross Disciplinary Conference on Web Accessibility (W4A), Lyon, April 2012,
- [16] Rello L, Baeza-Yates R., Good Fonts for Dyslexia, ASSETS'13, October 21-23 2013, pp. 14-22
- [17] Sassoon R., Computers and Typography, Intellect book, 1993
- [18] Shaywitz, S.E., Escobar, M.D., Shaywitz, B.A., Fletcher, J.M. & Makuch, R. (1992). Evidence that dyslexia may represent the lower tail of a normal distribution of reading ability. *New England Journal of Medicine*, 326(3), 145-150, Shovman, M., Ahissar, M.. Isolating the impact of visual perception on dyslexics' reading ability. *Vision Research*, 46, 3514-3525.

[19] Trauzettel-Klosinski S, MacKeben M, Reinhard J, Feucht A, Dürrwächter U, Klosinski G., Pictogram naming in dyslexic and normal children assessed by SLO., *Vision Res.* 2002 Mar;42(6):789-99

[20] Zikl, P, Bartosova I, Viskova, K., Hvlickova K, Kucirkova A., Navratilova J., Zetkova B., The possibilities of ICT use for compensation of difficulties with reading in pupils with dyslexia, *IETC 2014, Procedia Social and Behavioral Sciences*, 175(2015), 915-922

## 9.5.5 An MDE-based framework to support the development of Mixed Interactive Systems

Emmanuel Dubois <sup>a,\*</sup>, Christophe Bortolaso <sup>b</sup>, Damien Appert <sup>a</sup>, Guillaume Gauffre <sup>a</sup>

<sup>a</sup>University of Toulouse, IRIT – Elipse118, route de Narbonne, 31 062 Toulouse Cedex 9, France

<sup>b</sup>Queen’s University – School of Computing, Equis Lab, 141 Collingwood St., Kingston, Ontario, Canada K7L 3X6

### HIGHLIGHTS

- Five benefits justify the use of MDE for developing Mixed Interactive Systems.
- A combination of MDE tools is proposed to develop complex interactive systems.
- MDE approach contributes to the increasing popularity of Mixed Interactive Systems.
- Multiple design facets of MIS are easily supported by MDE approaches.
- Early design decisions are strongly anchored within a development approach of MIS.

### Keywords:

Mixed Interactive System , Model-Driven Engineering, Development process, Domain-specific language, Flexible model transformation

### ABSTRACT :

In the domain of Human–Computer Interaction (HCI), recent advances in sensors, communication technologies, miniaturization and computing capabilities have led to new and advanced forms of interaction. Among them, Mixed Interactive Systems (MIS), form a class of interactive systems that comprises augmented reality, tangible interfaces and ambient computing; MIS aim to take advantage of physical and digital worlds to promote a more transparent integration of interactive systems with the user’s environment. Due to the constant change of technologies and the multiplicity of these interaction forms, specific development approaches have been developed. As a result, numerous taxonomies, frameworks, API and models have emerged, each one covering a specific and limited aspect of the development of MIS. To support a coherent use of these multiple development resources and contribute to the increasing popularity of MIS, we have developed a framework based on Model-Driven Engineering. The goal is to take advantage of Model-Driven Engineering (MDE) standards, methodology and tools to support the manipulation of complementary Domain Specific Languages (DSL), to organize and link the use of different design and implementation resources, and to ensure a rationalized implementation based

on design choices. In this paper, we first summarize existing uses of MDE in HCI before focusing on five major benefits MDE can provide in a MIS development context. We then detail which MDE tools and resources support these benefits and thus form the pillars of the success of an MDE-based MIS development approach. Based on this analysis, we introduce our framework, called Guide-Me, and illustrate its use through a case study. This framework includes two design models. Model transformations are also included to link one model to another; as a result the frameworks coverage extends from the earliest design step to a software component-based prototyping platform. A toolset based on the Eclipse Modeling Framework (EMF) that supports the use of the framework is also presented. We finally assess our MDE-based development process for MIS based on the five major MDE benefits for MIS.

. \* Corresponding author. Tel.: +33 561557405.

*E-mail addresses:* Emmanuel.Dubois@irit.fr (E. Dubois), bortolas@cs.queensu.ca (C. Bortolaso), Damien.Appert@irit.fr (D. Appert), Guillaume.Gauffre@irit.fr (G. Gauffre).

[https://www.google.fr/url?sa=t&rct=j&q=&esrc=s&source=web&cd=4&cad=rja&uact=8&ved=0ahUKEwitz-Hjx7fLAhXpJZoKHQgjAVgQFgg3MAM&url=http%3A%2F%2Foatao.univ-toulouse.fr%2F12423%2F1%2FDubois\\_12423.pdf&usg=AFQjCNEmVE7CeqKdnK6\\_yFtFEBXplLbrUg](https://www.google.fr/url?sa=t&rct=j&q=&esrc=s&source=web&cd=4&cad=rja&uact=8&ved=0ahUKEwitz-Hjx7fLAhXpJZoKHQgjAVgQFgg3MAM&url=http%3A%2F%2Foatao.univ-toulouse.fr%2F12423%2F1%2FDubois_12423.pdf&usg=AFQjCNEmVE7CeqKdnK6_yFtFEBXplLbrUg)





# Conception et évaluation de techniques d'interaction non-visuelle optimisées pour de la transmission d'information

## Résumé :

Dans des situations où la perception visuelle est fortement contrainte ou déficiente, il est nécessaire de rendre perceptible l'information sous une forme non visuelle tout en prenant en compte des capacités sensorielles et mnésiques humaines.

L'objet de cette thèse est de contribuer à la conception de techniques d'interactions permettant d'optimiser la transmission non visuelle d'information. Les travaux de ma thèse ont mené à la conception, à l'implémentation et à l'évaluation de techniques d'interaction non visuelle « multiforme », en réponse à différents contextes, dont en particulier ceux de la transmission d'information(s) de type <valeur>, <position> et <itinéraire>.

Les résultats de trois études que j'ai menées, ont démontré qu'une participation « Active » des utilisateurs pendant une transmission d'information ainsi que la présence d' « aides explicites », en faciliter la perception et la mémorisation.

**Mots-clés :** interaction non-visuelle, multimodalité, estimation de valeurs, technique d'interaction active, technique d'interaction avec références, technique d'assistance.

---

---

# Design and evaluation of techniques non-visual interaction optimized for the transmission of information

## Sum up :

In situations where the visual perception is strongly constraint or deficient, it is necessary to make perceptible the information with a "not visual form" while taking into account While taking into account human sensory and mnesic capacities.

The purpose of this thesis is to contribute to the design of interaction techniques allowing to optimize the transmission not visual of the information. The works of my thesis led to the design, to the implementation and to the evaluation of interaction techniques "non-visual" and "multiform", in answer to different contexts, whom in particular those of the information transmission of type <value >, <position > and <itinerary >.

The results of three studies which I have led, demonstrated that an "active" participation of the users during a transmission of information as well as the presence of "explicit help" make easier the perception and the memorization.

**Keywords:** non-visual interaction techniques, multimodality, value estimation, technique with references, technique with references, technique of assistance.

---