

UNIVERSITE JOSEPH FOURIER (GRENOBLE I)

Techniques de l'Ingénierie Médicale et de la Complexité - Informatique, Mathématiques et Applications, Grenoble, CNRS UMR5525, La Tronche

AGIM, Age – Imagerie – Modélisation, CNRS FRE3405, La Tronche

THESE

présentée et soutenue publiquement le 23 octobre 2015

Abdessalem CHEKHCHOUKH

pour l'obtention du grade de

Docteur de l'Université Joseph Fourier (Grenoble I)

Spécialité : Ingénierie de la cognition, de l'interaction, de l'apprentissage et de la création

au titre de l'École Doctorale

Ingénierie pour la Santé, la Cognition et l'Environnement de Grenoble

**Substitution visuelle par électro-stimulation linguale
Étude des procédés affectant la qualité de perception**

Directeurs de thèse

Nicolas GLADE
Nicolas VUILLERME
Yohan PAYAN

Jury

Mme Edwige PISSALOUX
M. Christian SCHEIBER
Mme Isabelle MARC
M. Nicolas GLADE
M. Yohan PAYAN
M. Nicolas VUILLERME

Rapporteur
Rapporteur
Examineur
Co-Directeur
Co-Directeur
Directeur

Résumé. Notre perception du monde est multimodale. Les différents sens (vision, audition, tactile,...) sont coordonnés entre eux afin d'accomplir nos tâches quotidiennes. Dans le cas où un organe des sens connaît un dysfonctionnement, le système nerveux central est capable de remplacer ou de renforcer l'information manquante en faisant appel à d'autres modalités sensorielles intactes. Le cadre conceptuel de ce remplacement d'un sens par un autre est le paradigme de substitution visuelle. Cette thèse se situe dans ce contexte et, pour explorer certains aspects de ce paradigme, exploite un dispositif d'électro-stimulation linguale, le *Tongue Display Unit* (TDU), une matrice de 12x12 électrodes disposées sur la langue et destinée à restituer sous la forme de voltages variables des scènes visuelles sur la langue.

L'efficacité de ce dispositif souffre de nombreux problèmes au rang desquels le problème de *fading* (atténuation) qui correspond à une perte rapide de perception du signal d'électrostimulation par les sujets et celui de la fusion de perceptions multi-modales qui, nous le verrons, se complique avec l'usage du TDU. Dans cette étude, nous proposons d'étudier l'effet sur la qualité de perception et sur l'atténuation du signal de l'adjonction au signal image de saccades et de tremblements, processus qui nous ont été inspirés par la vision. Nous étudions également la façon dont les sujets perçoivent un signal image en cas de multi-modalité perceptive (vision + substitution visuelle sur la langue) et l'effet que produisent des contradictions d'information dans ce contexte.

Ce travail met en évidence l'importance de la présence de saccades ou de scintillement dans le signal d'électro-stimulation pour contrer les effets de *fading* : grâce à ces procédés, même lorsque les stimulations sont longues, les sujets continuent de percevoir correctement le signal image. De plus, les expériences de perception multi-modale et de contradictions nous ont permis de montrer que la perception d'informations de type image par l'intermédiaire du TDU était comparable à celle de la vision. Enfin, nous avons pu mettre en évidence un effet de fusion d'informations contradictoires entre la perception visuelle et linguale, connu dans le cadre de la fusion d'information visuo-auditives sous le nom d'effet Mc Gurck.

La compréhension de l'effet de procédés destinés à améliorer la perception dans le cadre de la substitution visuelle et celle de l'importance des effets de fusion multi-modale dégagée au cours de thèse devrait trouver application dans de nombreux domaines où ces effets sont présents ou requis, en particulier dans le domaine émergent de la vision prothétique.

Mots clés : Vision, substitution visuelle, électro-stimulation linguale, *Tongue Display Unit*, scintillement, micro-saccades, perception multi-modale, contradiction d'informations.

Abstract. Our perception of the world is by essence multimodal. Various sensory cues (vision, hearing, touch, etc.) are hence integrated to allow us the execution of daily life activities. When one sense organ is missing or impaired, in some circumstances, the central nervous system can replace or reinforce the missing sensory information by using reliable sensory information coming from an other intact sense organ. This doctoral thesis was designed to explore this so-called ‘sensory substitution’ concept through the exploitation of a visual-to-tactile sensory substitution device. The Tongue Display Unit (TDU) is made of a 12x12 electro-tactile matrix capable of rendering visual scenes on the dorsal part of the tongue.

However, the effectiveness of this device suffers from numerous drawbacks among which two main problems: (i) the problem of ‘fading’ (mitigation) which corresponds to a rapid loss of perception of electrostimulation signal by the individuals and (ii) the problem of multimodal fusion. In the present work, we propose to investigate these problems and their effects on the quality of perception, by (i) applying some eye-inspired mechanisms (micro-saccades) to the electro-tactile signal on the tongue, (ii) studying how individuals perceive a picture signal with multi-modal perceptual (visual stimulus + substitution on the tongue) and (iii) the effect produced contradictions of information in this context.

This work evidence the importance of applying saccades and sparkle mechanisms in electro-tactile signal to resolve the problem of fading especially when stimulation periods are relatively long. Furthermore, the experiments we have conducted on multi-modal perception and contradictions suggested that the perception of image information through the TDU was comparable to that of vision. Finally, we were able to identify a conflicting information fusion effect between visual perception and electrotactile tongue perception. This phenomenon is known as Mc Gurck effect (fusion of visual and auditory information).

Understanding the effect of these processes to improve the quality of perception through the visual substitution paradigm and the importance of multi-modal fusion effects during this thesis could be applied in various areas where these effects are present or required, especially in the emerging field of prosthetic vision.

Key words: Vision, visual substitution, electro-stimulation on the tongue, Tongue Display Unit, sparkle, micro-saccades, multi-modal perception, contradiction information.

Valorisation du travail doctoral

Articles :

- Chekhchoukh A. N. Vuillerme, Y. P. N. G. Effect of saccades in tongue electrotactile stimulation for vision substitution applications *Engineering in Medicine and Biology Society (EMBC), 2013 Annual International Conference of the IEEE, Japan, 2013*
- Chekhchoukh A. M. Goumidi, N. V. Y. P. N. G. Electrotactile vision substitution for 3D trajectory following *Engineering in Medicine and Biology Society (EMBC), 2013 Annual International Conference of the IEEE, 2013*
- Chekhchoukh A., Glade N., Influence of sparkle and saccades on tongue electro-stimulation-based vision substitution of 2D vectors, *Acta Biotheor.* (2012)
- Chekhchoukh A., Glade N., Vuillerme N. Vision substitution and moving objects tracking in 2 and 3 dimensions via vestibular electrostimulation of the tongue. Conference ASSISTH 2011.

Posters:

- Comment spatialiser l'information visuelle sur un dispositif d'électro-stimulation linguale ? SFBT Autrans 2011.
- Effect of saccades in tongue electrotactile stimulation for vision substitution applications *Engineering in Medicine and Biology Society (EMBC), 2013 Annual International Conference of the IEEE, 2013*
- Electrotactile vision substitution for 3D trajectory following *Engineering in Medicine and Biology Society (EMBC), 2013 Annual International Conference of the IEEE, 2013*

Présentations :

- Présentation du poster : Comment spatialiser l'information visuelle sur un dispositif d'électro-stimulation linguale ? SFBT Autrans 2011.
- Présentation journée des doctorants du laboratoire TIMC-IMAG, Substitution visuelle par électro-stimulation linguale et perception en 3D "Application DMLA", Mars 2012.

- Présentation SFBT Saint Flour Juin 2012, Substitution visuelle par électro-stimulation linguale et perception en 3D "Paradigme & Applications".
- Effect of saccades in tongue electrotactile stimulation for vision substitution applications *Engineering in Medicine and Biology Society (EMBC), 2013 Annual International Conference of the IEEE, Japan, 2013*
- Electrotactile vision substitution for 3D trajectory following *Engineering in Medicine and Biology Society (EMBC), 2013 Annual International Conference of the IEEE, Japan 2013*

Rédaction :

1. Projet de recherche soumis au Comité de Protection des Personnes CPP du CHU de Grenoble : Étude de l'Influence de Procédés Hérités de la Vision sur des Taches Psychomotrices utilisant un Dispositif d'Electrostimulation Linguale dédié à la Substitution Visuelle.

Investigateur principal : Jacques Demongeot, PU/PH

Laboratoire AGIM, UJF / CNRS FRE3405, Faculté de Médecine de Grenoble
38700, La Tronche, France

Co-Investigateurs :

Nicolas Glade, MCU, Laboratoire AGIM, UJF / CNRS FRE3405

Yohan Payan, CR1 CNRS, Laboratoire TIMC-IMAG, UJF / CNRS UMR5525

Nicolas Vuillerme, MCU, Laboratoire AGIM, UJF / CNRS FRE3405

Expérimentateur :

Abdessalem Chekhchoukh : doctorant, Laboratoires AGIM, UJF / CNRS FRE3405 & TIMC-IMAG, UJF / CNRS UMR5525.

2. Projet de recherche soumis au CPP SUD EST IV : Perception spatiale via un dispositif de stimulation électrotactile lingual : étude comportementale et imagerie d'activation cérébrale par IRMf

Investigateur principal : Pr. Christian Scheiber

Centre de Neurosciences Cognitives UMR 5229 CNRS, 57 Boulevard Pinel, 69677 Bron

Co-Investigateurs :

Jean-Claude Dreher, CR1 CNRS, Centre Neurosciences Cognitives, UMR 5229 CNRS

Laurent Jomir, Neurologue, CHG de Nîmes

Olivier Chenu, Laboratoire TIMC IMAG UMR UJF CNRS 5525

Nicolas Glade, MCU, Laboratoire AGIM, UJF / CNRS FRE3405

Yohan Payan, CR1 CNRS, Laboratoire TIMC-IMAG, UJF / CNRS UMR5525

Nicolas Vuillerme, MCU, Laboratoire AGIM, UJF / CNRS FRE3405

Expérimentateurs :

Mohamed-Hedi DAOUAS, Ingénieur en Systèmes Embarqués, ECE Paris

Abdessalam CHEKHCHOUKH, Ingénieur en thèse, UJF Grenoble

Réalisation matériel et logiciel :

- Plateforme expérimentale intégrée (voir chapitre 3).
- TDU HD : N. Glade et moi-même avons réalisé un TDU en haute résolution (24x24 taxels) en partenariat avec le bureau d'étude SMEE à Voiron et impliquant un élève ingénieur de Polytech Grenoble en 4eme année de la filière 3I.
- Expériences en IRM fonctionnelle : dans le cadre d'une collaboration avec C. Schneider, CHU Bron, Lyon, je suis allé de nombreuses fois en mission à Lyon pour mettre en place le système d'expériences de substitution visuelle sous IRM fonctionnelle. Ce travail a demandé une redéfinition complète des protocoles expérimentaux pour les rendre compatibles avec les problèmes liés à l'aimantation de l'IRM, à l'étroitesse du tube du scanner IRM, aux contraintes temporelles sur la durée des expériences sous IRM pour les patients. J'ai codé également une nouvelle plate-forme logicielle, dans un langage de script, pour contrôler à la fois le TDU et le scanner IRM. Ce travail de prototypage a été complètement réalisé.

Table des matières

Valorisation du travail doctoral.....	4
Table des matières.....	7
Glossaire & Abréviations.....	10
Introduction Générale.....	1
I - État de l'art.....	9
1 Vision, pathologies de la vision et vision prothétique.....	10
1.1. Vision normale.....	11
1.1.1. Anatomie de l'œil et mécanismes permettant la vision.....	12
1.1.2. Détail des mouvements de l'œil.....	15
1.2. Anomalies de la vision (liée à la rétine).....	19
1.2.1. DMLA.....	19
1.2.2. Rétine pigmentaire.....	21
1.3. Solutions thérapeutiques.....	22
1.3.1. Vision prothétique.....	23
1.3.1.1. Stimulation de la rétine.....	26
1.3.1.2. Stimulation du nerf optique.....	31
1.3.1.3. Stimulation du cortex.....	32
1.3.2. Thérapie cellulaire.....	34
1.4. Conclusion de l'état de l'art sur la vision et ses pathologies.....	35
2 Suppléance perceptives.....	37
2.1. Perception tactile.....	37
2.1.1. Anatomie de la peau et des muqueuses.....	38
2.1.2. La langue et sa capacité perceptives.....	40
2.1.3. Méthodes de stimulation tactile.....	43
2.2. La suppléance perceptives et ses applications.....	45
2.2.1. Le paradigme de substitution sensorielle et la suppléance perceptives.....	45
2.2.2. Applications.....	46
2.2.2.1. Tâches liées à la vision et substitution de la vision.....	49
2.2.2.2. Premiers dispositifs de substitution sensorielle électrotactiles.....	53
2.2.2.3. Interactions avec du texte et des interfaces graphiques.....	55
2.2.2.4. Applications éducatives et de divertissement.....	57
2.2.2.5. Domaine médical.....	58
2.2.2.6. Domaine militaire.....	60
2.2.2.7. Commercialisation de l'unité d'affichage sur la langue (TDU Brainport) - restitution de la vision dans le cadre de la déficience visuelle.....	61
2.3. Conclusion de l'état de l'art.....	62
II – Matériel et méthodes.....	65
3 Plate-forme expérimentale.....	66
3.1. Fonctionnalités.....	66
3.2. Diagramme de classe.....	68
3.3. Implémentation.....	71
3.3.1. Librairies Externes.....	71
3.3.2. Ressources matérielles.....	72

3.4.	Interface utilisateur (GUI).....	75
3.5.	Génération des effets affectant la qualité de perception.....	79
3.5.1.	Effet de saccades.....	79
3.5.2.	Effet de scintillement (sparkle).....	80
3.5.3.	Atténuation de l'intensité lumineuse en fonction de la profondeur.....	81
3.5.4.	Atténuation de l'intensité lumineuse en fonction du voisinage.....	82
3.5.5.	Motion blur - Rémanence.....	83
4	Protocole Expérimental.....	84
4.1.	Le processus de calibrage	84
4.2.	Familiarisation.....	88
4.3.	Protocole d'hygiène.....	88
4.3.1.	Composition du désinfectant.....	89
4.3.2.	Efficacité du désinfectant.....	89
4.3.3.	Modalités pratiques d'utilisation	89
4.4.	Conservation.....	90
4.5.	Protection du personnel et précautions d'emploi.....	90
5	Traitement des données.....	92
5.1.	Mesures - Erreur d'angle et temps de réponse.....	92
5.2.	Analyse de variance et Tests de comparaison.....	93
5.3.	Recherche de réponses gaussiennes.....	93
5.4.	Qualité des modèles.....	96
5.5.	Bootstrap.....	96
5.6.	Comparaison des modèles générateurs.....	97
III	- Expériences.....	99
6	Perception d'informations statiques.....	100
6.1.	Expérience 1 - Qualité de perception de l'orientation d'une ligne 2D.....	100
6.1.1.	Objectif	100
6.1.2.	Protocole	101
6.1.3.	Résultats.....	103
6.1.4.	Traitement statistique et Interprétation sur l'Effet des Saccades et du Scintillement.....	112
6.1.4.1.	Interprétation en première instance	11
	2	
6.1.4.2.	Interprétation en deuxième instance	11
	4	
6.2.	Expérience 2 - Perception de l'orientation d'une ligne persistante.....	117
6.2.1.	Objectif	117
6.2.2.	Protocole	117
6.2.3.	Résultats.....	120
7	Multi-modalité et Contradictions d'informations.....	125
7.1.	Contexte général.....	125
7.2.	Inversion spatiale : effet miroir.....	129
7.2.1.	Problématique.....	129
7.3.	Multi-modalité TDU et VDU.....	135
7.3.1.	Objectif.....	135

7.3.2. Protocole	136
7.3.3. Résultats.....	136
7.4. Contradictions d'informations.....	142
7.4.1. Objectif	142
7.4.2. Protocole expérimental.....	142
7.4.3. Résultats	144
7.4.3.1. Analyse globale des résultats	14
4	
7.4.3.2. Analyse des erreurs d'angles - recherche de modèles	15
0	
IV - Discussion et Perspectives.....	167
V - Bibliographie.....	182
VI - Annexes.....	188
Questionnaire.....	188

Glossaire & Abréviations

TDU : *Tongue Display Unit*, unité d'affichage sur la langue.

VDU : *Video Display Unit*, unité d'affichage vidéo écran.

HMD : *Head Mounted Display*, unité d'affichage "tête haute", supporté par un casque

DMLA : dégénérescence maculaire lié à l'age

RP : Rétine Pigmentaire

TVSS : Système de substitution Visuelle par stimulation tactile

CPP : Comité de protection des personnes

CIC : Centre d'investigation Clinique

DLL : *Dynamic Link Library*, bibliothèque logicielle dynamique

Introduction Générale

Dans le monde vivant, les façons de percevoir l'environnement varient d'un être vivant à un autre. Cette diversité est étroitement liée aux propriétés biologiques des organes qui constituent chaque système ainsi que leurs modes opératoires. En dépit de cette variété, la fonction principale est d'assurer la survie et la coexistence de l'être vivant dans son environnement. La vision chez l'être humain (et les vertébrés supérieurs de façon assez générale) fait partie de ces systèmes de perception les plus importants et les plus complexes. La synergie du système de perception visuelle avec les autres systèmes de perception (auditif, tactile, etc.) permet de nombreuses tâches comme la navigation, l'interaction, l'équilibre, etc.

L'œil est l'organe biologique responsable de l'acquisition des objets visibles dans l'environnement grâce aux photorécepteurs de la rétine. Quand un des organes qui constituent le système visuel souffre d'un dysfonctionnement, la vision est affectée d'une cécité partielle ou totale. Parfois, une personne naît avec une cécité visuelle totale (« aveugles de naissance »). Dans d'autres cas, de nombreuses maladies peuvent affecter le système visuel au niveau de différentes parties de l'organe (globe oculaire, nerf optique, cortex visuel, ...). Parmi ces affections, la dégénérescence de la rétine est une maladie dont on peut distinguer 2 cas : la **Rétine Pigmentaire (RP)** et la **Dégénérescence Maculaire Liée à l'Age (DMLA)**. Dans le cas de la DMLA une partie de la rétine qui se charge de la vision centrale (**macula**) dégénère d'une façon incrémentale. La personne ne perçoit pas l'intégralité de la scène ; seule la vision latérale persiste. Cette maladie touche les personnes "âgées" à partir de 50 ans. Le champ de vision du malade s'éteint de plus en plus alors que ce dernier se prépare à la retraite après une carrière intense. Les conséquences sont dramatiques au niveau psychologique comme au niveau physique pour le patient, mais aussi pour son entourage et sa société. Cette maladie devient un enjeu de santé publique du fait de sa prévalence internationale importante. Les statistiques sont en effet effrayantes : plus d'un million et demi de personnes sont concernées en France (2.4 % de la population française).

En neurosciences, un paradigme appelé « suppléance perceptive » propose une solution à la cécité visuelle : la substitution visuelle issue du principe de remplacement d'un sens déficient (ex: vision) par un autre sens (ex: auditif, tactile). La stimulation électro- ou vibro-

tactile sur un organe sensible, la peau notamment (ex: le bout du doigt, le dos, l'abdomen, le front, ...), peut permettre la transmission d'une information de type visuelle acheminée par le système nerveux central jusqu'au cortex. C'est le cas pour la lecture du Braille dans laquelle l'aveugle reconnaît les lettres Braille par le toucher avec le bout de ses doigts. D'autres systèmes de substitution visuelle par électro-stimulation tactile, plus complexes, ont été conçus afin de recouvrer une partie de l'information visuelle perdue. Parmi ces systèmes, Paul Bach-y-Rita, un des fondateurs de cette discipline, a proposé le système de vision par stimulation tactile ou Tactile Vision Substitution System, composé d'une caméra qui assure l'acquisition des images, d'une unité de traitement de l'information et d'une partie de stimulation tactile posée sur le dos ou l'abdomen. Une version miniaturisée du TVSS utilise le Tongue Display Unit, ou l'unité d'affichage sur la langue. Il s'agit d'une matrice d'électro-stimulation linguale composée de 12 x 12 *pixels électrotactiles* (appelé Taxel), dont les niveaux de voltage peuvent permettre la restitution des scènes visuelles 2D ou 3D sur la langue.



Figure 1. L'unité d'affichage sur la langue TDU (Tongue Display Unit. Version utilisée au cours de notre étude.

En dépit des solutions de substitution visuelle développées, la vision prothétique ou artificielle offre aussi une possibilité efficace de recouvrement de la vision. Il s'agit là d'un ensemble de technologie de prothèses implantées de manière chirurgicale sur la rétine, le nerf optique ou le cortex visuel. Ces techniques, très invasives, se basent souvent sur l'électro-stimulation à l'aide de matrices d'électrodes qui ressemblent à celle d'un TDU. Les spécialistes de la vision artificielle proposent des prothèses positionnées en différents endroits du système de perception visuelle (rétine, nerf optique, ...), différentes résolutions encore très limitées (8 x 8, 10 x 6), et formes (rectangulaire, elliptique, ...). Actuellement les recherches pour remédier à la dégénérescence de la rétine se centralisent autour de cette technologie des implants (Vision Prothétique). Ces techniques fondées sur la stimulation électrique de la rétine dégénérée ont comme contrainte que le nerf optique soit sain afin que la transmission des signaux au cortex visuel soit assurée. Au-delà de la résolution limitée de ces technologies, de leur aspect invasif et des contraintes physiologiques (nerf sain), nous verrons que d'autres problèmes subsistent, notamment liés à la façon dont l'homme perçoit les signaux de substitution.

Il existe ainsi quelques solutions aux problèmes de déficience visuelle mais la plupart ne sont pas encore matures, et d'autres sont en cours de développement. Les évolutions technologiques des différents domaines (médecine, biologie, neuroscience, ...) ne parviennent donc pas encore à solutionner les problèmes de déficience visuelle. La substitution visuelle par électro-stimulation linguale peut être considérée comme une de ces solutions se développant en parallèle et en interaction avec la vision prothétique non seulement dans le cas du DMLA mais aussi dans d'autres affectations visuelles. Il y a en effet de nombreuses similitudes entre les 2 techniques. Le TDU présente l'avantage d'une interface cerveau-machine non invasive contrairement aux implants proposés par la vision artificielle. De plus, sa résolution peut être plus facilement augmentée (densité d'électrodes plus élevée sur une surface plus importante comprenant la surface de la langue et éventuellement le palais) par rapport à celle des prothèses implantées qui sont forcément de petite dimension. L'électrostimulation linguale est cependant difficile à accepter de la part des personnes mal voyantes du fait de l'encombrement du dispositif qui le rend visible aux yeux d'autrui (la personne est ainsi plus facilement assimilée "handicapée") et qui n'est également pas très agréable à utiliser (présence de la matrice dans la cavité buccale, sortie du câble, salivation, ...).

Ce travail doctoral s'inscrit au sein d'une dynamique de recherches et de développement liés au handicap puisque réalisé dans les équipes AFIRM (Acquisition Fusion d'Information pour la Médecine) et GMCAO (Gestes Médicaux et Chirurgicaux Assistés par Ordinateur) des laboratoires AGIM et TIMC-IMAG dédiés à l'informatique et aux technologies pour la santé, dans lesquels de nombreux travaux ayant trait au handicap d'une part et à l'électro-stimulation linguale d'autre part ont été réalisés. Il contribue à l'avancement de cette voie novatrice et prometteuse, en particulier sous la forme de l'investigation des mécanismes et des procédés qui aident à l'amélioration de la qualité de la perception via ces dispositifs d'électro-stimulation (linguale ou prothétique). L'information électro-tactile sera notre centre d'intérêt. Son encodage et la manière dont l'utilisateur se comporte avec un tel dispositif de substitution perceptive ont été mesurés, analysés et discutés dans ce travail.

Un des premiers aspects de ce travail est de nature STIC (Sciences et Technologies de l'Information et de la Communication). Il consistait à se doter des moyens matériels et logiciels nécessaires pour conduire les expériences d'électro-stimulation linguale. Il s'agit du développement d'une plate-forme expérimentale intégrant divers matériels d'acquisition (outils de pointage, joystick, caméras stéréoscopiques, ...) et de stimulation (écrans, TDU ...). Notre plate-forme expérimentale permet l'implémentation et le paramétrage de protocole expérimentaux ainsi que l'automatisation des séries d'expériences associées. Un ensemble d'outils de traitement statistique des données expérimentales a également été développé.

L'axe scientifique de nos travaux de thèse est centré autour d'une problématique de recherche des procédés qui améliorent la qualité de la perception d'un signal de type image (une scène de la vie de tous les jours, un pictogramme, une forme géométrique, un vecteur ...) émis par un dispositif de substitution visuelle sur la langue (TDU). En effet, la résolution limitée de ces dispositifs (le maximum actuel est de 12 x 12 électrodes réparties sur une surface de 3x3 cm) qui n'est pas comparable à celle offerte par la rétine (des millions de photorécepteurs sur une surface d'1 cm² environ) nous oblige à chercher un codage efficace du signal et de sa spatialisation. Une restitution basique par simple dégradation de la résolution et des niveaux de couleur des scènes visuelles n'est pas suffisante pour assurer aux aveugles ou malvoyants une qualité de perception suffisante pour rendre intelligible et précis le signal et

permettre à terme la réalisation de tâches réalisées en toute autonomie (ex: tache de guidage ou de suivi de la trajectoire d'un ou plusieurs points mobiles, identification de formes ou de directions ...). De plus, s'ajoutent à ce problème de résolution d'autres sources de dégradation de la perception, en particulier l'atténuation rapide du signal perçu (fading) ou au contraire la surévaluation de l'intensité d'un signal image, liées respectivement à la physiologie des biocapteurs (cellules vibrotactiles dans le cas de la peau ou de la muqueuse linguale ; photorécepteurs de la rétine dans l'œil) dont la sensibilité au signal chute rapidement lorsque celui-ci est maintenu, et au chevauchement de plusieurs signaux voisins (ex: signaux émis par 2 taxels voisins) dus à l'étalement des signaux sur la zone stimulée (par exemple par propagation du signal électrique sur la langue, loin de l'électrode, à cause de la conduction électrique de la salive). La vision normale intègre cependant des mécanismes comme les mouvements des yeux (micro-saccades, tremblements, dérives) qui sont indispensables pour la persistance de la vision. Les micro-saccades oculaires sont nécessaires pour la perception visuelle, en particulier lorsque les yeux fixent une scène statique. On peut ainsi évaluer l'influence de tels mécanismes en les reproduisant par électro-stimulation linguale avec le TDU, sous la forme de saccades ou de scintillement du signal image, au cours d'un ensemble d'expérience de psychomotricité et de reconnaissance de formes, et mesurer le gain dans la qualité de perception en cas de persistance d'un signal-image électrique sur la langue. Les sources de dégradation sont souvent semblables (résolution, fading, sur-intensité ...) dans l'électro-stimulation linguale et la vision prothétique. Cela fait du TDU un excellent dispositif pour mettre au point, dans le cadre de la substitution visuelle (et perceptive en général), des solutions qui seront applicables à la vision prothétique.

La perception visuelle par électro-stimulation linguale est un mode de perception qui vient remplacer ou renforcer d'autres modes de perception (vision, audition, etc.). Dans le cas où une information est ressentie par plus d'une modalité de perception, cela provoque une perturbation de la personne, surtout si l'information n'est pas en concordance avec les autres modes de perception. C'est typiquement le cas de la surcharge cognitive en situation de conduite automobile, induite par la sollicitation de nombreux signaux : visuels (avec la signalisation, les autres véhicules et passants en mouvement et les données visuelles d'un GPS), auditifs (avec la radio, la voix du GPS, la voix des autres passagers), et enfin le sens véhiculé par ces signaux, le GPS pouvant indiquer une direction, la signalisation une autre et

l'expérience du conducteur une troisième. Dans ces cas-ci, le temps de décision chute (ce qui est un bon argument pour l'interdiction de la téléphonie sous toutes ses formes au volant) et les actions peuvent résulter de la fusion de plusieurs actions possibles (on ne va ni à droite, ni à gauche, mais un peu entre les deux). Pour une personne qui a l'habitude d'utiliser un mode de perception plus qu'un autre pour prendre une décision, la surcharge informative ou la contradiction entre les modalités de perception peut n'avoir *a priori* que peu d'effet. On montre cependant qu'en pratique les temps de réaction chutent et qu'il en résulte une nouvelle information (une fusion de plusieurs informations) ou un mode hybride de perception qui se situe entre les autres modes. Il se peut que parfois, une forte concentration aidant, le sujet suive un mode de perception préférentiel mais souvent une perception/action intermédiaire sera suivie. C'est le cas du phénomène mis en évidence par [MacDonald and McGurk, 1978] dans leurs expériences au cours desquelles l'information perçue est une fusion entre l'information visuelle (mouvement des lèvres d'un sujet filmé) et l'information auditive (phonèmes prononcés par un sujet). Ce phénomène affectant l'intelligibilité de ce qui est perçu est fortement susceptible d'avoir un effet important dans les applications de la substitution visuelle pour la perception augmentée (ce qui est le cas dans la substitution visuelle pour un aveugle devenu expert dans l'utilisation d'autres modes de perception, en particulier l'audition et le toucher) ou lors de l'ajout d'informations rétinienne en réalité augmentée.

Dans ce manuscrit on présentera les travaux réalisés pendant mes trois ans effectifs de thèse. Ce document est découpé en quatre parties principales chacune partagée en plusieurs chapitres :

La **première partie** intitulée « **État de l'art** » introduit le domaine de la vision, puis les travaux de substitution visuelle avec ses problèmes et les solutions existantes. On commence la première partie avec le **chapitre 1** dans lequel on donne une description de l'anatomie de l'œil et des mécanismes qui aident à bien accomplir la tâche de perception visuelle. De plus, certaines anomalies sont décrites avec leurs possibles solutions médicales et technologiques. Le **chapitre 2** consiste à présenter une revue de l'état de l'art des travaux réalisés dans le domaine de la suppléance perceptive, des études en cours et à venir menées par d'autres équipes de recherche en ce qui concerne les nouvelles technologies de restitution de la vision et de réadaptation de l'information visuelle.

La **deuxième partie** et la **troisième partie** respectivement intitulées « *Matériel et Méthodes* » et « *Expériences* » constituent le cœur de ce rapport. La deuxième partie contient la description de notre plate-forme expérimentale (**chapitre 3**) incluant les outils, ainsi que (**chapitre 4**) les méthodes et les techniques utilisées pour la mise en œuvre des expériences de psychomotricité. Avant d'entamer les expériences proprement dites de notre étude, on commence par un chapitre, placé dans la **partie** « *Matériel et Méthodes* », le **chapitre 4**, dans lequel on explique des phases communes à tous les autres expériences comme le calibrage et la phase de familiarisation des sujets avec le dispositif. **Chapitre 5**

La **troisième partie**, « *Expériences* » consacrée exclusivement aux expériences est formée des **chapitres 6** et **7** qui contiennent chacun un objectif scientifique différent. Le **chapitre 6** s'adresse à la question de l'effet des saccades et du scintillement sur la qualité de la perception électro-tactile. Le chapitre suivant (**chapitre 7**) traite, de façon générale, des questions relatives à la multi-modalité perceptive, en l'occurrence de la double perception sur le TDU et sur le VDU (moniteur) en même temps, d'informations sensorielles cohérentes ou au contraire contradictoires. Ce chapitre traite également d'un phénomène apparu au cours des expériences décrites dans le **chapitre 6**, un phénomène d'inversion spatiale (ou *effet miroir*) qui se manifeste comme une inversion par certains sujets, dans certaines conditions, des repères spatiaux droite-gauche et avant-arrière. Ce phénomène pose un problème dans l'interprétation des données.

Les données sont traitées par des analyses statistiques décrites au chapitre 5, notamment par une méthode statistique spécialement mise au point pour nos expériences et qui repose sur la recherche de mélange de distributions gaussiennes sur un support compact périodique. Cette méthode a été développée initialement pour les expériences de contradiction d'informations visuo-tactiles, mais elle s'est avérée utile dans plusieurs situations au cours de ce travail.

Ce rapport s'achève avec la **quatrième partie** « **Discussions et perspectives** » dans laquelle nous revenons avec une analyse critique sur les expériences que nous avons menées et nous comparons ce qui a été réalisé dans ce travail avec ce qui existe déjà dans ce champ de recherche. Ce travail vise en effet à explorer de nouvelles manières de percevoir une information visuelle par le biais de systèmes électro-tactiles. Ces travaux ont été menés sur une plate-forme expérimentale fondée sur l'usage d'un afficheur électro-tactile sur la langue (TDU) mais la finalité de cette étude est de transférer les connaissances et technologies mises

ici au point vers la vision prothétique qui est un domaine en pleine évolution. Nous terminons par un ensemble de perspectives dévoilant quelques domaines d'application qui visent l'amélioration de la qualité de vie des personnes aveugles et malvoyantes.

I - État de l'art

1 Vision, pathologies de la vision et vision prothétique

La perception chez l'être humain s'accomplit grâce à plusieurs sens, principalement la vision, l'audition, le sens du toucher (tactile), la proprioception. Ces modalités de perception sont synchronisées entre elles au niveau du système nerveux central (Fig. 1.1) pour accomplir les tâches quotidiennes de la personne. Un système de perception est formé de trois parties : (i) les capteurs d'information, (ii) les voies de transport de l'influx nerveux correspondant à l'information captée et (iii) une unité de traitement, le cerveau. La perception commence donc par une acquisition de l'information par l'organe via les capteurs dédiés. L'organe génère alors des pulsations nerveuses caractéristiques de l'organe et du signal capté. Ces pulsations transitent via le système nerveux périphérique, éventuellement la moelle épinière et le tronc cérébral, jusqu'au cerveau. Une des dernières étapes est le traitement de ces informations dans des aires corticales spécifiques (aires de la vision, de l'audition ...). Enfin, ces informations traitées sont fusionnées en une information intelligible par l'organisme qui peut alors les exploiter.

Par exemple, supposons que l'on demande à une personne de cliquer sur un bouton de la souris dès qu'elle voit un voyant vert sur l'écran. Lorsque ce stimulus visuel (voyant vert) s'affiche sur l'écran, l'image captée par l'œil, plus précisément la projection lumineuse du voyant sur la rétine, forme en environ 20ms des signaux nerveux générés par des neurones spécialisés de l'œil, les photorécepteurs exploitant la molécule de photo-rhodopsine pour convertir des photons en changement de potentiel membranaire, autrement dit en influx nerveux. Cet influx se propage via les axones du nerf optique jusqu'au cortex strié (aire cérébrale visuelle) en 10 ms environ. À l'intérieur des aires visuelles, un temps de traitement relativement long de 100 à 130 ms est nécessaire. Le résultat de ce traitement est envoyé vers la main via la moelle épinière en 20ms (parcours d'une distance de 1m environ). Enfin le doigt se contracte, l'activation et la contraction du muscle du doigt prenant aussi 20 ms. L'opération dure 170 à 200 ms du regard au clic.

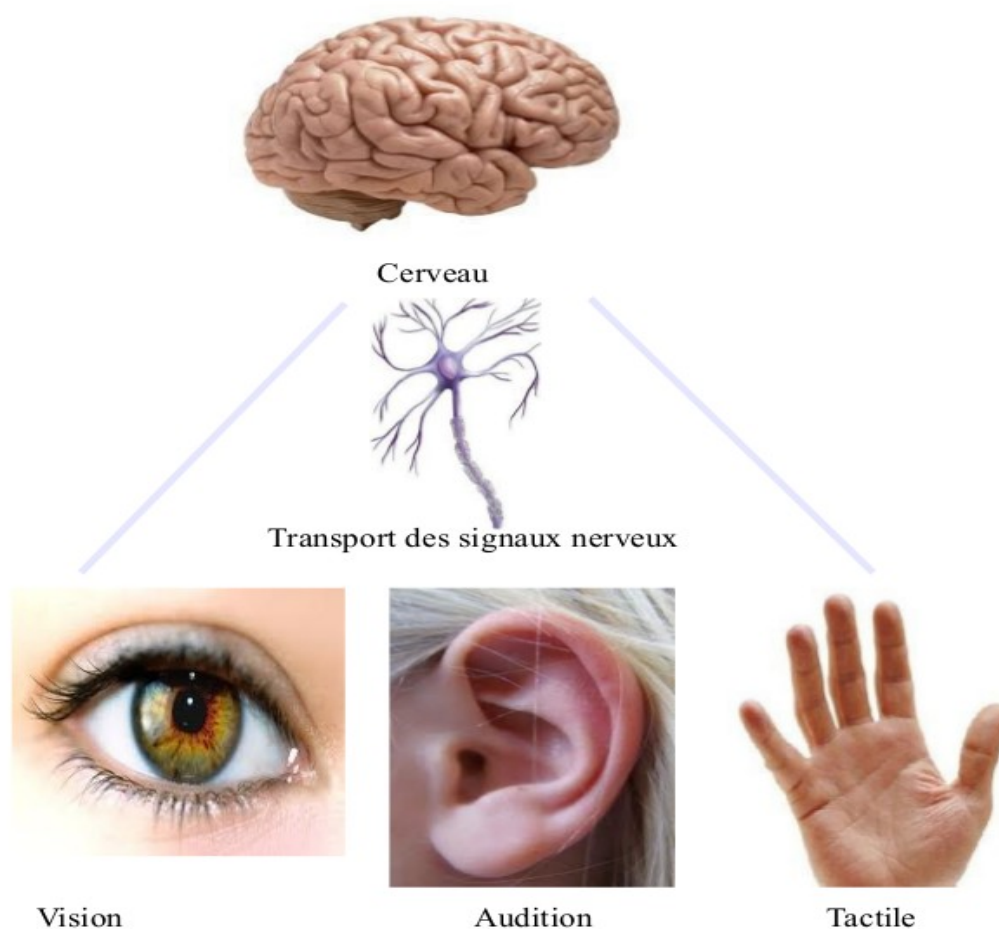


Figure 1.1. Le système de perception chez l'être humain

Dans ce chapitre, on s'intéressera à la vision dont on décrira le fonctionnement. On présentera certaines pathologies de la vision en rapport avec ce travail puis plusieurs technologies visant à rétablir la vision par des thérapies (chirurgicales ou cellulaires) sur l'organe de la vision (œil+ système nerveux associé).

1.1. Vision normale

L'œil est l'organe biologique qui s'occupe de la fonction de l'acquisition visuelle des stimulations lumineuses provenant de l'environnement (réflexion de la lumière sur les objets, émission de photons par des sources lumineuses). On explique ici la chaîne de transformation de l'information optique (lumière) vers une information chimique puis électrique.

Évidemment, les propriétés optiques de l'œil sont primordiales pour la vision normale mais dans cette section on détaillera surtout la phase post-optique, c'est-à-dire après le passage de la lumière dans le cristallin. La compréhension des mécanismes de la vision est importante puisque c'est ce qu'on cherche à reproduire ou à améliorer, souvent en s'inspirant directement de ces mécanismes naturels.

1.1.1. Anatomie de l'œil et mécanismes permettant la vision

La fonction visuelle est complexe et nécessite beaucoup de mécanismes associés à un ou plusieurs éléments anatomiques. Les éléments qui composent le capteur oculaire sont inclus pour la plupart dans le globe oculaire. À ceux-ci s'ajoutent les muscles oculaires qui assurent le contrôle des mouvements de l'œil.

Les éléments intra-oculaires proximaux (cornée, iris et cristallin) servent à produire une image projetée sur la rétine de façon nette et focalisée. La cornée et le cristallin forment la partie antérieure transparente de l'œil, ils assurent la fonction d'ajustement (à laquelle l'iris participe) et la focalisation (comme un objectif d'appareil photo). Le cristallin change de forme grâce aux muscles ciliaires. Il peut se bomber pour percevoir un objet proche ou s'aplatir pour percevoir les objets éloignés.

La rétine (encadré de la figure 1.2), l'élément intra-oculaire le plus distal, est l'élément le plus complexe de l'œil. Elle transforme l'information lumineuse en signaux nerveux. Organisée en plusieurs couches de cellules nerveuses ayant diverses spécialisations, la rétine est un prolongement du cerveau [Hubel and This, 1994]. On y distingue six types de cellules :

- les photorécepteurs **Cônes** : ce type de cellules est sensible aux couleurs. Leur grande densité au centre de la rétine (*fovéa*) assure la résolution des détails de l'image.
- les photorécepteurs **Bâtonnets** : plus nombreux que les cônes mais répartis en moins grande densité sur toute la surface de la rétine, ils assurent la vision latérale et sont sensibles même aux faibles intensités lumineuses.
- les cellules **bipolaires** reçoivent les informations électriques émises par les photorécepteurs. Elles projettent leurs axones vers les cellules ganglionnaires.

- les cellules **horizontales** relient les photorécepteurs avec les cellules bipolaires par des connexions relativement longues, parallèles avec le plan de la couche rétinienne.
- les cellules **amacrines** assurent la connexion entre les cellules bipolaires et les cellules ganglionnaires.
- les cellules **ganglionnaires** forment la couche la plus antérieure de la rétine, là où les axones traversent la surface de la rétine pour former le nerf optique. Il existe des connexions directes formées d'un photorécepteur d'une cellule bipolaire et d'une cellule ganglionnaire. Les connexions indirectes passent par les cellules horizontales et/ou amacrines [Hubel and This, 1994].

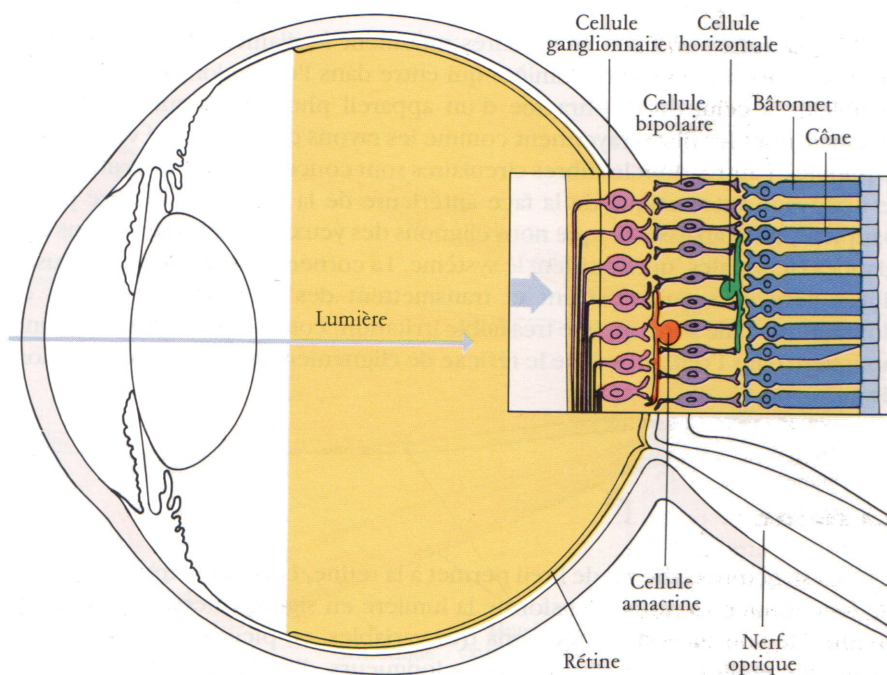


Figure 1.2. Anatomie de l'œil humain avec agrandissement d'une zone rétinienne [Hubel & This 1994]

Les photorécepteurs sont stimulés par une pluie de photons et régulés par le mécanisme de *photo-bleaching* : une molécule de pigment visuel, la rhodopsine, absorbe un photon. Via une cascade d'actions intra-cellulaires, cela provoque la dépolarisation de la membrane du photo-récepteur et une libération de neuromédiateurs au niveau des

terminaisons axonales des photorécepteurs. Lorsque la rhodopsine est stimulée, elle change d'état et perd le pouvoir d'absorber des photons avant d'être régénérée. Ce phénomène est nommé photo-bleaching et est responsable de l'atténuation (fading) de la perception de stimuli lumineux persistants.

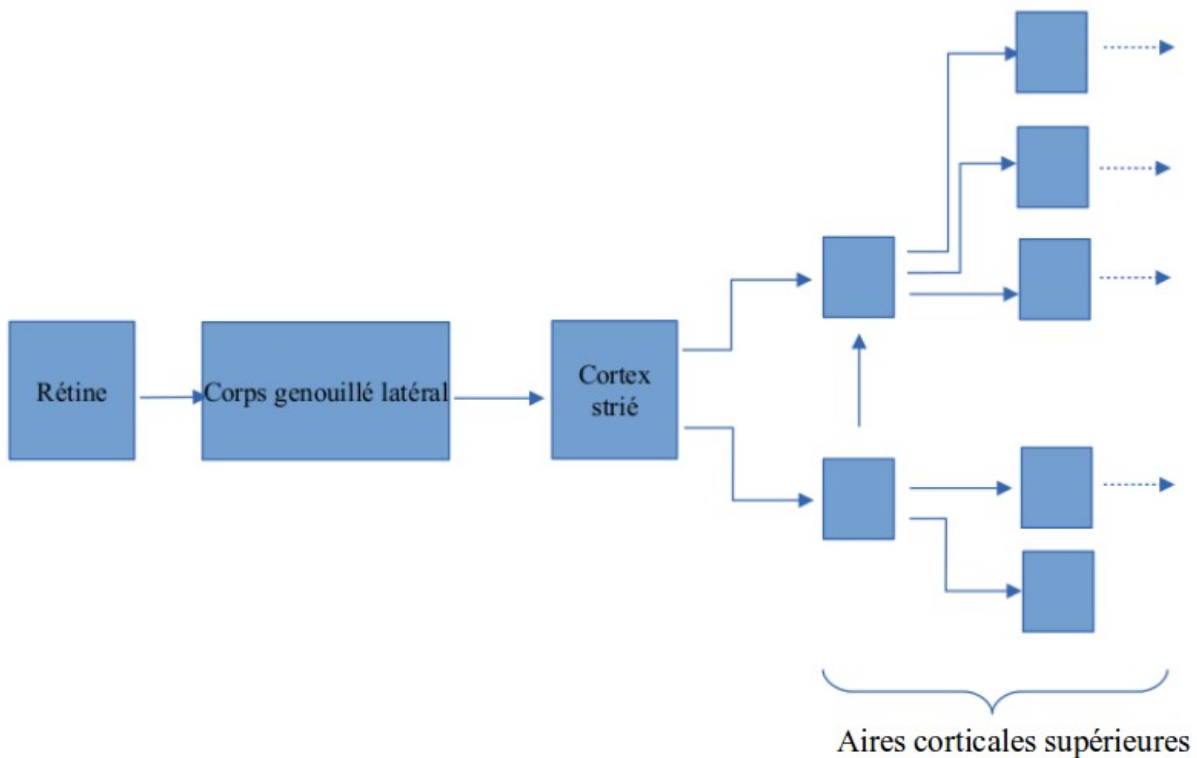


Figure 1.3. Acheminement de l'information depuis la rétine jusqu'au cortex visuel (Hubel and This, 1994).

Les photorécepteurs des neurones de type cônes et bâtonnets transforment l'énergie lumineuse en impulsion électrique qui passe par les deux couches de cellules bipolaires, horizontales et amacrines pour rejoindre enfin les cellules ganglionnaires dont les axones constituent le nerf optique. En sortie de la rétine ce sont environ un million d'axones de cellules ganglionnaires qui constituent le nerf optique. On compte cependant plus de 125 millions de photorécepteurs (bâtonnets + cônes). Cela signifie une certaine dégradation de la résolution du signal brut reçu à la surface de la rétine à cause de son intégration par la couche des cellules horizontales et amacrines.

Ces signaux sont alors transportés par le nerf optique vers le corps genouillé du thalamus puis vers le cortex visuel primaire (cortex strié situé à l'arrière du cerveau) et

secondaire afin d'y être traités. Les régions cérébrales activées diffèrent selon le type d'information et de tâche associée comme la reconnaissance d'objet, la détection, etc. Les structures post-rétiniennes du système visuel sont représentées dans la figure 1.3. Chaque structure comporte des millions de cellules qui coopèrent entre elles afin d'accomplir la tâche de vision.

Dans l'anatomie de l'œil, on doit aussi intégrer la périphérie du globe oculaire. Les mouvements de l'œil sont en effet assurés par six muscles extra-oculaires. Leur rôle est d'assurer les mouvements volontaires (ciblage, suivi ...) d'une part et les mouvements involontaires (micro-saccades, tremblements et dérives oculaires) d'autre part.

1.1.2. Détail des mouvements de l'œil

Tandis que vous lisez ces lignes de texte, vos yeux bougent d'un mot vers un autre (de gauche à droite) avec une certaine fréquence. Le titre d'une section en gras et/ou une figure attire un peu votre attention, alors vos yeux sautent vers cet endroit et peuvent ainsi explorer diverses parties de la page (bordures, coins, titres, ...) qui constituent des repères visuels. Le point d'attention retourne ensuite à l'emplacement de la dernière ligne lue pour retrouver le reste des mots que vous êtes en train de lire. De même, on sait grâce aux expériences de suivi du regard effectuées par A. Yarbus [Yarbus et al., 1967] que les yeux bougent d'une manière continue dans la scène visuelle (Fig. 1.4).

Ce sont des mouvements très rapide de nos yeux assurés par les contractions des muscles oculaires. Parfois, ces mouvements sont volontaires pour fixer le regard sur une région bien précise mais souvent ces mouvements sont involontaires. Les types de ces mouvements ainsi que leurs intérêts pour la vision sont décrits dans cette section. On se référera ici, pour la distinction des différents mouvements oculaires, aux travaux pionniers du laboratoire « Visual Neuroscience » de l'institut Barrow Neurological (Phoenix, Arizona).

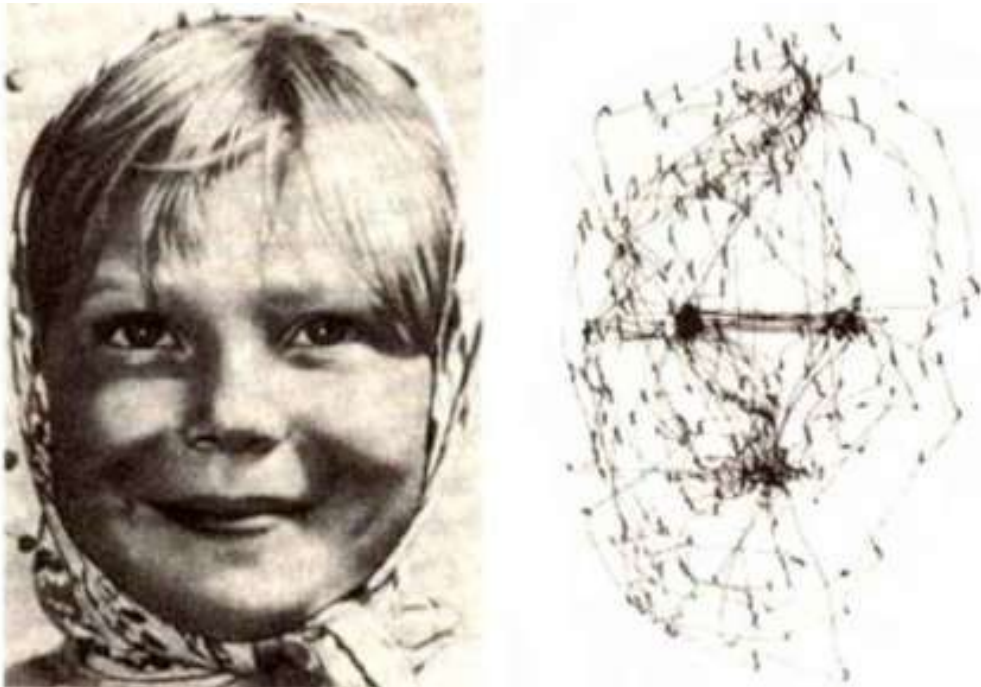


Figure 1.4. Enregistrement des mouvements des yeux (trajectoires à droite) effectué par Alfred Yarbus quand le sujet regarde ce visage féminin.

Dans la littérature on trouve plusieurs termes qui expriment les mouvements des yeux. Les types de mouvements des yeux pendant la fixation (Figs. 1.4 et 1.5) sont :

- les **Tremblements** (ou *Tremor*, parfois appelé *Nystagmus*) : sont des mouvements en forme d'ondes avec une fréquence d'environ 90Hz. Ce sont les plus petits mouvements en terme de diamètre. Leur amplitude est de l'ordre de la taille d'un photorécepteur de type cône de la fovéa. Comme il est difficile de suivre leur trajectoire par un suivi de trajectoire oculaire, la contribution des tremblements au processus de vision n'est pas encore claire.
- les **Dérives** (*Drifts*) : ce type de mouvement de grande amplitude survient simultanément avec les tremblements mais avec une fréquence plus lente. Elles se produisent entre deux mouvements de micro-saccades (Figs. 1.4 et 1.5).
- les **Microsaccades** : elles sont représentées par des mouvements rapides et d'assez faible amplitude durant la fixation volontaire d'un objet. Elles permettent de déplacer l'image rétinienne sur un grand nombre de photorécepteurs (de quelques dizaines à plusieurs centaines) sur une durée de 25 ms. Les micro-saccades oculaires dont on parle dans ce chapitre et dans la suite de ce mémoire correspondent aux petites

saccades involontaires qui se produisent au moment de la fixation [Martinez-Conde et al., 2000].

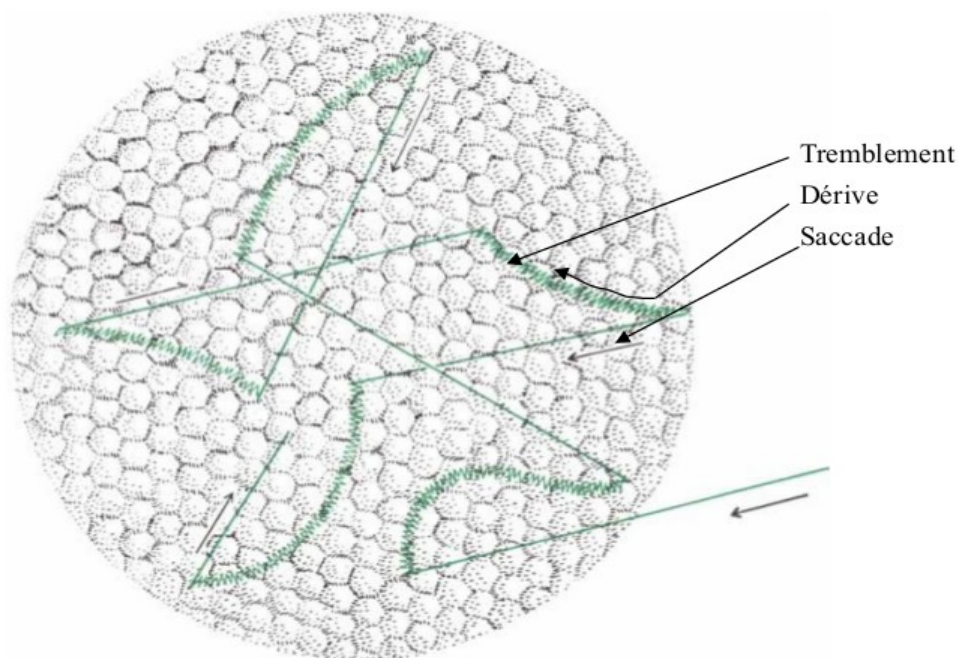


Figure 1.5. Types des mouvements de l'œil (trajectoire en vert sur une représentation de la rétine, chaque hexagone représentant un photorécepteur) pendant une phase de fixation d'un objet statique. Le temps de fixation se situe aussi autour de 10 secondes (Pritchard, 1961).

Le test de Troxler [Troxler, 1804] prouve qu'en fixant fermement (avec beaucoup de volonté) notre regard sur un seul point d'une scène visuelle on perd en quelques secondes le détail du reste de la scène. Vous pouvez réaliser vous-même cette expérience (Fig. 1.6): posez votre manuscrit de thèse sur une table, fixez le point rouge au centre du cercle bleu en vous forçant bien à ne regarder que celui-ci et à ignorer le reste ; au bout de quelques secondes le cercle commence à disparaître de votre champ de vision. On constate que même dans une scène assez simple comme celle figurant dans cet exemple, nos yeux n'arrivent pas à maintenir la perception de la scène visuelle lorsqu'on élimine (réduit tout du moins) la plupart des mouvements oculaires (dérive et saccades de fixation). En réalité l'œil est en mouvement permanent, mais ce mouvement est "fait" (*a été sélectionné au cours de l'évolution des organismes supérieurs*) pour fixer des objets.

L'apparition de ce mécanisme des micro-saccades est probablement à l'origine de nouveaux modes de prédation : une proie immobile ou très éloignée (donc ayant un mouvement apparent très lent) ne peut être perçue par un prédateur immobile que s'il dispose de ce mécanisme. Certains prédateurs (quelques reptiles) n'en disposent pas et ne peuvent

alors chasser que des proies en mouvement (et percevoir leur environnement que s'ils se déplacent eux-mêmes).

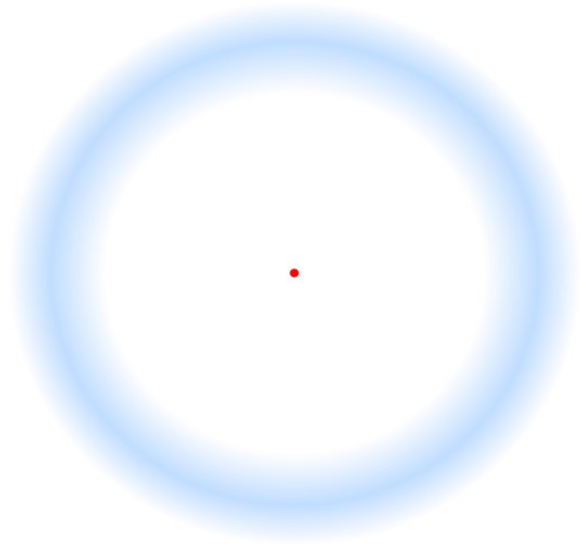


Figure 1.6. Test de Troxler : Après quelques secondes de fixation du point rouge au milieu le cercle bleu disparaîtra.

Les recherches concernant les mouvements de l'œil ont commencé depuis plus d'un siècle. Dodge a découvert dès 1907 que les yeux faisaient de petits mouvements permanents [Dodge, 1907]. Cinquante ans après, Cornsweet [Cornsweet, 1956] a commencé à étudier la fonction de ces mouvements. Au début, les rôles qui leur ont été attribués étaient limités à deux fonctions principales : (i) la prévention du phénomène de *fading* (atténuation du signal image par *bleaching* des photorécepteurs) et (ii) le contrôle du positionnement du regard pendant des longues périodes de fixation de l'œil. Par une succession d'études, le rôle des micro-saccades est devenu plus important pour les fonctions perceptives et cognitives. En particulier, on a reconnu leur importance dans l'acuité visuelle, dans la vision de petites régions spatiales, dans l'attention et dans la vision multi-stable. Dans les années 80, changement de cap : les recherches semblaient être arrivées à un consensus général que les saccades ne servaient à rien [Kowler and Steinman, 1980]. Cependant, à la fin des années 1990, les études à ce sujet se poursuivent sur les réactions neuronales qu'engendrent les mouvements de fixation des yeux dans la rétine et le cerveau [Martinez-Conde and MACKNIK, 2007]. En 1997 (seulement), le lauréat du prix Nobel David Hubel prouve dans son laboratoire le rôle des micro-saccades et la nécessité de ce mécanisme pour la vision. Voici un des éléments de son travail :

Mlle **A. I.** est un cas clinique décrit dans [Martinez-Conde et al., 2004]. Cette jeune femme n'a jamais eu de mouvements oculaires étant donné qu'elle est atteinte depuis sa naissance d'une maladie (fibrose musculaire extra-oculaire) qui affecte ses muscles extra-oculaires (ceux-ci sont paralysés). Malgré cette condition contraignante, elle a une vision normale. Elle peut lire avec une vitesse ordinaire, comme elle arrive à accomplir des tâches visio-motrices complexes comme la préparation d'un verre de thé sans aucun souci. En revanche, A. I. est obligée de suivre une stratégie afin de voir normalement: des mouvements saccadés de la tête entière. Ces mouvements incluent les deux types de mouvement, volontaires et involontaires (saccades et dérives). Cette stratégie est apparue chez A. I. pour compenser l'absence de mouvements de ses yeux. Le cas de A. I. aidera ainsi à solutionner le débat entre [Skavenski et al., 1979] et [Steinman and Collewijn, 1980] sur l'importance des mouvements de la tête et des mouvements de l'œil pour la perception visuelle.

1.2. Anomalies de la vision (liée à la rétine)

Selon [Flament, 2002], les maladies qui peuvent provoquer le dysfonctionnement de l'œil (le périphérique d'acquisition visuel) sont multiples. On peut notamment citer la rétine pigmentaire, le syndrome de Usher, l'amaurose congénitale de Leber, la maladie de Stargardt et la dégénérescence maculaire liée à l'âge. Dans cette section on présente deux de ces maladies qui affectent les cellules de la rétine. La première est la dégénérescence maculaire liée à l'âge ou **DMLA**. La deuxième est la rétine pigmentaire ou **RP**.

1.2.1. DMLA

La dégénérescence maculaire liée à l'âge est une maladie qui affecte le système visuel (Fig. 1.7), plus précisément la macula, la région de la rétine responsable de la vision centrale. La DMLA représente la première cause de malvoyance chez les personnes âgées de plus de 50 ans en France et dans les pays industrialisés. Sa définition clinique désigne des lésions précoces comportant des drusen (des dépôts de matériel amorphes dans la rétine) et/ou des migrations pigmentaires [Cohen and Desmetre, 2005].



Figure 1.7. Stades de développement de la dégénérescence maculaire liée à l'âge (DMLA) (academy for EyeCare excellence)

À partir de l'âge de 50 ans les photorécepteurs de la macula commencent à dégénérer progressivement, ce qui résulte en une cécité visuelle progressive. À part l'âge, les facteurs de risque de la DMLA sont multiples. Ils incluent le tabagisme, l'hypertension artérielle et même une couleur claire de l'iris. Cliniquement, le premier signe de DMLA est l'apparition de précurseurs, les *drusen*, qui sont dus à l'accumulation de résidus de phagocytose des photorécepteurs par les cellules de l'épithélium pigmentaire. Les drusen ont des formes et des tailles variables et ils apparaissent lors de l'analyse d'un fond de l'œil comme des amas ronds et blanchâtres [Cohen and Desmettre, 2005].

Après la phase initiale (apparition des drusens) la DMLA peut prendre une des deux formes suivantes :

- la DMLA atrophique qui est la plus prédominante (80% des cas) et mène à l'apparition d'un scotome (tache aveugle) central formé par la disparition progressive des photorécepteurs sus-jacents et la choriocapillaire sous-jacente.
- la DMLA excudative représente la deuxième forme de cette maladie. Elle est moins prévalente (20% des cas). Elle est due à l'apparition de néo-vaisseaux chorodiens qui

franchissent l'épithélium pigmentaire et se développent sous la macula. Cette néo-vascularisation choroïdienne entraîne un décollement excudatif de la rétine maculaire [Cohen and Desmettre, 2005; Parier and Soubrane, 2008].

À l'heure actuelle il n'y pas de traitement médical efficace pour la DMLA, surtout pour la forme atrophique, à part quelques solutions de rééducation en basse vision. Pour la forme excudative il existe quelques possibilités thérapeutiques (par photothérapie dynamique (ou PDT)) qui dépendent de la localisation des néo-vaisseaux chorodiens.

1.2.2. Rétine pigmentaire

La rétine pigmentaire (RP) est une maladie génétique dégénérative qui provoque une perte progressive et graduelle de la perception visuelle. L'évolution de la RP mène vers la cécité. La prévalence globale de cette maladie est de 1 pour 3500 personnes des deux sexes, quelles que soient leurs origines, âgées entre 10 et 30 ans dans la plupart des cas. Comme mentionné, la RP est génétique et due à l'altération (mutation) des gènes impliqués dans le fonctionnement et la régulation des photorécepteurs de la rétine. Dans la RP, la capacité de l'œil à absorber la lumière est affectée. Les bâtonnets sont détruits progressivement puis vient le tour de cônes (Fig. 1.8). Des scotomes se forment alors dans le champ visuel à cause des zones comprenant des photorécepteurs dégénérés.

Actuellement, il n'existe pas de traitement permettant de guérir de la RP. On parle plutôt de prévention et de précautions pour ralentir la progression de la maladie. Des solutions optiques d'amélioration de la vision sont proposées comme des lunettes grossissantes, des loupes, ainsi que des aides non optiques comme des livres à gros caractères.

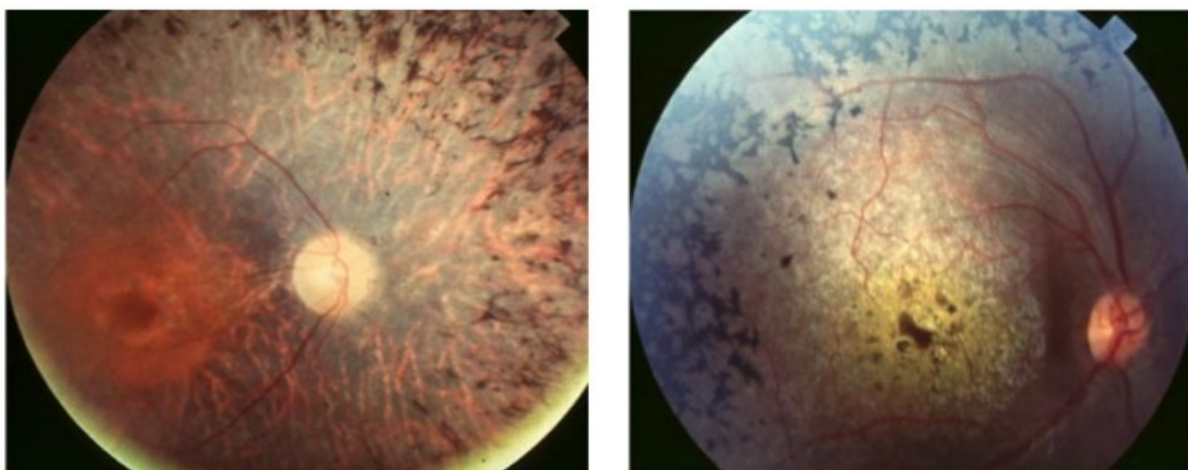


Figure 1.8. Rétinopathie pigmentaire à prédominance de bâtonnets (à gauche) et à prédominance de cônes (à droite). source : <http://www.snof.org>

L'aboutissement des recherches à ce sujet est probablement encore loin.. Les spécialistes visent à localiser les gènes responsables de la RP afin de comprendre les mécanismes impliqués dans la manifestation de la maladie et ainsi améliorer le traitement. Des travaux de recherche ont été réalisés afin de d'envisager d'autres pistes thérapeutiques comme la rétine artificielle (ce qu'on va voir dans la section 1.3.1) ou des cellules souches (section 1.3.2). Ces solutions ne sont pas encore matures pour être appliquées cliniquement et de façon systématique sur des malades de la RP.

1.3. Solutions thérapeutiques

Les dégénérescences rétiniennes par la RP ou la DMLA provoquent un dysfonctionnement de la perception visuelle et une baisse considérable de l'acuité visuelle. Selon le principe de fonctionnement de la rétine qui est la transformation de l'information lumineuse en des signaux électriques "compréhensibles" par le cortex visuel, les concepteurs d'implants rétiniens ont produit des prothèses capables d'assurer cette fonctionnalité. Ces travaux sur les prothèses visuelles progressent, mais ne sont actuellement pas suffisamment avancés pour que celles-ci soient systématiquement implantées chirurgicalement chez des patients présentant ces troubles de la vision. Précisons tout de suite qu'il s'agit d'implanter une matrice d'électrodes dans l'œil ou directement sur le cerveau (nerf optique ou cortex) ; autrement dit, il s'agit de technologies qui sont en plus extrêmement invasives.

Les recherches sur les thérapies géniques et cellulaires sont une excellente alternative aux dispositifs technologiques, surtout dans le cas des rétinopathies. Les résultats très prometteurs chez la souris condamnent peut-être d'avance tous les développements de prothèses rétinienne.

1.3.1. Vision prothétique

La majorité des concepts dédiés aux prothèses visuelles sont fondés sur l'électro-stimulation des neurones au niveau des différents emplacements critiques de la voie de perception visuelle. Le nommage de ces systèmes est étroitement lié à la région stimulée (région corticale, nerf optique ou rétine). La cécité due aux maladies comme la rétine pigmentaire ou la dégénérescence maculaire liée à l'âge peut être compensée partiellement à l'aide de l'électro-stimulation de la rétine ou du nerf optique. Une stimulation du cortex visuel peut aussi être utilisée pour restaurer une partie de la vision perdue à cause de maladies qui affectent le nerf optique (tumeur, inflammation, ...) ou le système nerveux central (les pathologies du cortex visuel primaire et secondaire). Dans de cette partie, on présentera une vue de l'ensemble des travaux actuels et futurs dans le domaine de la vision prothétique.

Le bon fonctionnement des dispositifs de vision prothétique repose sur la réponse à 2 questions : (1) la stimulation électrique d'une petite zone neuronale associée à la vision va-t-elle créer une perception lumineuse et (ii) cette dernière est-elle comparable à celle produite par la projection d'une pluie de photons sur les photorécepteurs de la rétine ? La réponse est "oui" à la première question et "non" à la deuxième. L'homme perçoit effectivement des petites zones lumineuses appelées **phosphènes** à partir de la stimulation électrique des neurones associés au système nerveux visuel. Mais ces phosphènes perçus ne sont pas représentatifs et ne sont pas comparables à ceux perçus dans une vision normale.

En ce qui concerne les dispositifs proprement dits, les spécialistes de la vision prothétique développent en permanence des techniques pour mieux rétablir l'information visuelle perdue en prenant en considération à la fois l'efficacité de la prothèse, la difficulté d'implantation et sa sûreté une fois implantée cliniquement. Les études menées pour s'assurer de l'efficacité des prothèses se font principalement par des expériences de psychophysique afin de trouver la résolution minimale et acceptable pour une vision effective.

En 1965, Bridgely GS suggère que 600 points sont suffisants pour une tâche de lecture de lettres latines individuelles imprimées normalement. Les études se poursuivent et confirment que 625 points (625 pixels dans 1 cm²) représentent le strict minimum pour une acuité visuelle de 20/30. Ces méthodes utilisent toutes un simulateur de phosphène portable. Les résultats montrent effectivement qu'avec une matrice de 25 x 25 phosphènes on peut lire efficacement quatre lettres de texte défilant à une vitesse proche de 170 mots/minute et à une vitesse de 100 mots/minute pour un texte fixe. Le même dispositif est utilisé pour des tâches de navigation à pied dans un labyrinthe (Fig. 1.9). La vitesse de marche s'améliore considérablement au cours des 3 semaines d'entraînement [Cha et al., 1992a, 1992b].

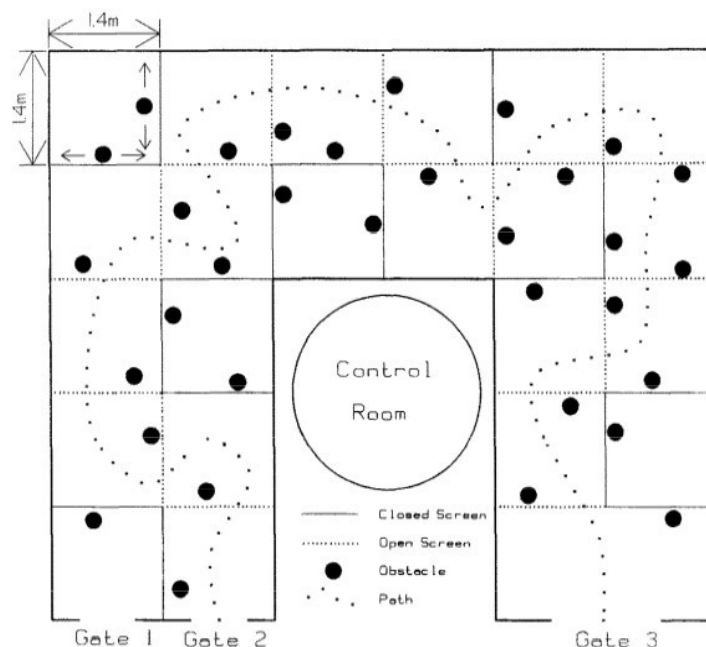


Figure 1.9. Labyrinthe utilisé au cours des expériences de Cha, Horch & Normann 1992b

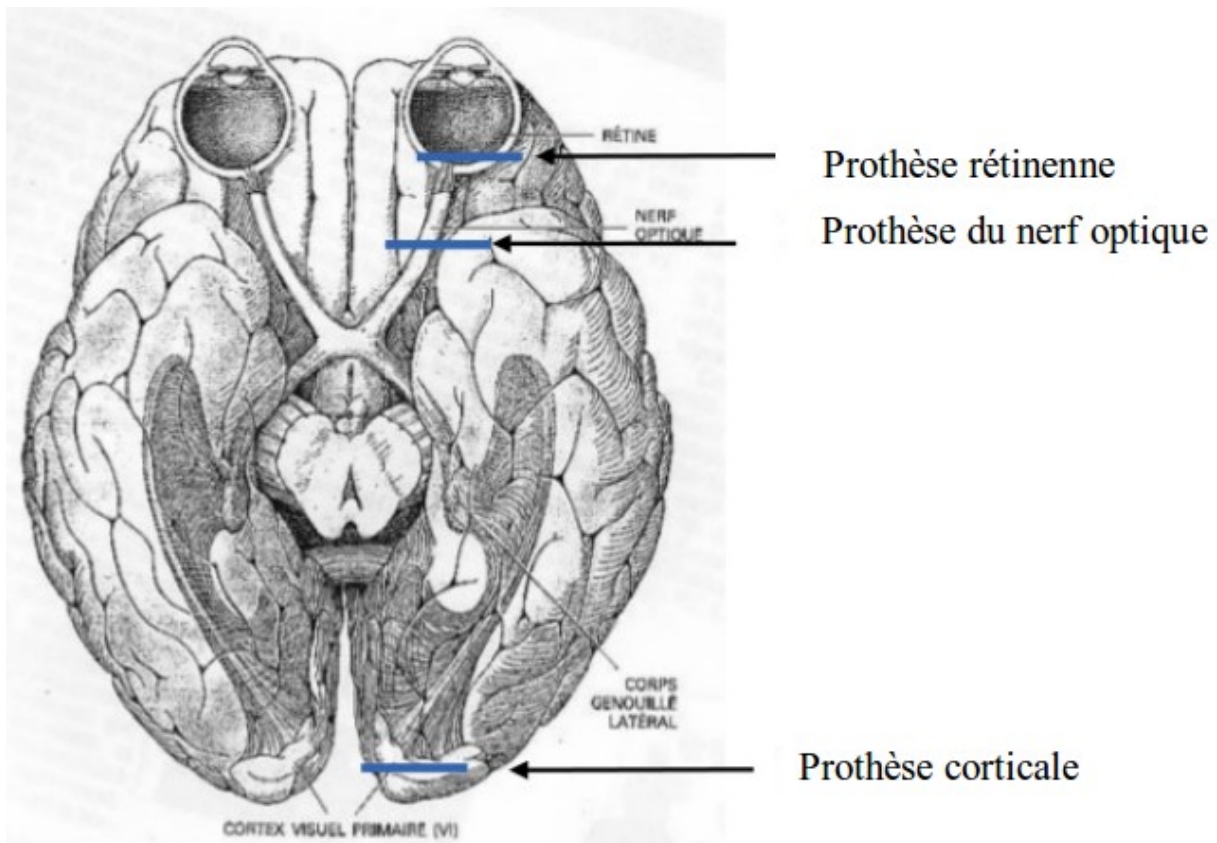


Figure 1.10. Positionnements possibles des prothèses en vision prothétique.

En pratique, les chercheurs souhaitent stimuler électriquement les neurones au lieu de stimuler la rétine avec des phosphènes. De nombreux travaux préalables ont donc été effectués en électro-physiologie. L'efficacité d'une stimulation électrique d'un tissu neuronal est influencée par 8 facteurs :

- le seuil de stimulation électrique dépend des propriétés électriques de la cellule neuronale et de la partie de la cellule qui va être stimulée (dendrite, axone ou soma) [Rattay, 1998],
- la distance entre la région stimulée et la cellule cible. Un modèle mathématique a été proposé par [BeMent and Ranck Jr, 1969].,
- la durée de stimulation,
- l'impédance du tissu et son degré d'anisotropie [Szlavik and de Bruin, 2000],
- l'état de la cellule au moment de la stimulation,
- la charge du signal,
- la polarité du signal,

- le format de l'onde.

Récemment, une équipe d'un laboratoire australien a proposé des modèles incluant l'ensemble de ces facteurs [Meffin et al., 2013, 2012].

La matière de fabrication des électrodes est aussi un problème. Des matériaux précieux ont été proposés : platine, iridium, rhodium, or, et palladium. Le platine et les alliages avec l'iridium sont les plus utilisés du fait de leur résistance à la corrosion et de leur capacité considérable en termes de charge électrique.

Nous allons maintenant voir les 3 types de prothèses visuelles (Fig. 1.10) (rétiniennes, sur le nerf optique, et corticales). Plus de détail sur la vision prothétique et les prothèses rétiniennes sont néanmoins disponibles dans [Javaheri et al., 2006; Kien et al., 2012; Margalit et al., 2002]. Ces publications donnent un bon aperçu des concepts, des différentes techniques utilisées ainsi que des travaux en cours et futurs du domaine de la vision artificielle.

1.3.1.1. Stimulation de la rétine

La majorité des études menées dans le domaine de la vision prothétique sont concentrées sur l'électro-stimulation rétinienne, avec souvent, comme objectif, le traitement technologique des maladies de dégénérescence de la rétine comme la RP et la DMLA. Quand la rétine perd ses photorécepteurs la fonction d'acquisition de la lumière est arrêtée. Cette couche de photorécepteurs dégénérés peut-être remplacée par une prothèse positionnée en différentes zones de la rétine (voir Fig. 1.11), notamment en position sous-rétinienne (*subretinal layer*) ou au contraire sur la surface interne de la rétine (position epi-rétinienne). Les avantages de cette dernière technique sont : (1) la possibilité d'utiliser la physiologie optique de l'œil ainsi que les éléments poste-rétiniens du système de perception visuel (le nerf optique et le cortex visuel) ; (2) la cavité de l'œil permet de voir la prothèse avec un simple enlèvement de la pupille ce qui facilite la consultation externe ; (3) l'intervention chirurgicale est moins complexe que celle nécessaire pour les autres types de prothèses sur le cortex ou le nerf optique. On peut néanmoins citer quelques inconvénients des prothèses rétiniennes : (1) la stimulation n'est pas suffisamment focalisée pour activer un seul axone des cellules

ganglionnaires à la fois (une électrode va stimuler un très grand nombre de neurones de la même manière) ; (2) la difficulté de fixation de la prothèse surtout pour le type épi-rétinien ; (3) les cellules bipolaires et ganglionnaires encodent bien plus d'informations qu'une simple stimulation pouvant provenir d'une électrode, notamment des propriétés de la lumière comme sa couleur (fréquence dans le spectre lumineux), intensité, ... ; (4) cette technique ne peut être utilisée que dans le cas de la rétinopathie (en effet, si l'affection touche le nerf optique ou le cortex visuel, stimuler la rétine ne résoudra rien).

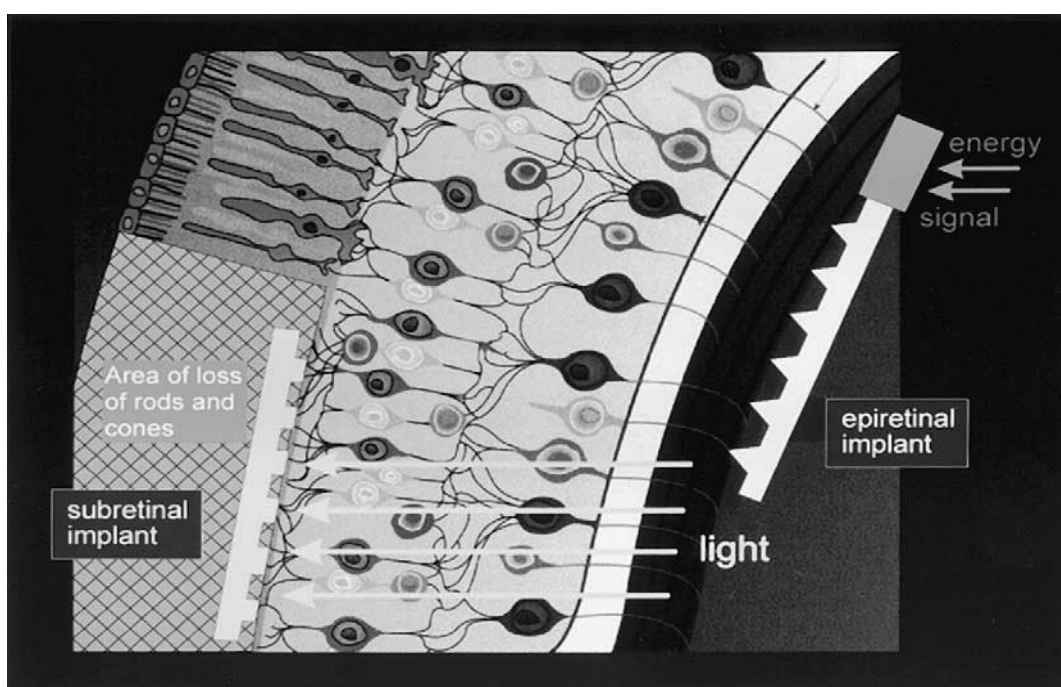


Figure 1.11. Positionnement des implants rétiniens sous-rétiniennes & épi rétiniennes [Margalit et al 2002]

Le système développé par [Peyman et al., 1998] fonctionne avec environ 3500 micro-photodiodes capables de produire du courant électrique à partir de la lumière ambiante. Cette prothèse nommée *Artificial Silicon Retina (ASR)* stimule la rétine grâce à un grand nombre d'électrodes auto-alimentées par les photodiodes qui les entourent (voir Fig. 1.11). Cette fonction de transformation de la lumière en électricité par les photodiodes remplace celle des photorécepteurs dégénérés à cause de la DMLA ou de la RP, mais elle stimule aussi les couches de cellules qui ne sont pas encore affectées.

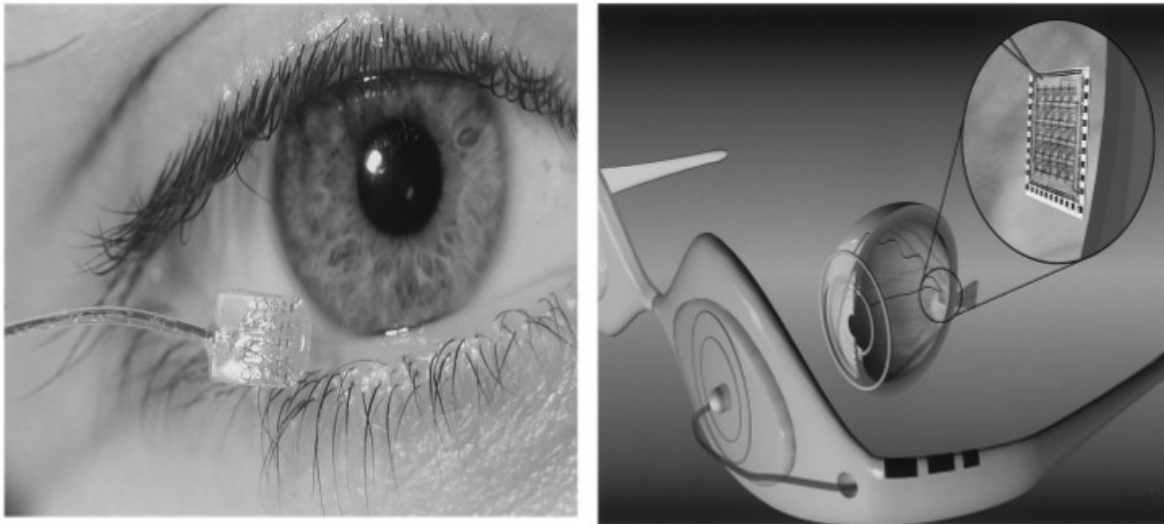


Figure 1.12. Système de prothèse épirétinienne, conçu par Humayun et al 1999

[Zrenner et al., 1999] propose un système similaire à l'ASR, le MPDA, avec une différence sur la façon dont on alimente les photodiodes. Ce groupe a montré que la quantité de lumière visible incidente qui peut atteindre le périphérique sous-rétinien est limitée et malheureusement insuffisante pour générer suffisamment de courant pour activer une surface de $200 \times 200 \mu\text{m}^2$ couverte de micro-photodiodes. En se fondant sur des calculs théoriques et une analyse empirique de l'efficacité de la photodiode et de l'impédance des électrodes, ils en déduisent que la quantité de lumière visible qui est nécessaire à cet effet est d'environ 2000W/m^2 [Schubert et al., 1999a, 1999b]. Empiriquement, cela a été montré par une expérience dans laquelle on faisait briller une tache de lumière (avec un faisceau laser) sur le dispositif MPDA et où on mesurait le courant à la surface des électrodes. Cette gamme est bien au-delà de l'intensité de la lumière du soleil sur la terre (la puissance surfacique d'une lampe fluorescente est de l'ordre de 10W/m^2 et celle de la lumière du soleil fait 100W/m^2). Pour surmonter ce problème, on utilise sur les photodiodes l'énergie supplémentaire fournie par la lumière du proche infrarouge [Schubert et al., 1999a, 1999b]. La lumière infrarouge est tolérée par la rétine jusqu'à des intensités de 2000W/m^2 , par opposition à la tolérance beaucoup plus faible de la rétine à la lumière visible. La viabilité de l'amélioration infrarouge pour augmenter la puissance de stimulation a été prouvée. Une poursuite de l'optimisation de ce dispositif est en cours [Schubert et al. 1999a]. La minimisation de la surface des électrodes de stimulation est une autre solution partielle pour ce même problème [Chow and Peachey, 1998].

Récemment, [Waschkowski et al., 2013] vise la maximisation du vecteur de stimulation sur la rétine en respectant la forme sphérique de l'œil tout en tenant compte des contraintes de miniaturisation. Ces travaux centrés sur le développement d'un *Very Large Array for Retinal Stimulation (VLARS)*. Les auteurs proposent quatre configurations différentes : (1) une forme circulaire continue ; (2) une forme spirale qui peut être déployée pendant la chirurgie ; (3) une forme en étoile qu'on peut replier ou recourber au cours de l'implantation chirurgicale et qui reprend forme une fois sur la rétine ; (4) une forme de "globe" qui peut se déployer sur la rétine une fois implantée (voir Fig. 1.13).

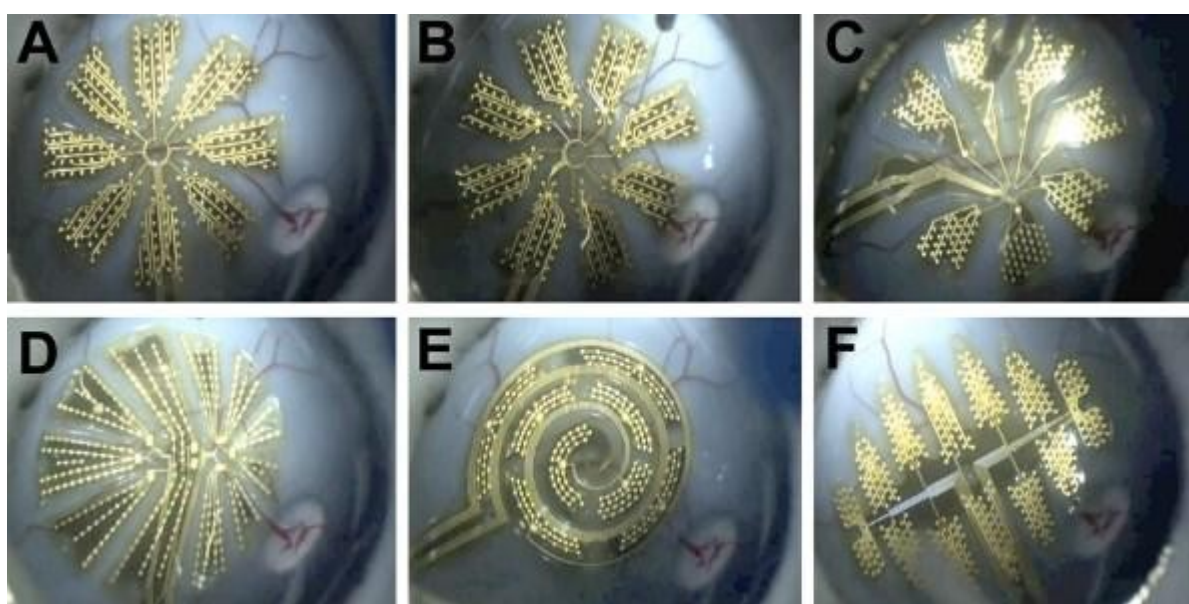


Figure 1.13. Implantation rétinienne du VLARS dans un œil de cochon. Formes A, B, C, D en étoile, E en spirale, F en forme de globe (source : Waschkowski et al. 2013).

Dans son dernier article, [Zrenner 2013] a étudié la substitution de la fonction des photorécepteurs pour les aveugles. L'implant développé comporte 1500 microphotodiodes qui entourent des microelectrodes dans une puce de 9 mm. Il était implanté chez neuf sujets aveugles. Trois sujets étaient capables de lire des lettres de l'alphabet spontanément, un sujet a dû passer par une période d'apprentissage. Ce dispositif de vision artificielle montre lui aussi que les implants sous-rétiniens sont capables de restituer une partie de la vision afin que des personnes atteintes de rétinopathies puissent accomplir à nouveau certaines tâches utiles de la vie quotidienne.

À l'issue de recherches sur les prothèses épi-rétiniennes est née la société *Second Sight*. Son objectif est de développer, fabriquer et commercialiser des prothèses visuelles implantables. Leur dernier produit, le système Argus II (voir Fig. 1.14) est formé de deux parties : une partie extérieure qui comporte l'unité de traitement reliée par un câble à des lunettes. Ces dernières sont équipées d'une caméra et d'une antenne. Une matrice de 8x8 électrodes est implantée sur la rétine. Elle fait partie, avec une antenne de réception (qui reçoit les informations de stimulation provenant du sous-système externe) placée sous le globe oculaire des sous-systèmes internes de l'Argus II. Ce dispositif est en phase d'essais.



Figure 1.14. : Le système ARGUS II comporte un implant rétinien, une caméra pour l'acquisition des images et une unité de traitement qui transforme l'information et gère la stimulation (2-sight.eu/en/system-overview-en)

Signalons enfin les travaux de [Tatur, 2011] qui visaient à tester, sans implantation chirurgicale certains concepts de la vision prothétique. Le dispositif employé est un casque d'immersion, ou *Head Mounting Display (HMD)*, avec projection d'une image fortement dégradée de 8x8 phosphènes sur la rétine (de patients sains). Ce dispositif a permis de tester notamment plusieurs techniques pour la mobilité des personnes aveugles en exploitant des données stéréoscopiques. En particulier, cette équipe s'est intéressée aux déplacements de sujets dans une pièce contenant des obstacles, la distance de l'utilisateur avec les éléments de

l'environnement pouvant être déduite de l'image provenant de 2 caméras. Les données spatialisées étaient ensuite restituées selon un mode sonar. On fournit en boucle une succession d'images de plus en plus distantes au sujet qui se fait ainsi une idée des obstacles qu'il va rencontrer. Avec ce même dispositif, cette équipe s'est également intéressée à la reconnaissance d'expressions faciales. Ce qui est important pour nous, dans cette thèse, c'est que l'utilisation qu'ils font de ce dispositif, qui n'est pas une prothèse, s'apparente à celle que nous voulons faire avec le TDU, à savoir tester des techniques ou des concepts qui seront utiles à la vision prothétique (par exemple la vision stéréoscopique en mode sonar pour Tatur et al 2011, les saccades oculaires dans notre cas).

1.3.1.2. Stimulation du nerf optique

Les recherches en vision artificielle ont montré qu'on pouvait stimuler directement le nerf optique [Veraart et al., 1998] (Fig. 1.15). Le nerf optique est un "câble compact" formé des axones des cellules ganglionnaires (voir la section anatomie de l'œil). Ce câble de transfert de pulsations nerveuses est positionné dans une région anatomique assez facile d'accès pour être stimulée. Avec cette technologie, des phosphènes colorés ont été perçus dans le champ visuel d'une personne affectée par la rétine pigmentaire.

Néanmoins, avec plus de 1.2 millions d'axones compactés dans environ 2 mm de diamètre, de nombreuses difficultés ont été rencontrées comme la focalisation des *stimuli* et la perception détaillée. D'autres problèmes existent, liés à l'intervention chirurgicale qui nécessite une dissection de la dura (membrane externe des méninges qui entoure le SNC dont fait partie le nerf optique) et qui provoque un certain nombre d'effets secondaires comme une circulation sanguine perturbée dans cette région, une infection possible du système nerveux central ... Il faut de plus, pour que cette technique soit applicable, que la couche des cellules ganglionnaires (voir anatomie section 1.1.1) soit saine.

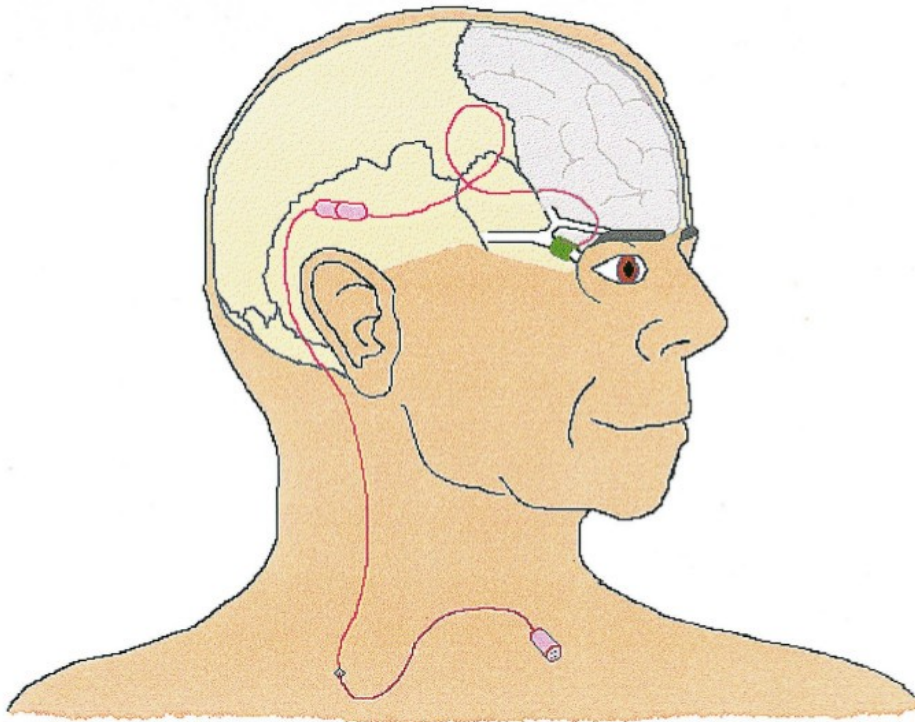


Figure 1.15. Système de stimulation du nerf optique proposé par Veraart 1998

1.3.1.3. Stimulation du cortex

La première cible de stimulation en vision artificielle était le cortex visuel (avant que les recherches ne s'attaquent au nerf optique puis à la rétine). Cette catégorie de prothèses se décline en deux techniques :

- *stimulation sur la surface corticale* (Fig. 1.16 et 1.17) : dans une des premières expériences qui incluait trois patients aveugles, deux d'entre eux ont été capables de localiser une source lumineuse grâce à un système formé d'une photo-cellule qui capte la lumière, d'un câble qui pénètre le cuir chevelu et le crâne pour arriver au cortex visuel et d'une matrice de 80 électrodes [Button and Putnam, 1962]. Ce système avait beaucoup d'inconvénient comme l'interaction entre les phosphènes, une même électrode pouvant induire plusieurs phosphènes, les électrodes ayant une grande taille. Dobelle et ses collègues ont alors proposé un nouveau système formé d'une caméra vidéo, d'une unité de calcul "intelligente" et d'un stimulateur du cortex visuel [Dobelle, 2000]. Leur matrice d'électrodes est formée de 64 disques en platine reliés à une caméra. Le couplage entre la caméra et les électrodes est assuré par une matrice de 100x100 photo-transistors. Le travail de Button et

Putnam 1962 était une étude préliminaire afin de tester les réponses du cortex visuel aux stimulations électriques alors que Dobelle a développé un système presque complet pour faire de la stimulation avec une caméra et une unité intelligente qui traite l'information avant de stimuler le cortex visuel. Avec de tels systèmes, l'aveugle est capable d'accomplir quelques tâches comme la détection d'objets (comme des lettres) et la reconnaissance de mannequins.

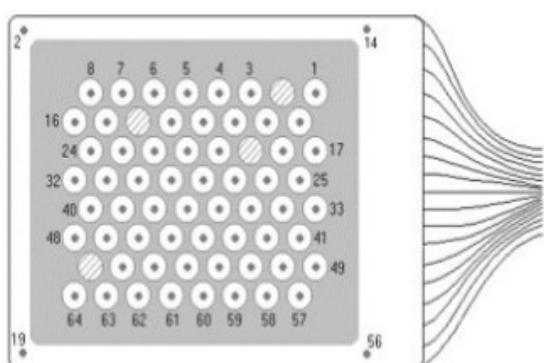


Figure 1.16. Prothèse corticale (à gauche). Matrice d'électrodes implantée chez un volontaire et vue sous radiographie X (à droite). Dobelle et al. 2000



Figure 1.17. Système de vision artificielle conçu par Dobelle et al 2000 (à droite) avec la caméra vidéo miniaturisée (à gauche). Dobelle et al. 2000

- *micro-stimulation intra-corticale* : la taille réduite des électrodes utilisées dans cette technique offre plus de précision car la stimulation du tissu cortical est plus locale. Théoriquement, ce type de stimulation réduit l'interaction entre phosphènes, l'induction multi-phosphène par une même électrode et l'inconsistance des phosphènes. Les résultats sont meilleurs que la première approche (stimulation sur la surface corticale). Des micro-électrodes en iridium, de 37.5 µm de diamètre, ont déjà été insérées dans le cortex occipital d'un patient, un sujet âgé de 42 ans et aveugle depuis 22 ans. L'équipe de recherche [Schmidt et al. 1996] en charge de ces travaux a ainsi effectué une étude de faisabilité d'une prothèse formée de 38 micro-électrodes dans une zone de cortex visuel de 40.8 x 19.2 mm. Ces travaux ont permis l'optimisation de la technologie en terme de seuil de stimulation, espace entre les électrodes et niveau de stimulation. Ce dernier peut induire une perception de différentes couleurs qui varient entre le blanc, le gris et le jaune. Les avantages de cette méthode (micro-stimulation intra-corticale) ont causé l'abandon de la solution de Dobelle.

1.3.2. Thérapie cellulaire

Des scientifiques britanniques de l'*University College* de Londres dirigés par R. Ali ont franchi un nouveau cap. Ce ne sont plus des cellules de soutien qui ont été conçues depuis des cellules souches embryonnaires (CSE), mais bel et bien des photorécepteurs. Comme ils l'expliquent dans la revue *Nature Biotechnology*, une fois transplantées chez l'animal, ces cellules semblent occuper leur position et jouer leur rôle en établissant les liaisons avec le nerf optique [Gonzalez-Cordero et al., 2013].

Les auteurs ont réalisé une culture en trois dimensions de CSE de souris de manière à les pousser vers une destinée de photorécepteurs. Pour s'assurer de la bonne différenciation des cellules, ils ont vérifié que les gènes exprimés étaient biologiquement équivalents. Une rétine de synthèse a ainsi pu croître en boîte de culture.

Environ 200 000 photorécepteurs artificiellement créés ont alors été transplantés dans la rétine de souris aveugles. Trois semaines après l'opération, les cellules avaient migré et avaient trouvé leur position dans l'œil de leur hôte. Elles avaient l'aspect de cellules en

bâtonnet matures normales et étaient toujours sous cette forme six semaines plus tard. Des synapses, les connexions avec le système nerveux, ont même pu être observées.

1.4. Conclusion de l'état de l'art sur la vision et ses pathologies

Au cours des sections précédentes, on a présenté quelques principes essentiels du fonctionnement de la vision normale, certaines pathologies de la vision et les développements en vision artificielle.

On retiendra que la vision est un processus dynamique qui a besoin du mouvement ; le rôle des micro-saccades oculaires en particulier est souligné dans la section 1.1.2. La plupart des expériences faites sur les dispositifs de vision artificielle comme les implants rétiniens sont des stimulations avec des scènes dynamiques (ex: on fait défiler du texte, la camera bouge en suivant les mouvements de la tête du patient, ou bien les sujets doivent regarder un objet en mouvement passant devant une caméra). Mais dans le cas de scènes visuelles statiques, quel sera le résultat de la perception ? On citera deux cas qu'on trouve dans la littérature où l'atténuation (fading) des images perçues est remarqué chez des sujets lorsque la stimulation persiste [Zrenner et al. 2011, Stingl et al. 2013]. Ce phénomène est dû à la saturation (bleaching) des cellules photoréceptrices stimulées. Les périodes de saturation sont différentes d'un sujet à un autre. Comme solution, dans ces articles, les auteurs effectuent une diminution de la fréquence de stimulation à 1-2 Hz au lieu de plus de 15 Hz pour les autres patients. Cela correspond à un clignotement de l'image perceptible par l'utilisateur, ce clignotement permettant la régénération de l'activité des photorécepteurs. [Fornos et al 2012] rencontre le même souci avec l'implant ARGUS II. La solution est de bouger la tête pour "rafraîchir" les cellules stimulées. Ce mouvement de tête nous rappelle le cas clinique de Mme A.I. abordé à la section 1.1.2 : il permet à chaque instant de fournir aux photorécepteurs une stimulation différente, ce qui empêche le phénomène d'atténuation. Ainsi, on peut dire que le mouvement des scènes visuelles projetées sur la rétine est primordial pour leur perception. Un des objectifs de cette thèse est de trouver une solution autre que le mouvement de la tête ou l'abaissement des fréquences de stimulation pour maintenir la perception, même dans le cas de scènes visuelles peu dynamiques ou statiques. C'est ainsi qu'on appliquera des traitements semblables à ce que produisent les mouvements des yeux aux signaux-images de stimulation : des saccades ou des scintillements des images [Chekhchoukh and Glade, 2012].

Comme alternatives à la stimulation directe des neurones du système visuel, d'autres approches transforment l'information visuelle en d'autres types d'informations, auditive ou tactile, afin que la personne aveugle ou le mal-voyant perçoive cette information avec une modalité de perception autre que la vision. Ces dispositifs de suppléance perceptives sont portables et non implantables ce qui représente un avantage majeur. Cependant, nous verrons qu'ils n'ont pas reçu une acceptation importante par les patients concernés car (i) ils ne rendent pas vraiment le sens de la vision (ii), leur basse résolution ne permet pas une perception nécessaire pour être autonome (iii), ils monopolisent un autre sens qui ne peut alors plus être utilisé normalement et (iv) ils demandent souvent une période d'apprentissage qui peut durer longtemps.

Dans le chapitre suivant on fera une introduction au domaine de la suppléance perceptives et plus précisément de la substitution visuelle. On présentera le paradigme de substitution sensorielle avec un aperçu des applications développées à ce jour fondées sur ce paradigme. On exposera aussi les problèmes rencontrés pour lesquels on pourra proposer certaines solutions.

2 Suppléance perceptive

Comme on a vu au début du chapitre précédent, notre perception du monde est multimodale. Chaque modalité perceptive utilise des organes spécifiques du corps humain pour acquérir l'information (notion de capteur). Ensuite le signal généré par ces organes est transféré par les cellules nerveuses puis traité dans des aires corticales spécifiques au niveau du cerveau.

D'autre part, on a vu aussi comment fonctionnaient plusieurs dispositifs de vision prothétique. Soulignons tout d'abord que les interventions chirurgicales associées sont complexes et que les performances des prothèses ne sont pas encore suffisamment au point pour que ces technologies de vision prothétique soient utilisées systématiquement dans le grand public (on rappelle la forte propension de DMLA dans la population mondiale). Ajoutons à cela que la population de malade sur laquelle on peut appliquer ce genre de technique est réduite par rapport au nombre total de malvoyants, même dans les populations de malades concernées par ces technologies à cause notamment de l'ensemble des contre-indications à la chirurgie et du degré parfois trop avancé de certaines pathologies.

Le concept de substitution visuelle vise à remplacer la vision par une autre modalité comme l'audition ou le sens tactile. Cette alternative à la prothèse sensorielle se présente sous la forme d'un ensemble de solutions non invasives qui nécessitent souvent un apprentissage, mais qui donne des résultats prometteurs. Au cours de ce chapitre nous allons aborder le concept de suppléance perceptive.

2.1. Perception tactile

Le système somesthésique est formé d'au moins deux sous-systèmes : un premier sensible à des *stimuli* mécaniques comme les vibrations, la pression et même un toucher léger ou une distension cutanée. L'autre sous-système s'occupe de la détection des *stimuli* douloureux et de la température [Purves et al., 2005]. Cette variété de capteurs est capable

d'acquérir de multiples types d'informations (mécanique, thermique, ...) qui, à la fin, sont tous traités au niveau du cortex somatosensoriel.

Le système associé à la perception tactile respecte la même architecture que celle des autres systèmes de perception sauf que ses capteurs sont distribués de partout sur la surface de la peau (et des muqueuses en général, notamment la langue) ainsi que dans plusieurs couches sous la peau. De plus, la densité en capteurs n'est pas la même dans les différentes régions du corps, raison pour laquelle le seuil de perception n'est pas le même dans une zone comme la main par rapport au reste du corps, ou même entre différentes parties de la main (le bout du doigt est plus sensible que le reste de la main). Nous verrons que c'est aussi le cas sur la langue.

2.1.1. Anatomie de la peau et des muqueuses

Notre corps est couvert par la peau (et à l'intérieur des muqueuses), un organe formé de plusieurs couches ayant différentes fonctions. La fonctionnalité tactile est l'une des plus nécessaires pour l'interaction de la personne avec l'environnement. Afin d'assurer la perception des *stimuli* provenant de l'environnement, les tissus cutanés et sous-cutanés sont très denses en capteurs. Selon le type de stimulus reçu, on distingue trois types de capteurs dans la peau et les muqueuses : les mécano-récepteurs, les nocicepteurs et les thermorécepteurs.

L'épiderme est la couche extérieure protectrice de la peau. En dessous se trouve le derme qui contient toute une gamme de capteurs comme les corpuscules de Meissner, Pacini, Ruffini, les disques de Merkel et des terminaisons nerveuses libres (Fig. 2.1). Globalement, le mécanisme de ces différents capteurs est le suivant : des *stimuli* appliqués à la peau déforment ou modifient de quelque façon les terminaisons nerveuses, ce qui affecte la perméabilité ionique de la membrane du récepteur. Ces changements de perméabilité produisent dans les terminaisons nerveuses un courant dépolarisant qui donne naissance à un potentiel de récepteur (potentiel générateur) déclenchant à son tour un potentiel d'action. L'ensemble de ces processus par lesquels l'énergie d'un stimulus est convertie en un signal électrique dans le neurone sensoriel constitue la transduction sensorielle.

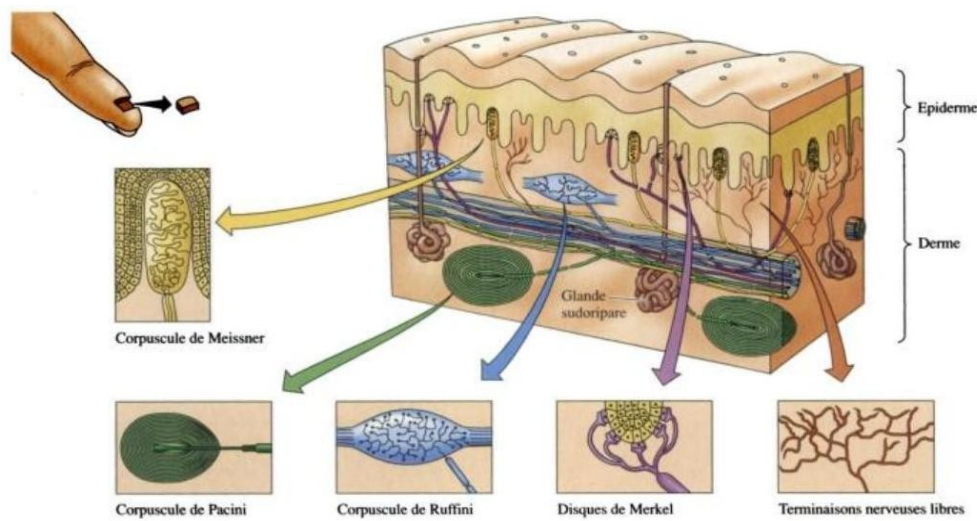


Figure 2.1. Types de capteurs (mécanorécepteurs) existant dans la peau au bout du doigt [Purves 2005]

Chaque capteur a une forme spécifique, afin d'assurer l'acquisition d'une information tactile avec des fréquences différentes et des forces variables. Le type d'information que chaque capteur est capable de percevoir ainsi que sa gamme de fréquences (vibrations) et l'étendue de son champ de réception sont divers. Les caractéristiques de ces capteurs sont résumées dans le tableau (Tab. 2.2). Les disques de Merkel (SA-I) et les corpuscules de Ruffini (SA-II) capteraient principalement les informations sur la forme spatiale et la texture des objets (comme des lettres en relief ou le Braille) alors que les corpuscules de Meissner (RA-I) et les corpuscules de Pacini (RA-II) capteraient surtout les mouvements à la surface de la peau (comme le glissement d'un verre tenu dans la main) [Blake et al., 1997]. Les deux premières structures (SA-I et SA-II) ont une réponse fréquentielle comprise entre 1 Hz et 200 Hz et s'adaptent donc plutôt lentement, mais montrent une résolution spatiale importante. Les deux autres structures (RA-I et RA-II) ont une réponse fréquentielle comprise entre 10 Hz et 1000 Hz et s'adaptent rapidement mais au prix d'une résolution spatiale plus faible. La distribution spatiale de ces récepteurs est non uniforme sur la main et leur pourcentage relatif est respectivement de 43% (corpuscules de Meissner), 25% (disques de Merkel), 13% (corpuscules de Pacini) et 19% (corpuscules de Ruffini) [Shimoga, 1992].

Récepteur	Perception	Forme	Durée de perception	Taille du champ de réception	Fréquences perçues
Merkel	Pression lente: légère tape, petits détails	Disque	Tant que le stimulus est présent	petit (1 mm)	0.3 à 3 Hz
Meissner	Pression rapide: tremblement, frottement	Boudin	Ne dure pas	petit	3 à 40 Hz
Ruffini	Étirement	Cylindre	Tant que le stimulus est présent	grand (8 mm)	15 à 400 Hz
Pacini	Vibration rapide	Oignon	Ne dure pas	grand	10 à 500 Hz

Tableau 2.2. Propriétés des capteurs présent dans la peau et capables d'encoder l'information vibrotactile en signaux nerveux.

2.1.2. La langue et sa capacité perceptive

Avant de décrire les méthodes de stimulation tactile nous allons décrire brièvement l'anatomie de la langue et sa capacité multi-perceptive. La sensation chimique est certainement la plus importante pour la majorité d'entre nous. Nous allons cependant nous intéresser ici plutôt à la perception tactile de la langue.

La langue de l'homme [Larousse Médical] est divisée en deux parties (Fig. 2.3) : une portion fixe, ou base de la langue et une portion mobile (ou libre).

- La *face supérieure* ou *dorsale* de la langue est recouverte de muqueuse et présente à l'union de son tiers postérieur et de ses deux tiers antérieurs un sillon en V ouvert en avant, le *sillon terminal*, vestige du développement embryologique qui sépare les portions buccale et pharyngienne de l'organe. Le sommet du sillon est marqué par une dépression, ou *foramen cæcum*.

La partie buccale de la face supérieure de la langue présente une surface inégale, marquée par la saillie de nombreuses papilles.

La partie pharyngienne de la face dorsale de la langue est presque verticale et regarde le pharynx. Sa surface est inégale, marquée par la présence d'amas de tissu lymphoïde, correspondant à l'amygdale linguale.

- La *face inférieure* de la langue est recouverte par une muqueuse lisse, transparente, soulevée sur la ligne médiane par un repli muqueux, le frein de la langue, qui se recourbe jusqu'à la partie médiane du maxillaire inférieur. De part et d'autre du frein se situent les caroncules sublinguales, orifice des glandes salivaires (sublinguales).

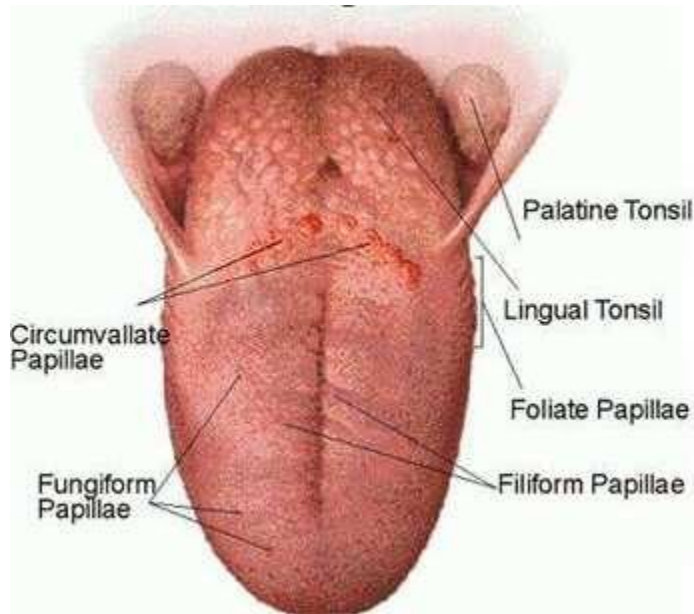


Figure 2.3. Face supérieure (ou dorsale) de la langue humaine

La langue est extrêmement sensible aux stimulations tactiles comme cela a été montré et mesuré par de nombreux travaux [Ringel and Ewanowski, 1974; Trulsson and Essick, 1997]. L'étude menée par l'équipe de Trulsson en 1997 a ainsi défini une cartographie des régions les plus sensibles à l'électro-stimulation de la langue, par stimulation mécanique des différentes régions de la langue et des mesures simultanées au niveau du nerf innervant ce muscle (Fig. 2.4).

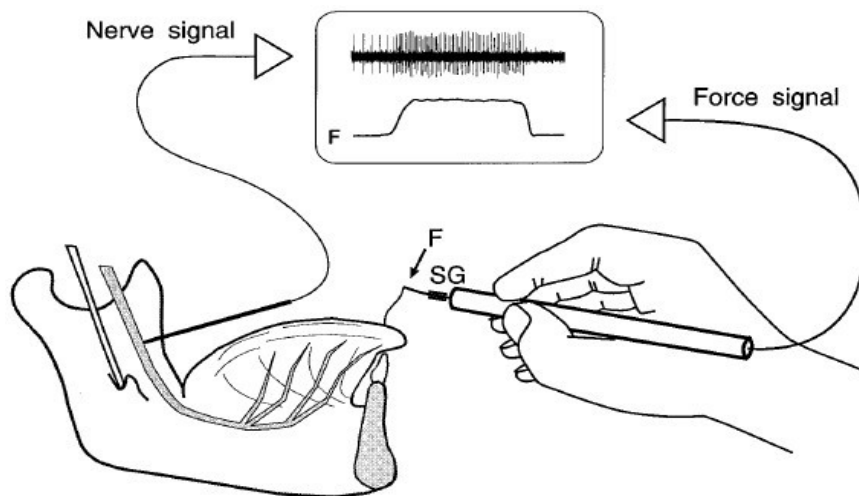


Figure 2.4. Stimulation de la partie dorsale de la langue et mesure des signaux sur le nerf de la langue [Trulsson & Essick 1997]

L'ensemble des résultats de cette étude (Fig. 2.5) suggère une tendance à l'asymétrie de la perception tactile sur la langue ainsi que la mise en évidence qu'une grande majorité des récepteurs superficiels de la peau du dos de la langue sont des récepteurs à adaptation rapide (RA) du type « corpuscule de Meissner ». Le reste étant des récepteurs à adaptation lente (SA). De plus, les récepteurs étaient caractérisés par un seuil d'activation extrêmement bas.

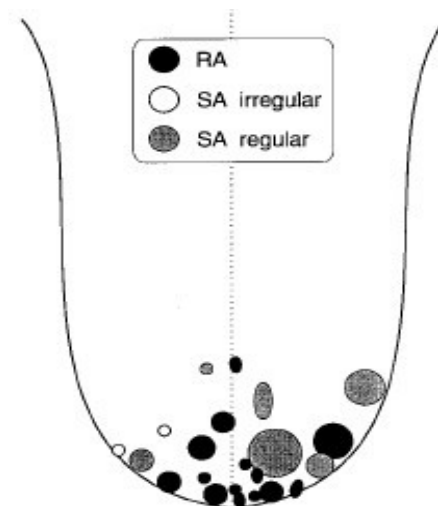


Figure 2.5. Types des capteurs du champ récepteur de la langue [d'après Trulsson & Essick 1997]

La langue est donc un organe intéressant pour percevoir de l'information tactile mécanique ou électrique (vibro-électrique) avec de nombreux avantages notamment pour la substitution visuelle par électro-stimulation de la langue. C'est ce que nous allons détailler dans la section 4.1.

2.1.3. Méthodes de stimulation tactile

Pour stimuler de façon tactile le corps humain il existe plusieurs méthodes. Ces méthodes sont étroitement liées aux propriétés du stimulateur utilisé, au temps de stimulation et au type des capteurs biologiques.

- **Vibro-stimulation** : consiste à faire vibrer un objet contre la surface de la peau ou d'une muqueuse. Typiquement, on utilise des fréquences minimales de 20 à 40 Hz qui parviennent à activer les récepteurs de Pacini. En ce qui concerne les hautes fréquences, [Roberts et al., 2000] étudient le choix des fréquences optimales en stimulation vibro-tactile haute fréquence. Ils obtiennent une valeur optimale de l'ordre de 250 Hz, valeur utilisée dans l'OPTACON (voir section 2.2.2.3).

En termes d'applications, dans [Bliss et al., 1970] on trouve une méthode pour la conversion de l'information optique d'une image (acquise par camera) vers une matrice vibrotactile de 24x6 pattes piézoélectriques. Ce système est destiné aux aveugles.

- **Electro-stimulation** : la majorité des études confirment que le courant électrique peut passer à travers la peau et qu'il stimule directement les afférents nerveux [Butikofer and Lawrence, 1978; Pfeiffer, 1968; Rollman, 1974] suggère aussi que des petites électrodes de 1 mm² puissent stimuler directement des récepteurs individuels. L'électro-stimulation directe du nerf du doigt a aussi été envisagée pour réaliser une prothèse substitutive pour encoder un signal auditif [Blamey and Clark, 1985].

Les sujets qui reçoivent une électro-stimulation décrivent la sensation comme un picotement, une démangeaison, une vibration, un *buzz*, une sensation de toucher, une pression, un pincement ou une forte sensation de brûlure ; tout dépend du voltage utilisé durant la stimulation, de la taille des électrodes, du matériau utilisé pour les électrodes et de l'hydratation à l'interface électrode-organe sensoriel [Collins, 1970; Gibson, 1968; Mason and MacKay, 1976; Pfeiffer, 1968].

- Il existe d'autres méthodes de retour sensoriel tactile comme l'haptique, le thermique et les flux d'air ou liquides, mais dans la suite de ce manuscrit on se focalise sur les systèmes à base d'électro-stimulation et de vibro-stimulation.

Les aires corticales du cortex somesthésique s'occupent du traitement des informations envoyées par les différents capteurs. Le terme de *somatotopie* est utilisé pour désigner la représentation de la surface cutanée sur la surface du cortex, plus précisément au sein de l'aire somesthésique. En d'autres termes, la somatotopie désigne la représentation du corps, point par point, sur une carte de la discrimination spatiale des informations sensibles permises par le système nerveux. L'importance de la représentation des parties du corps est aussi en rapport avec la densité des capteurs dans chaque région du corps. Des études antérieures ont cartographié cette projection des récepteurs, notamment les mécanorécepteurs, sur ce qu'on appelle l'*homonculus* sensitif (Fig. 2.6).

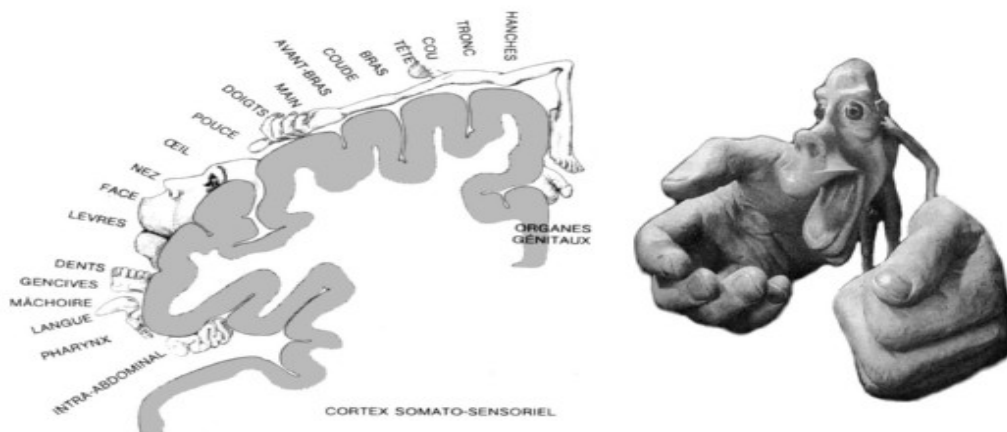


Figure 2.6. (à gauche) Représentation du cortex somatosensoriel et la répartition des capteurs somesthésiques. (à droite) corps humain représenté en fonction de la densité de mécano-récepteurs

De plus, une étude menée par [Solomonow et al., 1977] qui traite de la discrimination entre deux points électro-stimulés montre que le bout du doigt est la région la plus discriminante avec une valeur de 7.25 mm de séparation minimale des points. Les valeurs obtenues varient selon les organes et sont liées au mode de stimulation utilisé (vibro ou électro-stimulation) (Tab. 2.7).

Organe	Electrostimulation (mm)	Vibrostimulation (mm)
Bout du doigt (Finger tip)	7.25	2
Orteil (Sole)	7.67	-
Paume de la Main (Palm)	7.73	-
Avant bras (Forearm)	8.93	-
Mollet (Calf)	9.06	-
Épaule (Shoulder)	9.17	-
Bras supérieur (Upper arm)	9.48	-
Ventre (Belly)	9.78	-
Dos (Back)	9.79	11-18
Cuisse (Thigh)	9.88	-
Poitrine (Chest)	10.23	-

Tableau 2.7. Distances de discrimination entre deux points stimulés par un signal électro ou vibrotactile.

2.2. La suppléance perceptive et ses applications

La déficience perceptive chez l'humain est fréquente et a pour origine une malformation de naissance (problème de développement ou génétique), une maladie et bien entendu le vieillissement. Après un certain âge les organes de perception (l'œil, oreille, ...) commencent à être fatigués et leur rendement devient insuffisant pour satisfaire les besoins quotidiens de la personne. Notre rôle en tant que chercheurs dans ce domaine est notamment de retrouver des solutions à ces problèmes et de développer des technologies pour remédier à ces déficiences. Les spécialistes de la suppléance perceptive proposent un paradigme pratique et efficace pour ces personnes.

2.2.1. Le paradigme de substitution sensorielle et la suppléance perceptive

De manière générale, la substitution sensorielle ou suppléance perceptive se fonde sur le principe de remplacement et/ou suppléance d'un sens déficient ou occupé par un autre sens sain. Ce paradigme a été avancé par Paul Bach-y-Rita dans les années 1960 dans le cas particulier de la substitution visuelle. Il stipule que ce n'est pas l'organe qui voit, mais le

cerveau. Cela signifie que si on transmet des informations visuelles par un autre canal que le nerf optique, alors ces informations pourront être traitées comme telles par le cortex visuel.

Nous avons donné une définition utilitaire de la substitution sensorielle et de la suppléance perceptive ci-dessus en plaçant la substitution sensorielle comme paradigme général et la suppléance comme son application. Précisons toutefois qu'il existe un débat d'utilisation des deux termes suppléance et substitution au point qu'on considère actuellement qu'il ne s'agit pas exactement de la même chose. Ce débat vient pourtant de la traduction du terme substitution de l'anglais vers le langage français, la traduction ayant généré 2 termes auxquels on a associé 2 définitions différentes. Le terme suppléance a été proposé par [Sampaio and Dufier, 1988] étant donné que le système proposé venait pour suppléer une information visuelle manquante. [Lenay et al., 2000] donnent un autre argument qui défavorise le terme « substitution » disant que les dispositifs existants ne parviennent pas réellement à combler le sens déficient mais qu'ils facilitent l'acquisition de l'information par une autre modalité de perception. La perception par un organe biologique ne produit pas la même sensation psychologique que les dispositifs techniques de couplage sensori-moteurs. Les systèmes de suppléance sont cependant toujours en cours de développement provoquant une évolution technologique des capteurs, processeurs et stimulateurs, ce qui influence la qualité et la valeur émotionnelle de la perception et qui, à terme, amoindrira peu à peu ce *distinguo* entre substitution sensorielle et suppléance perceptive.

2.2.2. Applications

Le principe de fonctionnement des systèmes dits de substitution sensorielle ou suppléance perceptive consiste à transformer des *stimuli* propres à une modalité sensorielle en *stimuli* d'une autre modalité sensorielle (ex: vision traduite en audition). Le système comprend des capteurs et des unités de traitement qui permettent de convertir une forme d'énergie en d'autres signaux pouvant être interprétés par un sous-système responsable de l'activation coordonnée de stimulateurs (Fig. 2.8).



Figure 2.8. Suppléance perceptive - principe

Un nombre important d'aides sensorielles et de prothèses utilisant ce principe de suppléance perceptive ont été élaborées depuis les premiers travaux dans ce domaine. Certaines sont encore à l'heure actuelle en voie de développement. La qualité et les performances des capteurs et celles des actionneurs se développent de plus en plus. Chaque année sont mis au point des circuits électroniques plus complexes pour assurer des fonctions plus complexes avec des temps de traitement plus rapides. Les logiciels évoluent en parallèle de ces dispositifs électriques donnant également à ceux-ci plus de fonctionnalités.

Ces dispositifs, techniquement très variés (diversité des types de capteurs, de stimulateurs, etc.) (voir Fig. 2.9) répondent à des objectifs de recherche appliquée (aide aux personnes handicapées, prévention des troubles du développement, ergonomie des interfaces, implémentation et validation de dispositifs) et/ou de recherche fondamentale (étude de la cognition en général, des procédures expérimentales dans ce domaine et du processus d'apprentissage).

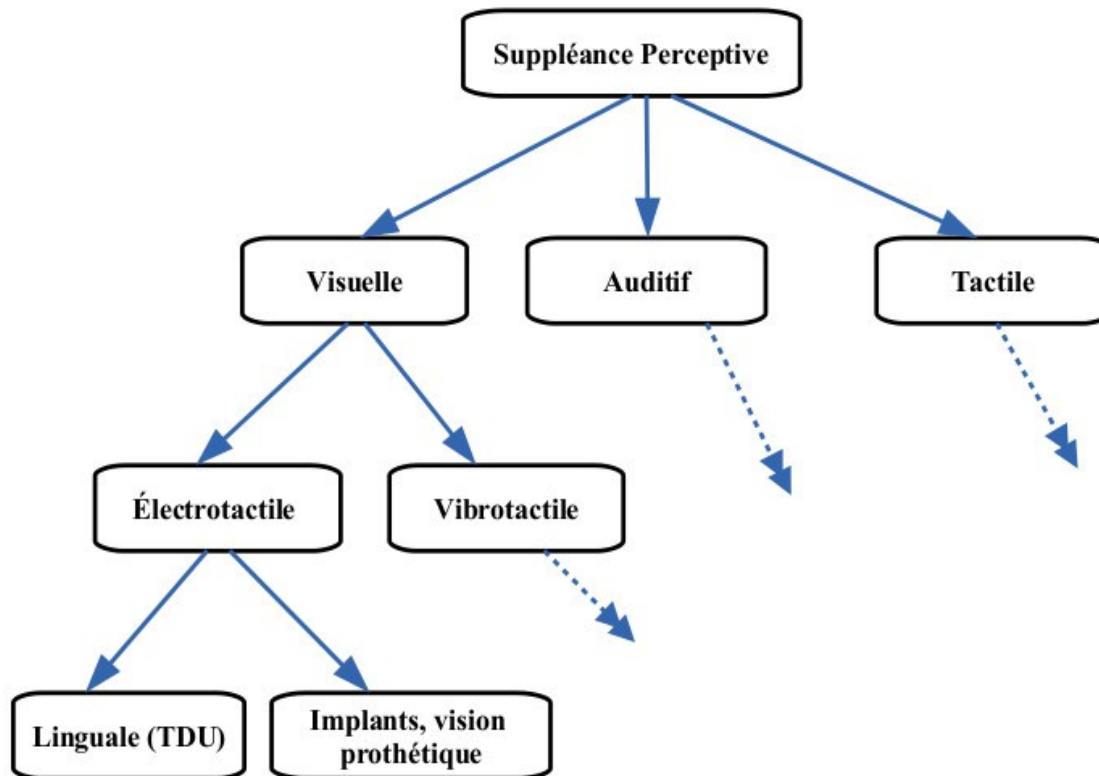


Figure 2.9. Suppléance perceptive - diversité des approches

Le tableau suivant (Tab. 2.10) représente une chronologie historique des travaux liés à la substitution sensorielle.

Année	Tâche souhaitée	Information suppléante (position)	Référence
1963	Vision	Vibrotactile	[Palacz and Kurcz, 1978]
1965	Audition	Électrotactile (dos)	[DIES, 1965]
1969	Vision	Vibrotactile (dos)	[Bach-y-Rita et al., 1969]
1981	Vision	Vibrotactile (dos)	[Jansson and Brabyn, 1981]
1988	Vision (enfants)	Vibrotactile (doigt)	[Miletic et al., 1988]
1992	Vision	Audition	[Meijer, 1992]
1998	Vision	Audition	[Capelle et al 1998]
2003	Vestibule	Électrotactile (langue)	[Tayler et al 2003]
2003	Vestibule (verticale)	Vibrotactile (veste)	[Erp et al., 2003]
2005	GPS	Vibrotactile (ceinture)	[Van Erp et al 2005]
2006	Vision	Vibrotactile (main)	[Velázquez et al.,

			2006]
2006	Guidage	Électrotactile (langue)	[Robineau et al 2006]
2006	Proprioception	Électrotactile (langue)	[Vuillerme et al., 2006]
2007	Pression plantaire	Électrotactile (langue)	[Pinsault et al 2007]

Tableau 2.10.: Chronologie des études avec des systèmes de substitution sensorielle.

2.2.2.1. Tâches liées à la vision et substitution de la vision

Quand une personne est privée de sa vision (mal-voyance, cécité ou lorsque le sens de la vision est occupé par un autre sujet d'intérêt comme une surveillance visuelle), la personne est privée d'une information cruciale et peut s'en trouver très handicapée. La recherche d'un autre mode de perception pour acquérir l'information perdue est une obligation pour accomplir la tâche souhaitée. Illustrons ceci par une expérience souvent vécue : notre attitude face à une coupure d'électricité la nuit. Plongés dans le noir, on va se mettre à chercher une autre source de lumière comme les allumettes et les bougies qui sont dans le placard. On dispose d'un avantage : on a en effet en mémoire une image mentale préalable de l'architecture de la maison et on sait où se trouve la cible. Cependant pour se déplacer jusqu'à cette cible il faut le faire en absence de toute vision (navigation à l'aveugle). C'est une tâche pénible : on se déplace doucement mettant nos mains devant pour toucher les murs ou les obstacles qui peuvent nous déstabiliser ou nous blesser. Arrivés à la cible, on commence à fouiller parmi les objets du placard : il s'agit maintenant d'une tâche de reconnaissance d'objets, encore une fois assurée par le sens de toucher. Dans ce genre de situations, des images mentales des objets et de la pièce nous parviennent. Le cortex visuel est activé suite à l'évocation mnésique des scènes visuelles et des formes mémorisées provoquée par la reconnaissance des lieux et formes due au toucher.

Cette tâche était relativement simple et la privation de vision était momentanée. Les aveugles de naissance peuvent développer une capacité extraordinaire à utiliser leurs autres sens. C'est le cas des savants aveugles (parfois nommés *super aveugles*). On cite le cas d'un savant d'Andalousie, Ibn Sidah (1007-1066), qui était aveugle de naissance (il s'agissait vraisemblablement d'un trouble héréditaire ; son père l'était aussi). Ibn Sidah était spécialiste de l'histoire et de la langue arabe. Malgré son handicap visuel il a pu apprendre les sciences par l'audition et le toucher. Parmi ses ouvrages, on cite le *Muhkem* composé de 28 volumes.

Ibn Sidah avait développé une technique de reconnaissance des livres par le toucher. Le détail de sa technique n'est pas connu mais il est possible qu'il ait pu suivre le relief des lettres imprimées (l'impression, dont l'impression xylographique qui laisse un relief négatif, étant connue dans le monde musulman depuis l'an 900) ou gravées.

On trouve un autre cas dans la *Lettre sur les aveugles à l'usage de ceux qui voient*, essai rédigé par Diderot en 1749 [Diderot, 1901]. Diderot présente le cas d'un mathématicien britannique, Nicholas Saunderson (1683-1739), lui aussi aveugle depuis l'âge de 12 mois. Saunderson donne des conférences sur la physique, notamment la lumière et les couleurs de l'arc-en-ciel, des phénomènes que Saunderson n'a jamais vu. Dans ce contexte, se pose alors une question, à la fois philosophique et scientifique, quant à notre capacité à percevoir le monde et les phénomènes qui nous entourent. Cette question a été formulée par le fameux problème de Molyneux en 1688 (et inspirée par le livre d'Ibn Tufayl 1110-1185). Molyneux a envoyé une lettre à John Locke, philosophe anglais du 17^{ème} siècle, en lui posant la question suivante :

« A Man, being born blind, and having a Globe and a Cube, nigh of the same bignes, Committed into his Hands, and being taught or Told, which is Called the Globe, and which the Cube, so as easily to distinguish them by his Touch or Feeling; Then both being taken from Him, and Laid on a Table, Let us Suppose his Sight Restored to Him; Whether he Could, by his Sight, and before he touch them, know which is the Globe and which the Cube? Or Whether he Could know by his Sight, before he stretch'd out his Hand, whether he Could not Reach them, tho they were Removed 20 or 1000 feet from Him? ».

Ce que demande Molyneux, c'est est-ce qu'un aveugle de naissance qui aurait appris à distinguer parfaitement un cube d'une sphère par le sens de toucher et par l'exploration des deux objets dans ses mains, s'il retrouvait par miracle la vue un jour, serait cette fois capable de distinguer les deux objets par la seule perception visuelle et sans les toucher ? La réponse était *a priori* non, puisque l'on peut raisonnablement penser que la représentation cérébrale des objets perçus par le sens du toucher n'est pas la même que celle obtenue par le sens de la vision. Ainsi l'interprétation des deux tâches se ferait de deux manières différentes dans deux régions différentes de notre cerveau. Comme nous le verrons, le paradigme de substitution sensorielle va à l'encontre de cette affirmation.

Diderot, dans son essai sur les aveugles [Diderot, 1901] présente aussi un langage fondé sur des matrices dotée d'épingles avec des têtes épaisses (que l'aveugle peut toucher) pour représenter des informations visuelles (nombres, opérations algébriques et formes géométriques). Ce système conçu par Saunderson permet de réaliser des calculs en utilisant des grilles équipées ainsi d'épingles ayant des positionnements différents pour coder les chiffres de 0 à 9 (voir Figs. 2.11 et 2.12). Ce codage était bien spécifique à Saunderson et n'est valable que pour des opérations mathématiques. Il restait alors à trouver un langage commun pour permettre la lecture à un maximum d'aveugle. En 1829, le français Louis Braille (1809-1852) invente un système d'écriture tactile qui porte son nom. Le langage Braille est utilisé aujourd'hui presque partout (livres, médicaments, transport en commun, ..., etc).



Figure 2.11. Système de codage des chiffres conçu par Saunderson. À gauche, les différentes position que peuvent prendre les épingles pour désigner un chiffre. À droite, un exemple de code avec le numéro 4.

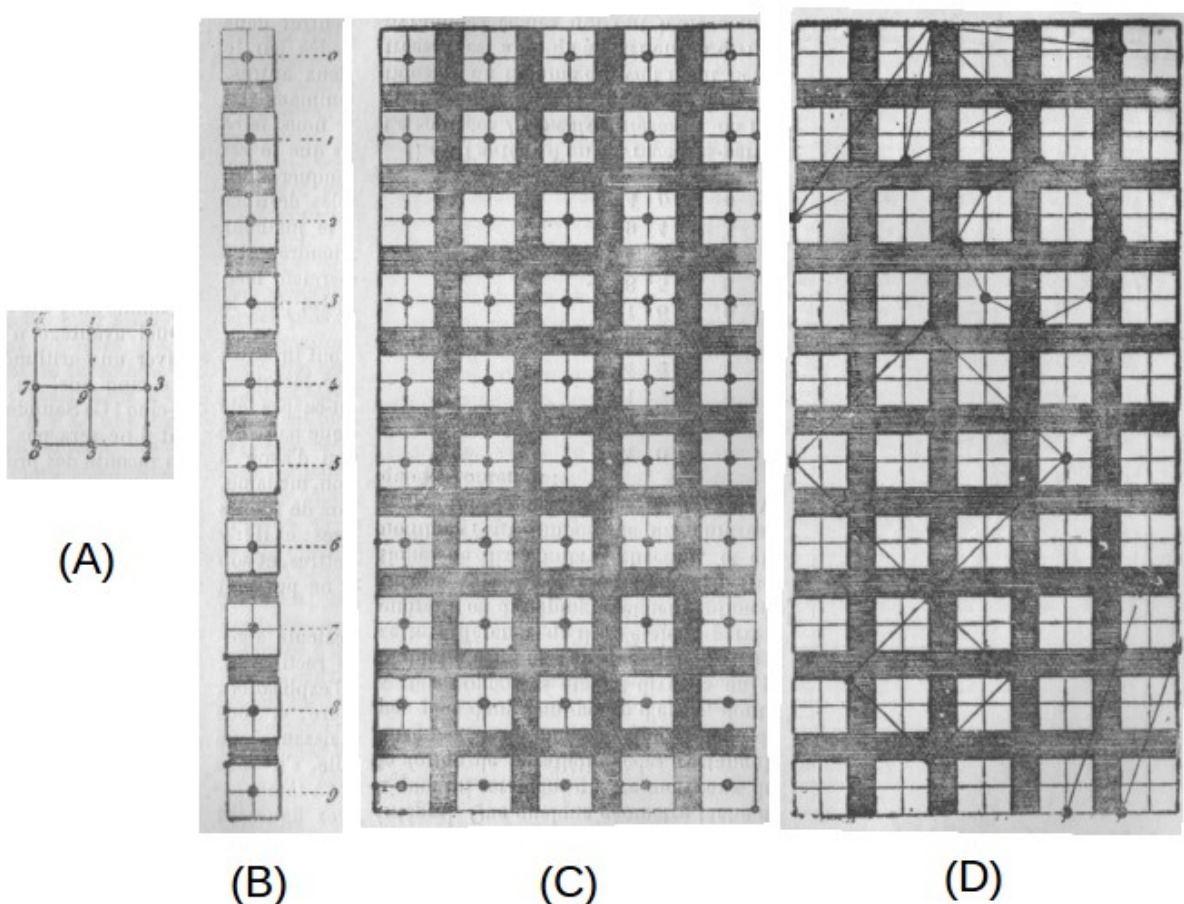


Figure 2.12. Le système de Saunderson d'après [Diderot 1901]. (A) les différentes positions qui peuvent prendre les épingles pour désigner un chiffre. (B) les numéros de 0 à 9 codés par ce système, (C) une succession de codes (D) des fils de soie qui achevaient de former des limites de ses formes géométrique.

Le principe du remplacement de la vision par le sens du toucher, par exemple dans le cas du Braille, a conduit les chercheurs en neurosciences à proposer comme généralisation le paradigme de substitution sensorielle : la perception d'une information perdue à cause de la déficience et/ou de l'occupation d'un sens peut être assurée par un autre sens de telle manière que les aires cérébrales associées qu'on souhaite suppléer sont activées d'une façon semblable à si elles avaient été sollicitées par ce sens. Les systèmes développés sur ce principe sont multiples, de même que leur domaine d'application. Souvent leur objectif est l'amélioration de la qualité de vie des personnes en besoin d'une information sensorielle manquante. Parfois, il s'agit de rajouter des informations sensorielles d'un type nouveau (par exemple un signal de position obtenu par GPS). Citons quelques exemples d'applications :

2.2.2.2. Premiers dispositifs de substitution sensorielle électrotactiles

« Il est possible d'envoyer des informations optiques au cortex visuel par un chemin autre que les yeux ». Ce challenge technologique a été relevé et remporté par plusieurs des groupes de chercheurs. En 1897, l'ophtalmologue polonais Noiszewski a initié ces travaux avec le développement d'un dispositif appelé l'*elektroftalmu* (que l'on appelle *œil artificiel* à cette époque). L'*elektroftalmu* permet le traitement de *stimuli* lumineux et leur transformation vers des stimulations auditives et tactiles. Noiszewski a écrit sur ce sujet environ 180 œuvres originales en polonais, russe, allemand et français. En 1963, Starkiewicz publie une étude qui traite de la stimulation sur la peau du front avec le même dispositif, l'*elektroftalmu*, composé de 80 canaux de stimulation vibro-tactile (Fig. 2.13).

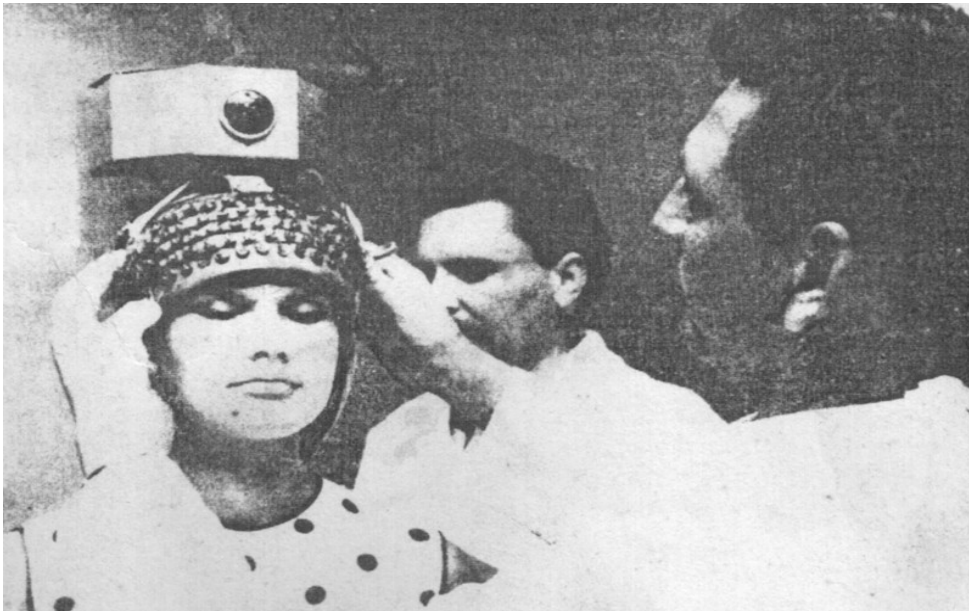


Figure 2.13. Le Dr W. Kuprianowicz (un des étudiants du professeur Starkiewicz) et le technicien F. Petruczenko lors de la dépose de l'*elektroftalmu* sur la tête du patient (1969)

D'autres études dans ce domaine ont été publiées par Samsó Dies, en particulier une étude dans laquelle un chien sourd retrouve une certaine perception "auditive" grâce à un dispositif de stimulation tactile [DIES, 1965] : le chien peut à nouveau utiliser les informations sonores de l'environnement.

Étant donné son travail important sur la substitution visuelle et les nombreux systèmes qu'il a créé depuis les années 1960, on peut raisonnablement dire que Paul Bach-y-Rita est le

fondateur de cette discipline (bien que d'autres y aient contribué à la même époque et avant). C'est en effet lui qui a produit le premier système visuo-tactile pour les personnes aveugles avec une matrice d'électrodes fixée sur l'abdomen (Fig. 2.14). En 1992, le laboratoire de communication tactile et la neuro-réhabilitation (Tactile communication and Neurorehabilitation Laboratory) a été fondé par l'équipe de Bach-y-Rita. Ses travaux sont aussi poursuivis par le laboratoire TIMC-IMAG,.

Le système de vision tactile (*Tactile Vision Substitution System* ou *TVSS*), mis au point par Bach-y-Rita, est un des premiers systèmes dans le domaine de la suppléance perceptive [Bach-y-Rita et al., 1969; Collins, 1970] (Fig. 2.14). Le TVSS transforme l'information visuelle acquise par une caméra vers une information tactile perceptive par les sujets. Un fauteuil de dentiste est équipé d'une matrice de stimulation tactile en deux dimensions de 400 (20x20) points (ou *taxels* pour *tactile pixels*). La matrice stimule ainsi le dos des sujets. Des expériences de perception visuelle ont été faites à l'aide de ce système: discrimination des orientations entre des lignes (suite linéaire de taxels actifs) horizontales, verticales ou diagonales, reconnaissance d'objets posés sur une table ou à proximité dans la pièce d'expérimentation. La caméra posée sur un trépied permet de changer l'angle d'acquisition et la position de la caméra par rapport aux objets, permettant ainsi d'explorer leurs formes. C'est le sujet lui-même qui bouge la position de la caméra avec ses mains. Le TVSS permet aussi de faire des zooms sur les objets.

Des versions portatives ou simplement plus petites ont été développées depuis. Différentes zones du corps sont stimulées, en particulier, comme nous le verrons dans un prochain chapitre avec le TDU, sur la langue qui est particulièrement sensible et discriminante.

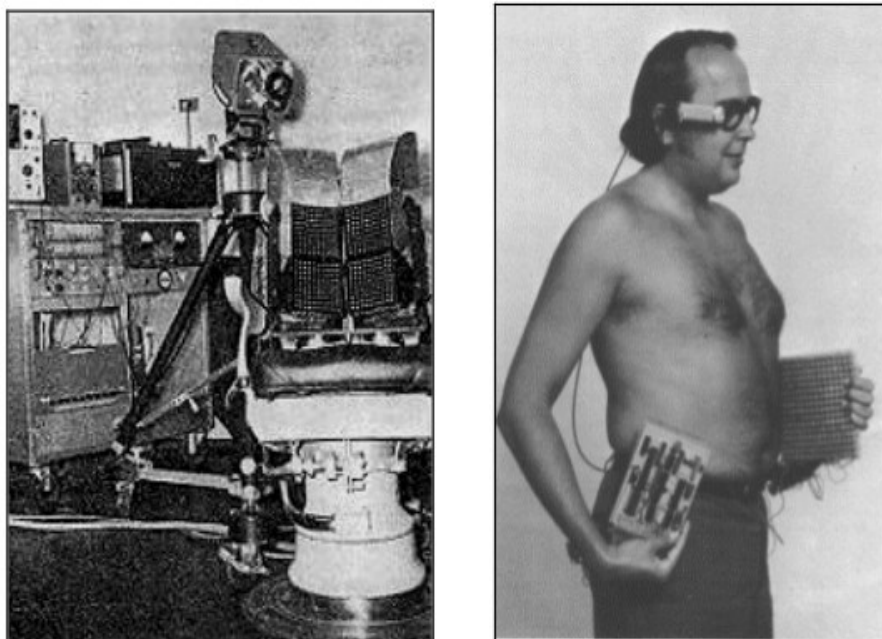


Figure 2.14. A gauche, l'un des premiers modèles TVSS. La caméra capte des images qui sont retransmises au sujet via une matrice située dans son dos. Ce dispositif était si encombrant qu'une pièce complète était nécessaire pour l'installer (d'après Bach-y-Rita, 1969). A droite, un sujet supposé aveugle (ici, il s'agit de P. Bach-y-Rita) utilisant une version portable du TVSS. La caméra est fixée sur des lunettes et les images sont retranscrites par stimulation électrotactile sur l'abdomen du sujet (d'après Bach-y-Rita, 1972)

2.2.2.3. Interactions avec du texte et des interfaces graphiques

Les dispositifs tactiles permettent à des utilisateurs de lire du texte depuis un écran tactile actualisable et obtenir du texte en format braille [Ramstein, 1996; Roberts et al., 2000; Yobas et al., 2003]. [Hughes and Forrest, 1996] proposent une souris tactile pour une manipulation directe des ordinateurs.

L'OPTACON [Goldish 1974] est l'un des premiers systèmes vibro-tactiles permettant la conversion du texte du format papier (livre, document) vers des lettres braille. Ce système est composé d'une caméra que l'utilisateur passe avec une main sur le texte et d'un système tactile utilisé avec l'autre main par laquelle le sujet perçoit les lettres braille (figure 2.15).

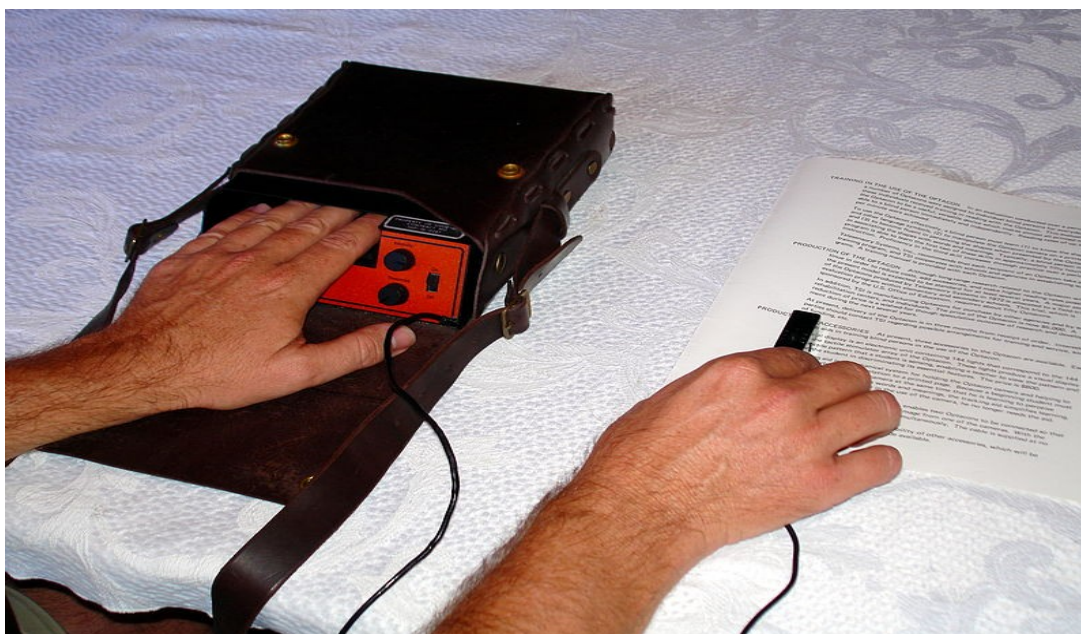


Figure 2.15. OPTACON : Dispositifs de lecture et conversion en braille [Goldish 1974].

La souris VirTouch en est une version ergonomique (Fig. 2.16). Elle contient trois interfaces tactiles de 32 épinglettes arrangées au format matriciel 4x8. Les épinglettes bougent verticalement en fonction des valeurs des pixels lors du balayage du curseur sur un écran (pixel par pixel). En utilisant ses trois doigts l'aveugle peut comprendre les formes affichées sur l'écran. L'encodage des informations d'image se fait selon 4 nuances de couleurs (blanc, gris, gris foncé et noir). Pour assurer l'interaction homme-machine, cette souris est équipée de six boutons pour donner des ordres d'exécution à l'ordinateur.



Figure 2.16. VirTouch Mouse (VTM) : une souris tactile qui intègre les fonctions d'affichage tactile à l'utilisateur ainsi que les fonctions d'une souris normale

Le système HyperBraille est introduit par [Kieninger, 1996] ; il donne accès au web pour les aveugles via une interface graphique adaptée avec des menus personnalisés avec un affichage en lettres braille. Le projet ACCESS [Petrie et al., 1996] est un autre exemple d'accès au *World Wide Web* dans lequel toutes les informations sont représentées en braille. D'autres projets comme *GUIB (Graphical User Interface for the Blind)* [Petrie et al., 1995] ou *Mercator* [Edwards and Mynatt, 1994] visent l'accès des aveugles aux interfaces graphiques à l'aide du son qui vient ici s'ajouter au sens tactile.

2.2.2.4. Applications éducatives et de divertissement

Dans un domaine proche, quelques applications logicielles ont été développées pour des objectifs d'éducation ou de divertissement. La majorité d'entre elles se fondent sur le principe de dessiner/effacer des images tactiles [Kobayashi and Watanabe, 2002]. Le montage montré dans la figure 2.17 représente une version expérimentale du *MIMIZU*. L'affichage tactile est assuré par une matrice vibro-tactile. Cette dernière récupère les données depuis une tablette. L'écriture est assurée par le mouvement du stylet sur la tablette.

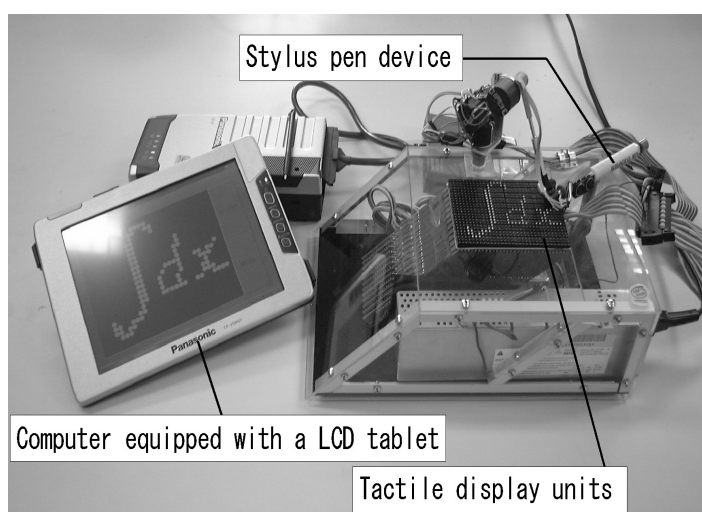


Figure 2.17. Le système MIMIZU est composé d'un stylet pour effectuer des opérations de dessin, d'un ordinateur équipé d'une tablette LCD et d'une matrice d'affichage tactile [Kobayashi 2002].

[Challis and Edwards, 2001] proposent un programme nommé *Weasel* qui permet aux aveugles de lire des notations musicales avec le toucher. Le principe nécessite bien entendu une connaissance préalable du solfège.

2.2.2.5. *Domaine médical*

On trouve plusieurs applications de la suppléance perceptive dans le domaine médical. Celles-ci se limitent souvent au stade de la recherche car il est difficile de les intégrer efficacement dans les protocoles (chirurgie, diagnostic ...) existants et car ils ne sont pas facilement acceptés (du fait de leur mode de fonctionnement notamment, mais aussi parce qu'ils consistent à donner davantage d'informations à quelqu'un (le médecin) qui en reçoit déjà beaucoup). Voici quelques-uns de ces dispositifs.

Un système de palpation qui capte l'information depuis l'intérieur du corps du patient et qui fournit un retour tactile au chirurgien en cours d'intervention a été développé par [Howe and Matsuoka, 1999]. Cette technique consiste à acquérir des informations tactiles depuis le corps du patient, plus précisément depuis la zone d'intervention. Sur son instrument chirurgical, le chirurgien dispose d'une sonde équipée de capteurs tactiles. Le flux d'informations tactiles provenant de la sonde stimule le bout du doigt du chirurgien. Cet outil de palpation permet une perception compensatrice au cours d'intervention chirurgicale. La

cohérence entre les variations de la force de retour et celles des vibrations de stimulation du doigt a été évaluée dans cette étude.

D'autres travaux concernent le guidage de gestes chirurgicaux de ponction. Il s'agit de fournir au chirurgien dont les sens usuels (vue, ouïe, toucher) sont saturés d'informations, d'autres informations (plus synthétiques) de guidage pour augmenter la qualité (précision et rapidité) de son geste. Au lieu de regarder son écran à chaque fois qu'il doit contrôler la trajectoire de son aiguille de ponction, des informations de guidage (gauche, droite, ...) lui sont envoyés par l'intermédiaire d'une matrice d'électrodes posée sur sa langue (Fig. 2.18). Les résultats obtenus sont probants : un utilisateur du TDU couplé au dispositif de ponction parvient sans entraînement à guider la pointe d'une aiguille de ponction sur une distance de 10 cm jusqu'à une cible (dans un *fantôme*) dans un tube d'erreur inférieur à 6 mm de diamètre, le plus souvent 2 mm [Robineau, 2009; Robineau et al., 2007].

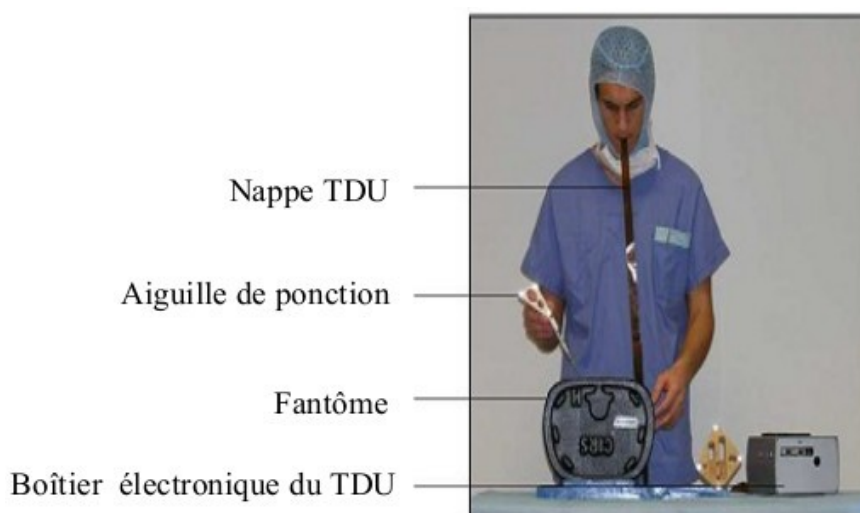


Figure 2.18. Le système de guidage proposé par Robineau et al 2007. Le chirurgien perçoit les instructions d'orientation de l'aiguille de ponction via la matrice de stimulation.

Une autre étude sur l'apprentissage de suivi de trajectoires et de guidage (d'un outil de ponction également) a été menée dans la même équipe GMCAO (Gestes Médicaux et Chirurgicaux Assistés par Ordinateur) du TIMC-IMAG. [Bluteau et al., 2010] ont évalué l'ajout d'un effet haptique sur le suivi de trajectoires. Leur protocole expérimental utilise le TDU comme un moyen de faire un retour sensoriel comparable à celui de la vision dans des tâches de suivi de trajectoires 2D et 3D [Bluteau, 2010; Bluteau et al., 2010].

Avec comme perspective la prévention des escarres chez le blessé médullaire paraplégique, [Chenu et al., 2009 et 2007] développent un système de suppléance perceptive (Fig. 2.19) fondé sur l'intégration de signaux provenant d'une nappe de pression (matrice de capteurs de pression) et un retour d'informations correspondant à des messages de changements de posture se faisant soit de façon visuelle (sur l'écran d'un moniteur ou d'un smartphone), soit via le TDU [Chenu et al, 2007 et 2009, Vuillerme et al 2007 et 2008] .

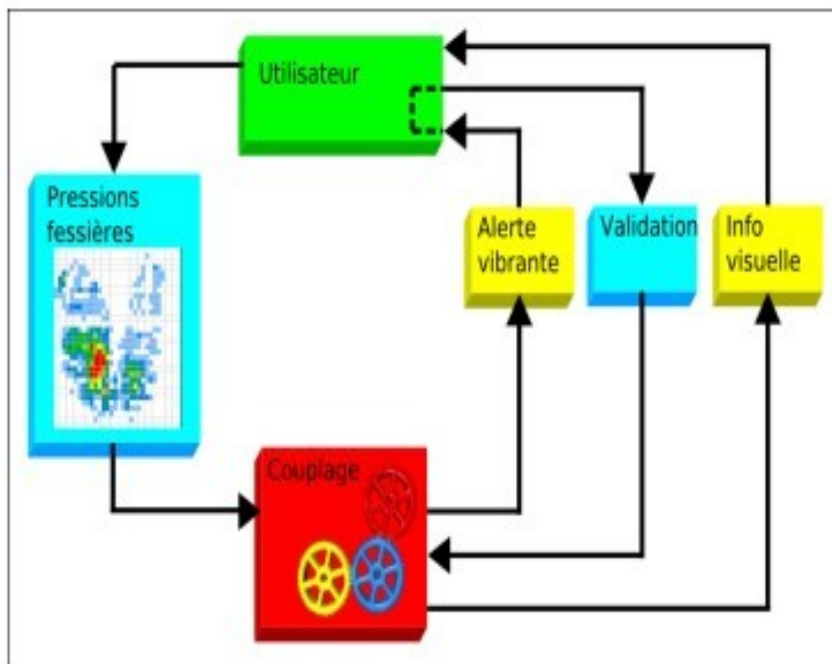


Figure 2.19. Schéma fonctionnel résumant le système de prévention des escarres [Chenu 2009]

2.2.2.6. *Domaine militaire*

Les technologies à base d'affichage tactile ont montré leur efficacité et, à défaut, leur utilité dans le domaine militaire. Le rapport de [Van Erp and Self, 2008] décrit l'état de l'art en ce qui concerne les affichages tactiles dans les environnements militaires. Les auteurs y détaillent les aspects neurophysiologiques, psychophysiques et les éventuels problèmes liés aux soldats, ainsi que les interactions homme/dispositif et les contraintes d'intégration. Van Erp et al terminent leur rapport par les travaux futurs avec ces technologies.

Soulignons déjà l'apport de ces systèmes pour la navigation sur terre, en mer ou même dans le domaine aéro- et astronautique. On mentionnera notamment la facilité d'interprétation

d'informations directionnelles (ex: situer une alerte comme la provenance d'un missile, ou garder un cap en situation de crise) sur des vêtements de type *Tactile Torso Display* (Fig. 2.20). L'information tactile sur le *TTD* est silencieuse et discrète ce qui représente un avantage.



Figure 2.20. Le système Tactile Torso Display proposé par Van Erp et al. pour les forces armées de l'OTAN

2.2.2.7. Commercialisation de l'unité d'affichage sur la langue (TDU Brainport) - restitution de la vision dans le cadre de la déficience visuelle

La société *BrainPort technologies* fait depuis plus de 10 ans des études très poussées dans le domaine de la substitution visuelle. Cette société a racheté les brevets déposés par P. Bach-y-Rita et en applique le même principe dans un dispositif compact disponible dans le commerce. Le dispositif est composé d'une caméra qui capte un flux vidéo et le restitue sous forme de signaux électriques sur la langue grâce à un TDU. Cela passe par une phase d'échantillonnage spatial (dégradation de la résolution) et de conversion des couleurs en niveaux de gris eux mêmes convertis après augmentation de leur contraste en niveaux de voltage. Cette société a réussi quelques 'coups médiatiques' avec un aveugle pratiquant l'escalade, le *Brainport* lui permettant de visualiser ses prises ; plus récemment avec un

vétéran britannique de la guerre d'Irak (Craig Lundberg) à qui le *Brainport* « permettait » de retrouver un peu de sa vision passée. Son intérêt majeur est de donner directement une image de l'environnement et non une interprétation (donc un guidage) comme c'est le cas dans les travaux d'O. Chenu, de F. Robineau et de J. Bluteau (Guidage du geste chirurgical ...).

2.3. Conclusion de l'état de l'art

Au cours de ce chapitre on a vu qu'une information tactile peut remplacer une information visuelle, les sensations tactiles pouvant être produites par des *stimuli* électriques ou vibrant, les capteurs à la surface de la peau et des muqueuses étant capable de percevoir ce type de *stimuli*. On a aussi exposé les limites organiques (et physiologiques) imposées à et par ces dispositifs (ex: nécessité des mouvements oculaires dans la vision).

Le paradigme de substitution sensorielle est valable (et applicable) mais n'est pas complet ! Il est vrai que l'on peut partiellement remplacer une information visuelle à l'aide du toucher et, plus généralement, de sens vibro-tactiles, mais principalement avec des objets que l'on connaît déjà (que l'on a appris à reconnaître) et que l'on peut déplacer. Les travaux réalisés dans ce paradigme offrent une certaine autonomie à la personne handicapée, mais on trouve dans la littérature des témoignages de personnes aveugles ou malvoyantes frustrées par la qualité frustrée des dispositifs de suppléance perceptive, en particulier visuelle.

La vision est associée au verbe *voir*. La vision représente une modalité de perception complexe. Quand on regarde une scène de la vie à travers nos yeux, on perçoit des images à la fois représentatives et significatives dans notre esprit. Cette perception, nommée *visuelle*, vient-elle du fait de *regarder* ou de *voir* ? Cette question est tout à fait légitime dans le cadre de la substitution visuelle. Dans le fameux livre sur la vision de [Rodieck 2003], on trouve deux chapitres intitulés *Voir* et *Regarder* qui séparent ces deux concepts. Dans le chapitre *Regarder* l'auteur présente le système visuel dans le sens de la stabilisation de l'image sur la rétine et de la cartographie du monde extérieur. Pour cela, la géométrie du regard entre en jeu, ce qui inclut les mouvements de la tête et la mise en œuvre des muscles oculaires. Par contre, dans le chapitre *Voir*, les termes utilisés sont orientés vers les composants de l'œil interne :

photorécepteurs, cellules ganglionnaires, et la sommation de l'information vis a vis du champ récepteur. Voir et regarder s'accomplissent de concert afin d'assurer la tâche de perception visuelle. *Regarder* correspond davantage au processus externe de cette tâche et *voir* concerne le traitement de l'information lumineuse (puis électrique dans les neurones) à l'intérieur du système visuel.

Avec l'utilisation des prothèses (vision prothétique dont rétine artificielle, ...) et des dispositifs dédiés à la suppléance visuelle, la restitution des images acquises à partir de caméras se fait par stimulation d'autres organes et/ou passent par des chemins inhabituels. De plus, le dispositif ne couvre que partiellement l'étendue des capacités de l'organe. Ainsi, si on prend le cas de la vision, le paradigme de substitution ne couvre que le concept *Voir*, puisque le regard (mouvements volontaires et involontaires, focalisation ...) n'est pas pris en charge par l'ensemble des dispositifs (prothèses et systèmes de suppléance). Par-dessus tout, même voir ne se fait pas correctement, puisque la résolution (et la qualité) du signal image offerte par ces dispositifs est infiniment éloignée de celle des organes. Il reste donc encore du travail, au-delà de l'augmentation de la résolution des dispositifs et de la qualité de leur implantation. En particulier, il faut maintenant travailler sur la nature du signal fourni (une restitution intelligente des images provenant de l'environnement) et sur les traitements de ces signaux pour le rendre plus perceptibles.

Dans une interface individuelle avec l'environnement, qu'il s'agisse de substitution sensorielle ou de réalité augmentée (ajout d'informations), la modalité tactile est un choix pertinent et viable. Toutefois, la mise en place d'une certaine interface tactile qui peut s'avérer appropriée sur ces endroits du corps pour une certaine application, peut ne pas l'être pour une autre. Par exemple, la main est particulièrement adaptée pour le placement d'une interface tactile car c'est le plus sensible et le discriminant des organes en ce qui concerne le toucher. Cependant, dans de nombreuses situations, la main est occupée ailleurs ; ce n'est donc pas une situation acceptable pour une interface tactile. Ceci conduit les chercheurs à se tourner vers d'autres endroits du corps, comme la langue, le palais, le front ou la voûte plantaire, pour la mise en place d'interfaces tactiles.

Enfin, comme un certain nombre de facteurs liés au dispositif peuvent influencer sur la sensibilité de la perception tactile (changement de l'épaisseur et de la largeur des vibreurs ou des électrodes, emplacement sur le corps, amplitude, fréquence et forme du signal de

stimulation ...) d'autres recherches sont faites pour identifier ces effets et leur importance. Une fois les limites physiologiques d'utilisation définies, les chercheurs seront en mesure de traduire ces connaissances sous la forme d'adaptation des composants (et de leur comportement) dans les dispositifs prothétiques ou de suppléance. Ceci permettra alors peut-être à terme une utilisation enfin effective des interfaces tactiles (ce qui n'est pas vraiment le cas actuellement).

Nous allons maintenant, dans la prochaine partie, entrer dans le vif de cette thèse en nous intéressant au cas spécifique de la substitution visuelle. Au cours d'une série d'expérience, nous montrerons que plusieurs adaptations sont nécessaires pour compléter, dans le cas de la vision, le paradigme de substitution visuelle. Nous allons donc y ajouter des éléments du regard inspirés du mouvement oculaire.

Nous montrerons qu'au cours de ces expériences, plusieurs effets pervers se manifestent, effets comme l'image miroir, la surcharge d'informations (surcharge cognitive) et les contradictions d'information, qui sont observés dans d'autres contextes comme par exemple la téléphonie au volant, les contradictions visuo-auditives, ...

II – Matériel et méthodes

3 Plate-forme expérimentale

Dans ce chapitre on va présenter l'architecture logicielle et matérielle adoptée pour la réalisation de nos expériences. Ceci est illustré de quelques captures d'écran de l'interface utilisateur ainsi que les outils intégrés. L'objectif escompté dès le début était la réalisation d'une plate-forme expérimentale qui soit flexible, ergonomique et bien sûr fonctionnelle. Cette plate-forme expérimentale est composée d'un ensemble de composants software (programmes informatique, bibliothèques logicielles, frameworks) et hardware (outils d'acquisition, dispositifs de stimulation visuels et électrotactile).

3.1. Fonctionnalités

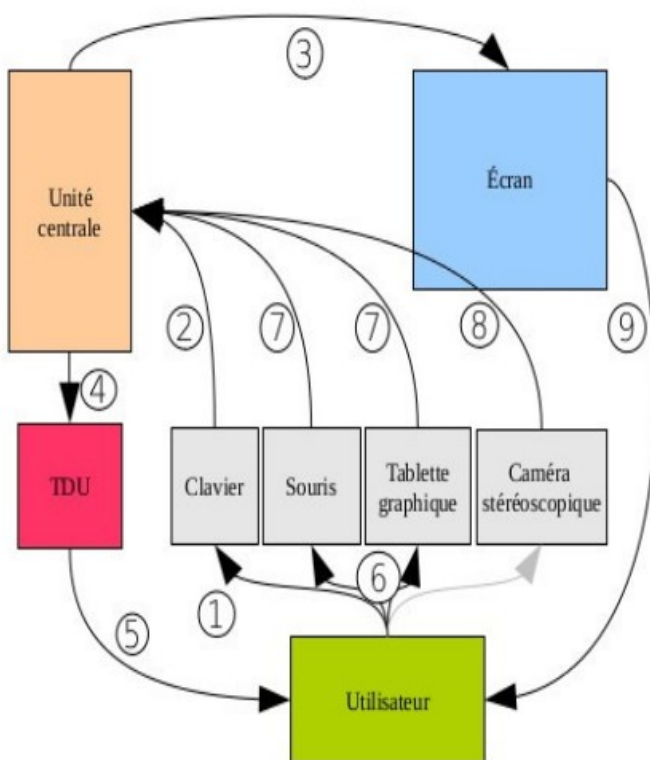
Notre plate-forme expérimentale assure un ensemble de fonctions afin d'accomplir les scénarios de nos expériences.

De manière générale, voici le déroulement du fonctionnement de l'application lors d'une expérience (fig. 3.1). Au départ, l'application connecte le TDU pour qu'elle puisse envoyer les stimulations à travers le port série sur lequel le TDU est connecté. Les paramètres de connexion et les opérations de réglages sont dans un module séparé. Pour assurer le bon acheminement des signaux vers la langue des sujets, une méthode spécifique se charge de l'envoi en prenant la matrice de données en argument. Cette matrice représente les voltages de chaque pixel tactile. Elle est calculée à partir des données images (matrice des intensités lumineuses comprises entre 0.0 et 1.0), de la matrice de calibrage et d'effets comme la sommation de plusieurs tampons d'images (*motion blur*), des translations aléatoires (saccades) et des fluctuations aléatoires de l'intensité lumineuse (scintillement). L'appel à cette méthode se fait d'une façon continue durant l'expérience à partir du lancement de l'expérience jusqu'à la fin (soit fin de durée d'une expérience, soit réponse acquise).

Les utilisateurs (expérimentateurs) peuvent contrôler l'ensemble des opérations (connecter/déconnecter le TDU, choisir un périphérique d'entrée ou de sortie, lancer une expérience, répondre à une question, modifier les paramètres de telle ou telle expérience, saisir les données utilisateur). On distingue les cas d'utilisation suivants :

- *Connecter le TDU.* Cela fait appel à toutes les fonctions de chargement et d'initialisation du TDU.
- *Changer le type de TDU.* Le TDU, dont notre exemplaire a été fabriqué en 1995, est livré avec une bibliothèque logicielle (une *DLL* non documentée (!) et dont tout le fonctionnement a été déduit du fichier d'entête). Cette bibliothèque permet de contrôler 3 types de TDU: dummy, fixed, wireless. Dans le premier type, aucune information n'est envoyée vers le TDU, mais toutes ses fonctionnalités sont simulées. Le type fixe correspond au TDU filaire connecté par le port série (RS232). Le dernier type correspond au mode sans fil (il y a au laboratoire un TDU sans fil mais son autonomie est limitée).
- *Gestion des sujets:* ajout des nouveaux sujets avec leurs informations personnelles (nom, age, sexe, profession, ...). Un identifiant est attribué à chaque sujet pour garder l'anonymat pendant l'expérience et le traitement des données. Ainsi, les fichiers de données ne contiennent pas le nom des individus. Seul l'expérimentateur peut faire le lien entre les identifiants et les noms des sujets.
- *Calibrer le TDU.* Cela permet de réaliser le calibrage (voir le détail dans le chapitre 4) des matrices de voltage de façon à uniformiser du mieux possible le signal perçu sur toute la surface de la langue au contact avec le TDU. Le calibrage est spécifique pour chaque sujet. Il est recommencé pour chaque sujet avant chaque expérience à partir des matrices de voltage préalablement déjà calibrées. Cette étape peut en effet s'avérer longue et affecter la sensibilité de la langue.
- *Choisir et réaliser une expérience.* Contient les fonctions de création d'un objet de type *expérience* avec tout ses attributs (nom du sujet, date, réponse attendue, paramètres de simulation, ...). Le lancement d'une mesure passe par une initialisation d'un compteur chronométrique (pas de temps 100 ms) pour connaître les délais de réponse dans les expériences. Certaines expériences ont des durées prédéfinies (ex: le suivi de trajectoire dure 30 secondes).
- *Modifier des paramètres d'une expérience.* Permet de tester l'influence de l'intensité de certains paramètres (scintillement, saccades, atténuation de luminosité, contradiction ...) sur les réponses obtenues à chaque expérience.

- *Donner une réponse.* Cela se fait en utilisant différents outils d'entrée, c'est-à-dire des périphériques comme le clavier, la souris, un stylet avec sa tablette graphique, ou encore via la caméra stéréoscopique (l'acquisition de points de référence sur une cible réfléchissante se fait via le contrôle par une autre DLL mise au point au laboratoire). Le mode de réponse peut être choisi dans certaines expériences. La réponse une fois acquise (la fin d'acquisition diffère selon les expériences) est enregistrée dans un fichier contenant tous les résultats pour ce type d'expérience avec les paramètres utilisés pour l'expérience et les données sur les sujets.



- 1.L'utilisateur lance une opération (début d'expérience, calibrage, ...) à partir du clavier.
- 2.Le clavier transfère la commande vers l'unité centrale.
- 3.Affichage des résultats de l'opération.
- 4.Transmet l'information vers le TDU.
- 5.Application des stimulations sur la langue de l'utilisateur.
- 6.Réponse de l'utilisateur via la souris, la tablette graphique .
- 7.Transfert de la réponse vers l'unité centrale.
- 8.Suivi par caméra stéréoscopique des réponses de l'utilisateur (nécessaire dans le cas 3D).
- 9.Visionnement des effets de ces manipulations ou stimulation dans le cas des tests en mode visuel.

Figure 3.1. Fonctionnalités générales de la plate-forme expérimentale

3.2. Diagramme de classe

Le diagramme de classe (Fig. 3.2) représente l'architecture interne et les relations entre l'ensemble des classes utilisées dans notre application.

On utilise ici un patron de conception (Design pattern) **Modèle Vue Contrôleur**, dans lequel les couches du modèle et les autres couches d'interaction avec l'utilisateur sont séparées. La liste des classes de notre modèle est la suivante :

- **Simulate** : cette classe permet d'assurer la stimulation en continu (un processus s'occupe de ça). Elle hérite de la classe **AV_Context** qui s'occupe du rendu graphique des scènes 2D/3D incluant les effets de saccades, de scintillement, d'ajustement de l'intensité lumineuse, de motion blur ...
- **User** : les utilisateurs de la plate-forme sont soit un expérimentateur, soit un sujet pour faire l'expérience. Cette classe est utilisée comme une abstraction des utilisateurs.
- **Subject** : le sujet est un élément majeur dans notre application. Ses propriétés (Sexe, age, profession) sont encapsulé dans la classe **Subject**. Les méthodes de cette classe permettent d'ajouter, supprimer, enregistrer des utilisateurs dans un fichier (ou une base de données pour le long terme), charger la liste des sujets sauvegardés.
- **Experimenter** : c'est celui qui gère l'expérience et assure le bon fonctionnement de la plate-forme. Pour l'instant c'est moi, Abdessalem Chekhchoukh, mais cela pourrait être d'autres personnes formées aux manipulations de cette application. Cela permet une certaine traçabilité des expériences.
- **Experience** : c'est le cœur de l'architecture étant donné que c'est cette classe qui associe les sujets avec les scénarios (tâches à faire) des expériences.
- **Scenario** : cette classe encapsule d'autres objets qui assurent le nombre de mesures a réalisé et le pattern de stimulation.
- **Stimulator** : classe abstraite héritant des fonctionnalités de stimulation de trois autres classes (actuellement) contrôlant divers systèmes de stimulation : stimulation Visuelle, Tactile ou auditive (**VDU** Vidéo Display Unit, **TDU** Tongue Display Unit, **ADU** Audio Display Unit).
- **Answer_Collector** : avec le même principe de classes abstraites, cette classe permet l'héritage vers des classes pour l'acquisition des réponses indépendamment des outils caméra stéréoscopique, souris, stylet, clavier. Une spécialisation est prévue en faisant un héritage de cette classe vers trois sous classes *Polaris*, *Mouse_Pad*, *Key_Board*.

- De plus, nous avons créé une classe nommée *TDU_Electrode_Matrix* qui est à la fois une matrice d'objets *DrawPanel*, dédiée au rendu visuel, et un objet de type *AVContext* dédiée au calcul de rendu graphique basse résolution (comme dans une matrice de 12x12 électrodes). La classe *TDU_Electrode_Matrix* encapsule toutes les fonctions de contrôle du TDU.
- *AVContext* est une classe jouant le rôle de pipeline graphique logiciel. Créé par N. Glade, elle est dédiée au TDU. Elle a été mise au point pour obtenir un meilleur rendu sur des périphériques d'affichage de très basse résolution (son rendu est meilleur qu'un échantillonnage basse résolution qu'on aurait fait à partir un rendu graphique standard, ex: *OpenGL*). Elle possède en outre tout un ensemble de fonctionnalités, dont la plupart travaillent en post-traitement (post-rasterisation et projection), permettant la réalisation facilitée de certaines expériences (scintillement, saccades, motion blur, atténuation exponentielle ou sigmoïdale du voltage en fonction de la profondeur...). Après le calcul des scènes, *AvContext* renvoie un frame buffer sous la forme d'une matrice float* au format souhaité (ici, nous récupérons uniquement la valeur de luminance (L du format couleur HLS), comprise entre 0.0 et 1.0, qui nous donne directement les facteurs multiplicateurs de la matrice de voltage (matrice de calibrage)). Cette matrice de luminance est passée directement en argument à la partie *DrawPanel*, pour l'affichage à l'écran, et à la partie TDUDLL pour la stimulation sur le TDU. Au-delà des fonctions spécialisées, l'intérêt de ce pipeline logiciel (donc ne dépendant pas des fonctionnalités des cartes graphiques) est de pouvoir fonctionner dans un système embarqué de type Arduino ou Raspberry Pi sans nécessiter de carte graphique et ainsi d'envisager la mise au point de dispositifs TDU portables.

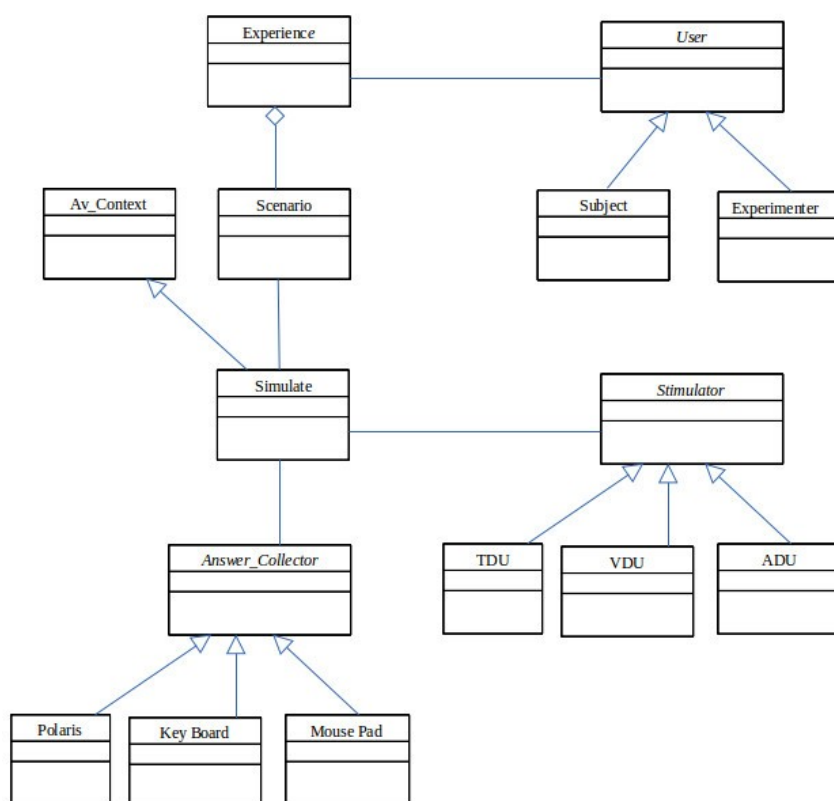


Figure 3.2. Diagramme de classe de la plate-forme logicielle expérimentale

Après avoir vu la conception du logiciel, passons à la partie réalisation qui décrit l'environnement du travail, les ressources utilisées et quelques captures d'écran qui montrent notre interface graphique.

3.3. Implémentation

Le travail a été principalement réalisé sur une station PC Intel sous Windows XP. La raison est que le TDU est contrôlé via une DLL. Nous avons cependant porté (réécrit) son code sous une forme multi plate-forme. Le langage de programmation choisi est le C++ en utilisant Code::Blocks comme environnement (IDE) de développement. Pour l'interface graphique on a fait appel aux widgets de Qt.

3.3.1. Bibliothèques Externes

- **Qt**: un framework orienté objet, pour la gestion de l'interface graphique (fenêtrage, boutons, zones de textes, etc).
- **Av_Context** : un pipeline graphique orienté objet fonctionnant sur un principe similaire à *OpenGL*.
- **TDUDLL** : une bibliothèque logicielle pour la gestion du TDU.

3.3.2. Ressources matérielles

Afin d'effectuer ces travaux dans les meilleures conditions, mes responsables des laboratoires TIMC-IMAG et AGIM ont mis à ma disposition tous les moyens que j'ai demandé, à savoir ceux que j'indique dans la liste suivante (voir la figure 3.4 montrant ces équipements en situation d'expérience de suivi de trajectoires 3D) :



Figure 3.3. (Gauche) Tongue Display Unit, Unité d'affichage sur la langue composé d'un boîtier électronique et une matrice d'électrostimulation comprenant 144 électrodes (12x12). (Droite) Tablette graphique et stylet pour tracer les réponses (lignes orientées). On distingue un abaque d'angles pouvant servir de référence pour les sujets.

- **Le stimulateur électro-tactile** ou **TDU** (Tongue Display Unit). C'est le stimulateur électrotactile de la langue. Le TDU comprend un boîtier électronique connecté au PC par le port série (RS232) et une nappe avec, au bout, la matrice d'électrodes que l'utilisateur doit poser sur le dessus de sa langue (Fig 3.3). La taille de la matrice que nous utilisons est de 3x3 cm et comprend 12x12 électrodes rondes de 1.2 mm de diamètre chacune et espacées de 1.2 mm. Le stimulateur électro-tactile de la langue avec une matrice de 144 électrodes est un appareil conçu dans le but d'une substitution

sensorielle de la vision par le sens tactile lingual tel que développé par le Pr Bach-y-Rita de l'Université de Wisconsin, USA. Il en existe très peu d'exemplaires : une copie de ce stimulateur a été développée à notre demande par K. Kaczmarek de l'Université de Wisconsin en 1998. Une autre copie (n'utilisant que des composants magnétiquement inertes et blindés) a été développée par la société ADEANE SARL par Jérôme Delachanal, sur la base d'un cahier des charges étudié par le Pr Christian Scheiber (avec qui nous collaborons) et le Pr Saint-Jalmes du laboratoire de Résonance Magnétique Nucléaire de l'Université Claude Bernard, spécialiste en électronique dédiée pour les applications RMN. Il existe d'autres versions du TDU : une version sans fil (wireless) avec une matrice de 6x6. Au cours de ma thèse, j'ai aussi participé au développement, avec N. Glade, d'une nouvelle version « haute résolution » (matrice de 3x3 cm comprenant 576 électrodes, soit 24x24) avec un partenaire industriel, la société SMEE (à Voiron). Au moment de l'écriture de ces lignes il nous reste à finaliser l'implémentation de l'application de contrôle (software) pour qu'il soit utilisable pour notre plate-forme. Sa mise au point, conséquente, doit aussi être faite.

- **La matrice d'électrode en or** qui est utilisée est celle brevetée par la société WICAB qui commercialise le Tongue Display Unit aux USA. La géométrie des électrodes a été étudiée par Kaczmarek et al [Kaczmarek, 2000]. De nombreuses configurations ont été proposées et la plus confortable est utilisée pour la version actuelle du TDU. Récemment Mitchell et al [Mitchell 2009] ont exploré les seuils de perception électro-tactile linguale et ils ressortent de leur étude une cartographie spatiale de la perception linguale permettant d'uniformiser la perception d'un signal, c'est-à-dire de percevoir les signaux de la même manière dans les différentes régions de la langue. La matrice d'électrodes est prévue pour des usages répétés. Une désinfection est effectuée avant et après chaque usage (voir 4.5. *Protocole d'hygiène*).
- **Le rendu visuel sur écran** est assuré par un outil de rendu graphique logiciel (AVContext), similaire à OpenGL. Il permet le calcul des scènes visuelles 2D ou 3D sur un écran d'ordinateur ou sur le TDU, en accord avec les faibles résolutions des TDU.
- **Les hauts parleurs** (ou des écouteurs) sont utilisés comme un moyen auditif pour donner aux sujets aveugles et mal-voyants la possibilité d'entendre les choix de

réponse. Dans certaines expériences, ils servent à émettre des « bips » indiquant le début et la fin d'une période d'acquisition.

- **Le bandeau noir** : On est amenés à utiliser un bandeau noir pour les yeux pour mettre les sujets dans l'incapacité de voir. Cela limite considérablement les problèmes de surcharge cognitive liés à la présence simultanée de signaux visuels et de stimulation linguale.
- **Le clavier** dont des touches spécifiques sont utilisées pour démarrer ou arrêter des expériences, y compris en cours d'expérience. Dans certaines expériences, les touches des flèches directionnelles servent à sélectionner ou modifier les coordonnées de la position actuelle (bouger à gauche ou à droite, en avant ou en arrière) comme dans une application vidéoludique.
- **La souris** est utilisée pour contrôler la plate-forme (expérimentateur) mais également pour que le sujet puisse calibrer sa matrice de voltages et enfin pour naviguer (changement des vues) dans certaines applications comme dans une application vidéoludique.
- **Le Joystick** est utilisé pour naviguer (changement des vues) dans certaines applications, comme dans une application vidéoludique.
- **La tablette graphique** est un des outils les plus utilisés dans les expériences (Fig. 3.3). Elle sert principalement à répondre à des questions comme l'angle que fait une ligne tracée en 2 dimensions. Le sujet utilise un stylet pour reproduire ces lignes en 2 dimensions. Elle a aussi été utilisée pour les expériences de suivis de trajectoires en 2 dimensions.
- **Le polaris** et ses **corps rigides** (ou **la Kinect** de Microsoft) est une caméra stéréoscopique qui permet d'acquérir les coordonnées 3D d'une ou plusieurs cibles (corps rigides tenus par la main de l'utilisateur ou fixés à l'utilisateur) dans l'espace par rapport à une référence (un autre corps rigide fixé au banc d'expériences). Les corps rigides sont des pièces en métal et en plastique possédant 3 réflecteurs de lumière infrarouge, utilisés fréquemment en chirurgie assistée par ordinateur.
- **Une station de travail fixe PC DELL Precision** équipé de :
 - Écran LCD 21 pouces.
 - Processeur Intel Core 2 Duo, CPU 2.33GHz.

- 2 Go de RAM et 100 Go de disque dur.
- Afin que notre plate-forme expérimentale soit aussi mobile nous avons aussi utilisé un **PC Portable (DELL Latitude E6400)**. Il a été utilisé pour réaliser des expériences dans d'autres contextes que le laboratoire, par exemple lors de déplacement chez des aveugles qui ont des difficultés à venir à la salle de manipulation au laboratoire. On note qu'on a utilisé des **adaptateurs USB-RS232** pour connecter le PC portable avec le TDU et le Polaris vu que leur connexion est en série (une technologie de bus qui n'est plus présente dans les nouveaux PC portables.)



Salle d'expérience (détails à droite)

Utilisateur avec le TDU sur la Polaris : on voit les 2 caméras langue, tenant la cible ducerclées d'émetteurs infrarouges Polaris. En bas à gauche, le corps rigide. Devant, le boîtier de contrôle du TDU et le PC.

Figure 3.4. Photographies prises pendant une expérience de suivi de trajectoires 3D devant la caméra stéréoscopique. A gauche une vue globale. Au milieu, des détails autour de l'utilisateur. A droite le Polaris.

3.4. Interface utilisateur (GUI)

L'interface utilisateur de la plate-forme a été faite en Anglais pour faciliter son utilisation future au sein des laboratoires AGIM et TIMC-IMAG (il y a beaucoup d'étudiants étrangers).

La figure 3.5 représente la fenêtre principale. Elle contient l'ensemble des paramètres liés au TDU (lancement, connexion, changement de mode, ...), et affiche une scène 3D (ici un cube vu de face) pour faciliter certains réglages.

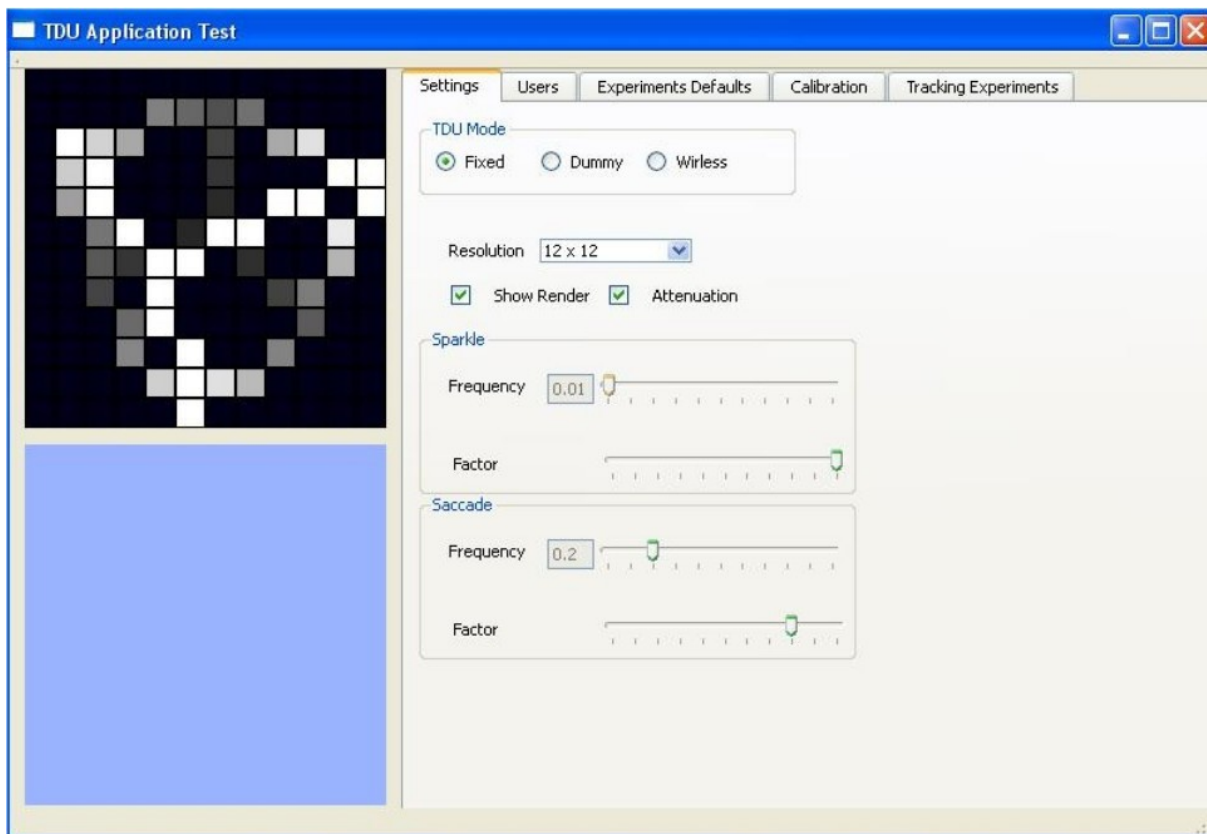


Figure 3.5. Fenêtre principale (connexion et réglages)

Dans l'onglet paramètres (*settings*) on peut modifier le type de TDU, la résolution du rendu graphique ainsi que des paramètres liés à l'affichage (activation, désactivation). Les réglages sont assez simplifiés dans ce logiciel et se contentent des paramètres qui nous servent dans le cadre de nos expériences. La plupart des autres paramètres ont été calibrés au cours de travaux antérieurs grâce à un autre logiciel (voir Fig. 3.6) permettant de régler l'intégralité des fonctionnalités qui nous intéressent (par exemple les paramètres de la fonction d'atténuation sigmoïdale de la luminosité en fonction de la profondeur, c'est-à-dire la position du point d'inflexion et la puissance de la courbure) mais que nous ne modifions pas au cours de nos expériences. Ici, l'intensité (facteur) et les fréquences des saccades et de scintillement sont ajustables par les deux réglettes en bas de la fenêtre.

La fenêtre comporte 2 contextes graphiques, un contexte DrawPanel dans lequel sont affichées les électrodes sous forme de carrés dont l'intensité lumineuse (allant du noir au blanc) représente le niveau de voltage, et un contexte OpenGL qui permet éventuellement d'afficher les scènes graphiques avec une haute résolution (voir Fig. 3.6), ou bien d'afficher les

directions saisies à la souris et au stylet sous forme de lignes. Le rendu graphique et la fenêtre OpenGL sont toujours présents dans notre interface et sous contrôle de boutons autorisant ou non l'affichage.

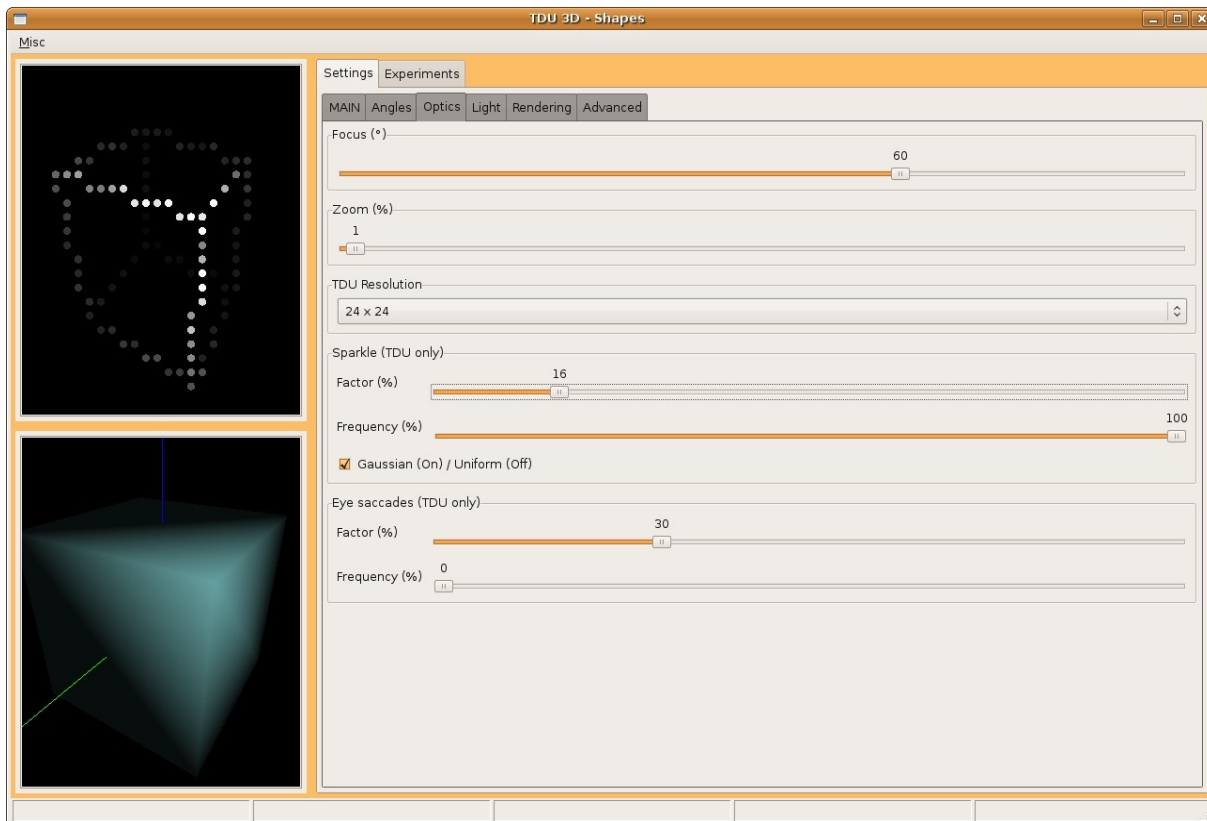


Figure 3.6. Logiciel de réglages fins et de test du TDU, de simulation et d'AVContext

Les paramètres de création et de modification des expériences sont dans un onglet indépendant. C'est ce qui figure sur la capture d'écran suivante (Fig. 3.7).

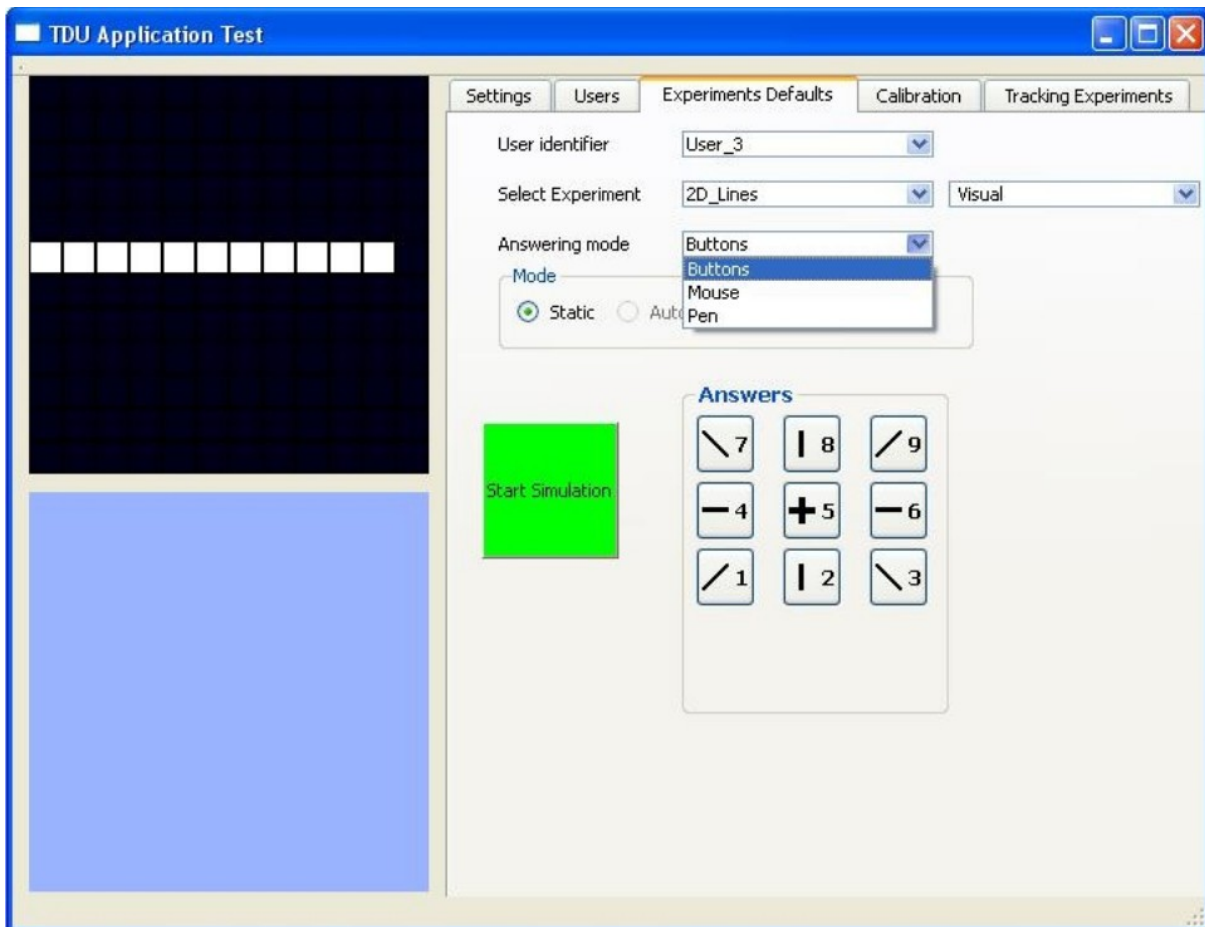


Figure 3.7. Interface graphique utilisateur qui permet de modifier les Paramètres liés aux expériences.

On y choisit notamment l'utilisateur qui subit l'expérience, et on y définit l'expérience en sélectionnant le type de formes 2D ou 3D à reconnaître, la modalité de rendu (mode visuel, mode TDU, ou bimodal : TDU+visuel), la possibilité ou non de contrôler la scène (statique ou dynamique), le mode de saisie (clavier, souris, stylet).

Pour les expériences qui contiennent des scènes dynamiques un onglet spécial (Tracking experiments) a été ajouté (Fig. 3.8). Ces expériences ont des paramètres très différents des expériences de mesure de sensibilité de l'orientation de formes 2D ou 3D. On y choisit la dimension dans laquelle la trajectoire est tracée (2D ou 3D), le numéro de la trajectoire (plusieurs trajectoires prédéfinies et déterministes utilisant des fonctions trigonométriques ont été définies), la vitesse de déplacement du point marquant la trajectoire.

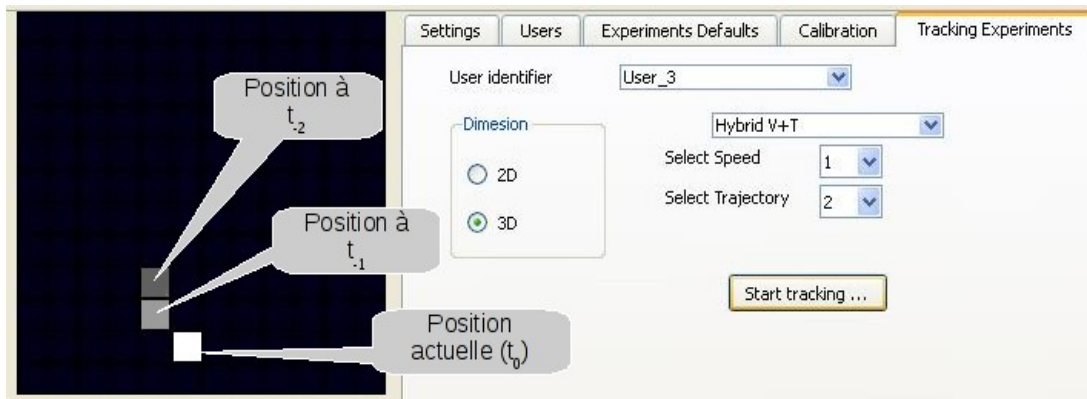


Figure 3.8. L'interface graphique montre l'expérience de suivi de trajectoires. On peut sélectionner des trajectoires différentes ainsi que la vitesse de déplacement des points. L'environnement 2D ou 3D est sélectionné avant le début de l'expérience.

3.5. Génération des effets affectant la qualité de perception

3.5.1. Effet de saccades

Comme nous l'avons décrit dans l'état de l'art, les micro-saccades oculaires sont des mouvements des yeux très rapides. Ce processus est nécessaire à la vision normale. Pour produire des saccades du signal image ressemblant aux saccades oculaires, nous utilisons une fonction qui provoque un déplacement aléatoire des points de la forme dessinée dans la scène à une distance régie par une loi de probabilité soit uniforme, soit gaussienne. Ces fonctionnalités sont un des traitements d'AVContext se situant soit (i) après l'étape de rasterisation (calcul de la position des pixels de l'image dans l'espace avant projection) et avant celle de projection, ceci permettant de faire trembler chaque objet de l'image indépendamment des autres, (ii) soit en fin de pipeline graphique, sur le buffer image (frame buffer), c'est-à-dire sur la totalité de l'image, ce qui est plus cohérent avec le mouvement de saccade oculaire qui provoque un déplacement uniforme de l'image entière sur la rétine.

Le principe est qu'avec une certaine fréquence f_s , l'image peut sauter, c'est-à-dire se trouver décalée par rapport à sa position d'origine. À chaque image calculée, on réalise un tirage aléatoire uniforme compris entre 0.0 et 1.0 ; si le chiffre tiré est inférieur ou égal à f_s , on réalise une saccade. Celle-ci est alors calculée comme un décalage de l'image par rapport à sa position d'origine. Nous avons utilisé une distribution gaussienne des sauts autour de la

position d'origine de l'image avec une moyenne centrée sur la position normale de la forme et un écart type ajustable permettant de quantifier la proportion de sauts inférieur à 1 pixel (donc pas de saut), celle de sauts compris entre 1 et 2 pixels, etc. Nous avons choisi un écart type assez faible, de manière à ce que 50% des décalages soient inférieur à 1 pixel (pas de saccade) et 50% environ soient supérieurs à 1 pixel (et en général inférieurs à 2 pixels). La fréquence d'actualisation du TDU étant de 10 Hz, si la fréquence de saccades est de 100% (tout le temps), cela signifie qu'une image sur deux subit une saccade, celle-ci étant donc ressentie à une fréquence de 5 Hz. La figure suivante (Fig. 3.9 gauche) illustre cet effet.

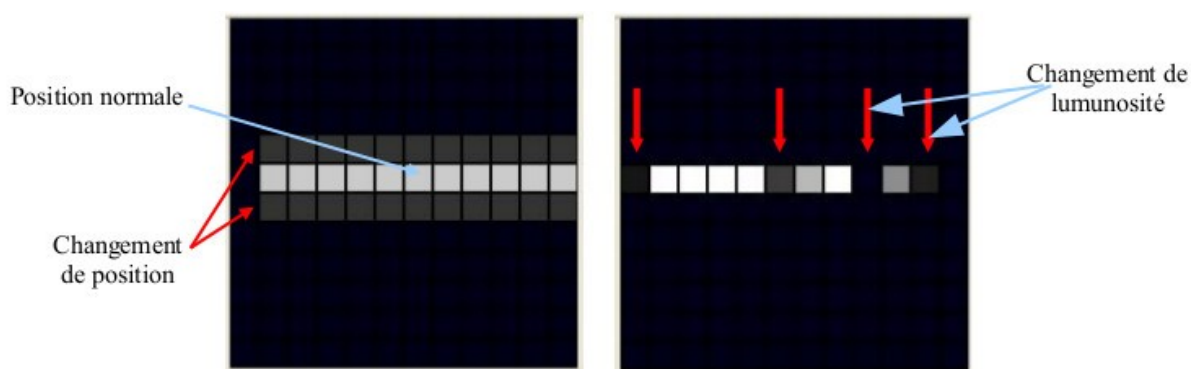


Figure 3.9. Effets de tremblements du signal électrique simulé dans un rendu graphique. (Gauche) Effet de saccade tel qu'il est perçu. L'intensité plus forte de la ligne centrale illustre le fait qu'elle est plus souvent affichée que les lignes latérales affichées 1 fois sur 2. (Droite) Effet de scintillement. Cette fois, c'est l'intensité des points qui fluctue autour d'une valeur proche de l'intensité normale du point.

3.5.2. Effet de scintillement (sparkle)

Le scintillement est un autre effet dynamique que nous utilisons pour générer un effet semblable à celui des saccades oculaires. Cet effet est local ; il est obtenu par modification de la luminosité de chaque point qui compose notre rendu graphique, comme cela est montré figure 3.9 à droite. Le processus est identique à celui des saccades, sauf qu'au lieu de faire trembler l'image par translation, on simule un tremblement en réalisant un tirage aléatoire (uniforme ou gaussien) modifiant l'intensité de chacun de ses points. Plus précisément, en chaque pixel de l'image, on évalue par un premier tirage aléatoire uniforme l'occurrence d'un scintillement. Si un scintillement doit se faire en un pixel, on effectue un deuxième tirage aléatoire (que nous avons choisi comme suivant une distribution gaussienne dont la moyenne est égale à l'intensité d'origine du pixel) qui détermine une fluctuation d'intensité (distance par

rapport à l'intensité d'origine), celle-ci ne pouvant dépasser 1. Ceci a pour but d'atténuer l'effet de fading tel que proposé par Zrenner et al. 2011 et Stingl et al. 2013.

3.5.3. Atténuation de l'intensité lumineuse en fonction de la profondeur

Pour encoder la profondeur dans une image en 3 dimensions projetée dans un périphérique d'affichage à 2 dimensions, classiquement un écran, on utilise d'une part la représentation perspective, d'autre part des artifices augmentant la sensation de distance avec les objets à mesure de leur éloignement. L'un de ces artifices est l'atténuation de la visibilité en fonction de la profondeur (éloignement), en particulier par une atténuation de la luminosité ou un brouillard (effet de fog). Dans le TDU qui est un périphérique de très faible résolution affichant des images sur un organe bien moins discriminant que la langue, nous ne savons pas encore si la représentation perspective est ressentie et aide à concevoir l'orientation en 3 dimensions d'un objet. Nous ne savons pas non plus si cela sera amélioré par l'augmentation de la résolution du TDU (TDU 24x24). En revanche, nous avons pensé, à juste titre, que l'atténuation de l'intensité lumineuse en fonction de la profondeur serait efficace pour cela. C'est en tout cas ce qui ressort des expériences préliminaires de suivis de trajectoires en 3 dimensions et de déplacement dans un environnement virtuel (jeu vidéo).

Ainsi, pour encoder la profondeur, nous affichons les objets les plus proches avec un maximum d'intensité (voltage élevé) et les objets les plus éloignés dont progressivement éteints. La fonction d'atténuation que nous avons choisi, devant une atténuation linéaire trop régulière et ne permettant pas une assez grande discrimination entre 2 positions d'éloignement différents et une atténuation exponentielle trop rapide, est une fonction sigmoïde (de Hill) dont on peut régler la position du point d'inflexion qui correspond ici à la position à laquelle 50% de l'intensité du signal est perdue, et la puissance de la sigmoïde qui indique à quelle vitesse l'intensité du signal diminue (la pente de la tangente à la courbe au niveau du point d'inflexion). L'usage de cette fonction permet au sujet de bien percevoir les objets proches et de vite comprendre l'éloignement de certains objets qui se situent à la limite du point d'inflexion.

3.5.4. Atténuation de l'intensité lumineuse en fonction du voisinage

La stimulation électrique d'une région de la langue par une électrode n'est pas aussi localisée que nous le souhaiterions. La cavité buccale est mouillée de salive et assez conductrice. Ainsi, lorsqu'on espère stimuler seulement une région de 1 à 2 mm de diamètre au contact avec une électrode, on stimule en fait une région beaucoup plus large pouvant atteindre plusieurs mm à 1 cm. Cela est d'autant plus ennuyeux lorsqu'on stimule la langue avec plusieurs électrodes voisines (par exemple une ligne d'électrodes). Dans ce cas, le ressenti est un cumul de voltages pouvant être élevé et rendre peu à peu la langue insensible à la perception de tout signal (Fig. 3.10).

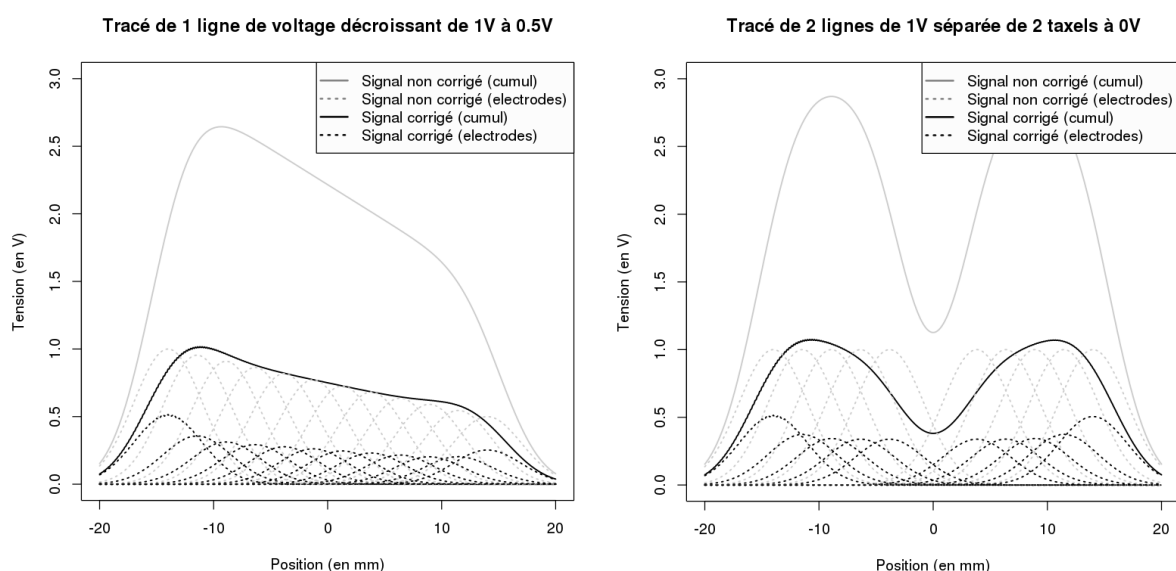


Figure 3.10. Simulation de la correction de l'intensité du signal (voltage) de chaque électrode en fonction de celle de son voisinage pour restituer un signal proche du signal souhaité. Les voltages non corrigés et corrigés de chaque électrode sont représentés par les signaux individuels gaussiens (représentant la diffusion de l'énergie électrique sur la langue) en pointillé. La somme des voltages perçus par l'ensemble des électrodes est représenté par les courbes en traits continus. Les signaux non corrigés sont en gris ; ils sont en noir après correction. On a ainsi (Gauche) le signal d'une ligne complète d'électrodes de voltage décroissant que l'on souhaite percevoir dans la plage 1V à 0.5V (et son signal non corrigé qui grimpe à 2.5V). De même (Droite) le signal apparent de 2 lignes de voltage constant à 1V séparées par 2 taxels mis à 0V (au milieu).

Il est possible de remédier à cela (Fig. 3.10) en corrigeant l'intensité du signal fournie par chaque électrode en fonction de son voisinage (c'est-à-dire de l'intensité du signal fournie par les électrodes voisines par une renormalisation). Bien que nous n'ayons ni implémenté ni expérimenté ce traitement, nous pensons qu'il serait très efficace pour améliorer sensiblement

la qualité du signal perçu. Le contrôle qu'aurait l'expérimentateur (le programmeur) sur l'intensité restituée souhaitée serait notablement meilleur.

3.5.5. Motion blur - Rémanence

Afin d'augmenter la perception de mouvement de certains objets affichés dans le champ de vision TDU, nous avons implémenté une fonction permettant le motion blur (le flou de mouvement). Il s'agit simplement de la possibilité de réserver plusieurs tampons d'image (profondeur du tampon d'accumulation) et d'effectuer une sommation de ces tampons avec une atténuation de l'intensité des tampons les plus anciens. Cela permet l'affichage de 'objets avec une trace rémanente (Fig. 3.8). Cependant, à part pour les représentations de points mobiles (trajectoires) sur le moniteur (pour que le sujet comprenne bien le sens de déplacement du point), nous ne l'avons pas appliqué au TDU pendant les expériences car nous avons vite constaté que la stimulation sur la langue provoquait d'elle-même un effet de rémanence.

4 Protocole Expérimental

4.1. Le processus de calibrage

La langue est un organe musculo-membraneux situé dans la cavité buccale, dont les rôles principaux sont la phonation, la déglutition et la mastication. Certaines caractéristiques de la langue en ont fait l'organe vers lequel Bach-y-Rita s'est naturellement dirigé en développant le TDU (pour Tongue Display Unit ou Unité d'affichage linguale). En effet la langue, en tant que support de stimulation, possède plusieurs avantages :

- Une grande sensibilité : ses récepteurs sont proches de la surface de la langue ; en outre ils sont à l'abri des agressions extérieures.
- Une forte discrimination spatiale avec une distinction spatiale de signaux de l'ordre d'1 mm.
- Une bonne représentation corticale (de l'ordre d'un centimètre carré) avec une grande précision de l'innervation (mapping).
- Sa mobilité : elle lui permet d'explorer librement les stimulations (bien que ce ne soit pas notre objectif).

La sensibilité de la langue est variable. C'est ce qu'on ressent naturellement en s'alimentant (on localise les aliments et leur saveurs finement au niveau de son extrémité pointue, l'apex), ou en mâchonnant un stylo (on en ressent l'extrémité plus précisément avec l'apex qu'à l'arrière de la langue). Cette observation confirme les études faites précédemment par (Robineau, 2008). La figure 5.1 représente les seuils de sensibilité électro-tactile sur toute la surface de la langue. Sa sensibilité est de plus en plus basse dans le sens antéro-postérieur et le sillon médian possède un seuil plus bas que les bords de la langue. De plus, la partie latérale droite apparaît légèrement plus sensible que la gauche. Ces résultats viennent des travaux de F. Robineau (Robineau, 2008) dont les résultats sont conformes à ceux de Marlow et al. 1965, Ringel & Ewanowski 1965, Grossman 1964, Lass et al. 1972, Maeyama 1989 et Trulsson & Essick 1997.

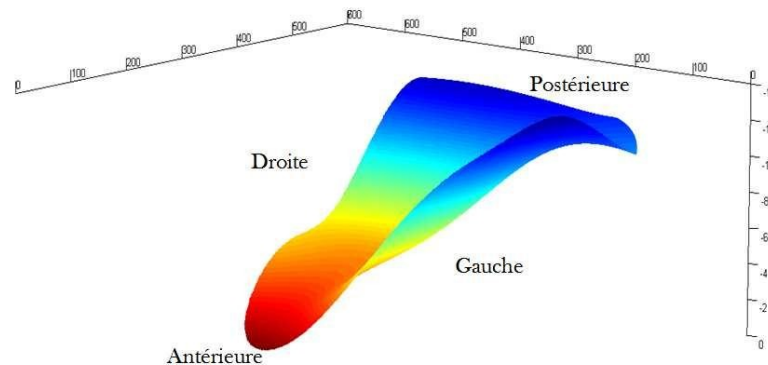


Figure 4.1. Sensibilité de la langue : représentée en niveaux de seuil de confort. La sensibilité est plus basse vers l'arrière et elle augmente vers l'avant (l'apex) [Robineau 2009]

L'utilisation du dispositif d'électrostimulation linguale n'est pas exempte de difficultés. La principale vient notamment de cette sensibilité spatialement différente de la langue. Le ressenti d'un même signal (d'un même voltage) ne sera pas le même selon la zone de stimulation. Ce signal doit donc être corrigé selon l'endroit de la stimulation. Par exemple, 1 volt pourra suffire à l'apex tandis qu'il faudra près de 10 volts pour un même ressenti deux centimètres en arrière. On réalise donc une étape de calibrage afin que l'utilisateur ressente le même stimulus de partout. Pour cela, on génère une matrice de voltages compris entre 0 et 15 volts. C'est le produit terme à terme de cette matrice avec une matrice de signaux compris entre 0 et 1 qui constitue l'information envoyée à la matrice d'électrostimulation linguale.

Le protocole des expériences est décrit ci-dessous. Chaque expérience commence par le calibrage des voltages appliqués à la matrice d'électrodes pour chaque sujet.

La sensibilité de la langue aux voltages n'est pas la même pour chacun des sujets et n'est pas la même pour une même personne en divers endroits de sa langue. Cela implique qu'un même signal (des électrodes activées avec un certain voltage) ne sera pas ressenti de la même manière (avec la même intensité) par une même personne selon la zone de la langue qui reçoit ce signal. Il ne sera pas non plus ressenti de la même manière par deux personnes différentes. Le processus de calibrage se fait comme suit:

1. Initialisation d'une matrice de calibrage temporaire construite sur la base d'un gradient antéro-postérieur de voltages. Le voltage est plus faible devant (1 volt) et plus fort en arrière

de la langue (10 volts), conformément à ce que l'on sait de la sensibilité de la langue (Fig. 4.1).

2. Affichage des voltages de la matrice de calibration pour chaque électrode selon une disposition en matrice (0,5 volts minimum, 15 volts maximum). Pour réaliser l'opération de calibration on a ajouté des boutons pour augmenter ou diminuer les valeurs de la matrice en chaque point, soit en cliquant sur l'électrode elle-même avec la souris ou le stylet graphique (accroissement avec le bouton gauche, diminution avec le bouton droit), soit sur un des boutons qui agissent sur toute la matrice, sur une ligne ou sur une colonne entière.

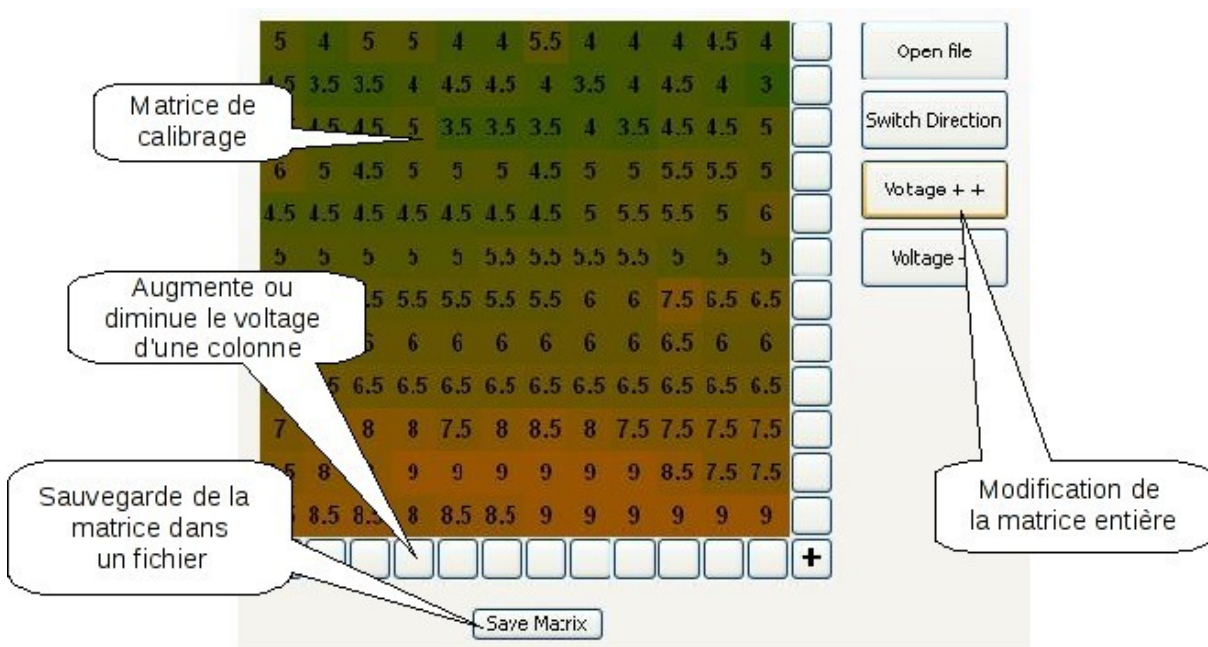


Figure 4.2. Interface graphique dédiée à la phase de calibration, ce pad matriciel de taille 12x12 représente les pixels tactiles, les valeurs affichées ainsi que les couleurs associées indiquent le voltage qui passe par un taxel.

3. L'utilisateur teste le ressenti : on trace une ligne complète d'électrodes activées qui se déplace par aller-retours successifs à vitesse constante, horizontalement ou verticalement (choix que fait le sujet) (Fig. 4.2).

4. Quand l'utilisateur pense ressentir la ligne sur toute la surface de sa langue avec la même intensité et confortablement, il sauvegarde cette matrice dans un fichier qui porte son nom (Fig. 4.3). Sinon il reprend le calibration jusqu'à satisfaction. L'enregistrement des matrices de calibration permet leur réutilisation pour de futures manipulations. On propose notamment aux

sujets de vérifier et éventuellement modifier leur matrice de calibrage à chaque expérience à laquelle ils participent.

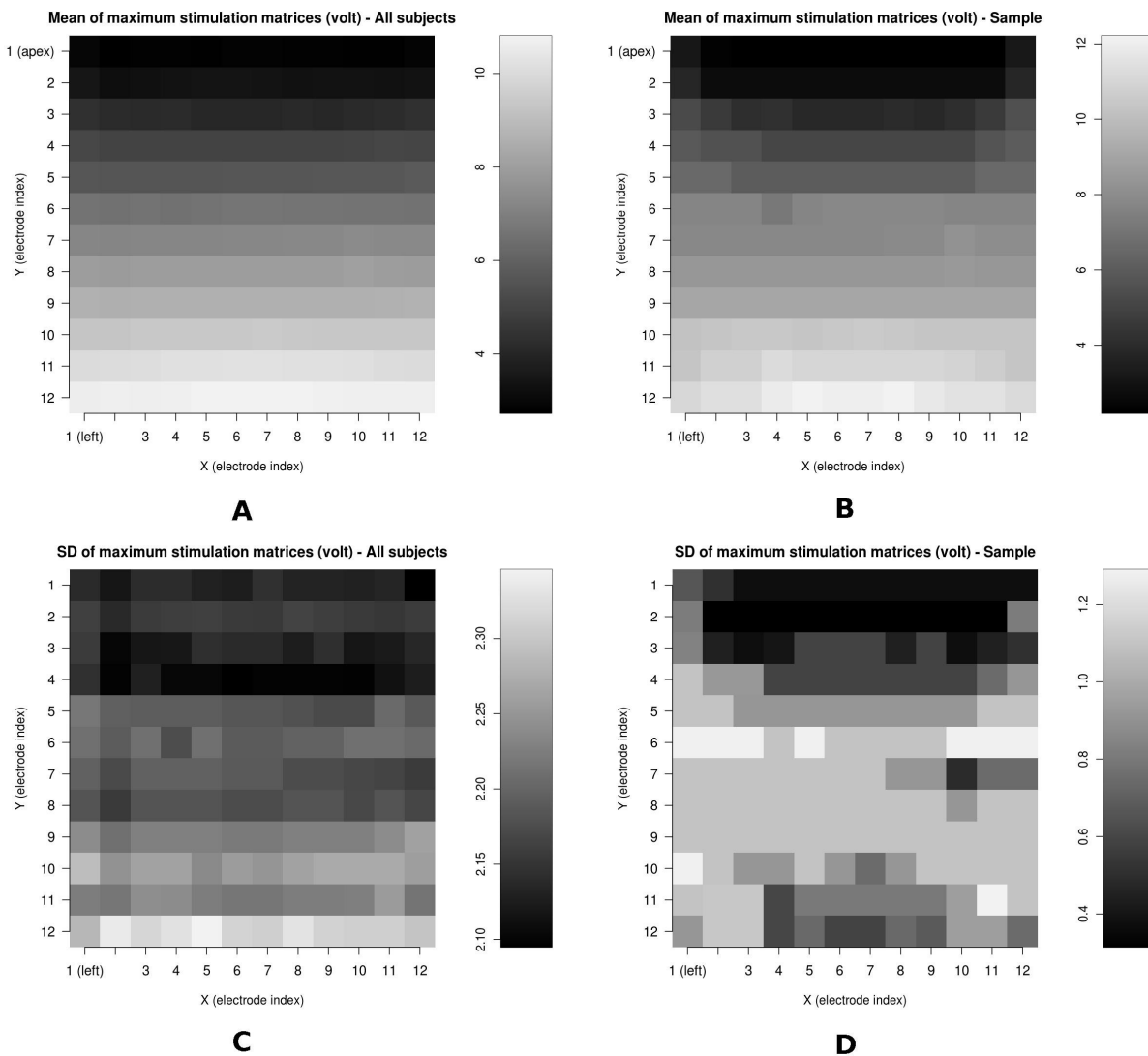


Figure 4.3. Résultats du processus de calibrage : (A) matrice de calibrage moyenne de tous les sujets (B) matrice de calibrage moyenne d'un sujet type (C) matrice de déviation standard de tous les sujets ; 4 matrices par sujet (D) matrice de déviation standard d'un sujet type (sur 4 matrices).

L'efficacité du calibrage est difficile à évaluer. Néanmoins, on s'aperçoit qu'un même sujet calibre à chaque fois sa matrice de manière assez semblable. De plus, dans un questionnaire, nous avons demandé aux sujets de juger l'efficacité du processus de calibrage et l'ergonomie générale du dispositif expérimental (Fig. 4.4). La plupart des sujets ont jugé le

processus de calibrage comme très efficace et l'ensemble du dispositif comme moyennement ergonomique pour la moitié d'entre eux, très ergonomique pour l'autre moitié.

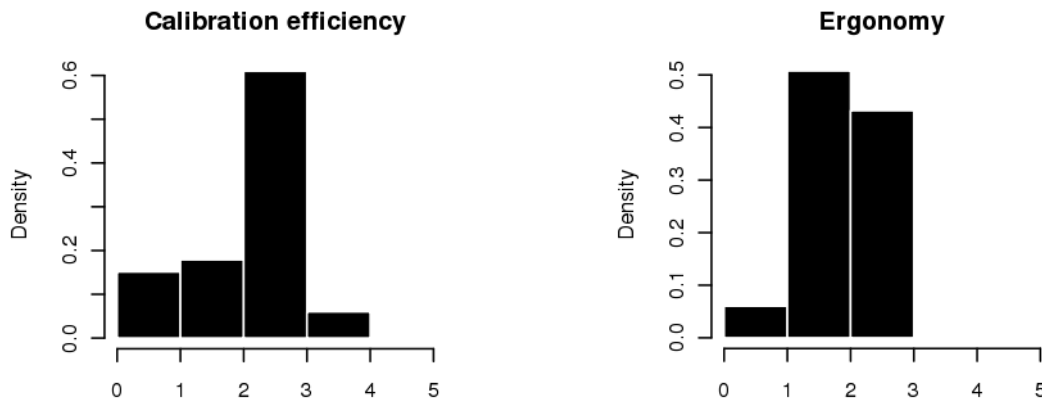


Figure 4.4. Résultats du questionnaire sur l'efficacité (subjective) du processus de calibrage et sur l'ergonomie générale du dispositif expérimental. Les notes vont de 0 (très mauvais) à 4 (excellent).

4.2. Familiarisation

Avant toute expérience, les sujets passent par une étape de familiarisation avec le concept de substitution sensorielle notamment et avec le fonctionnement de l'expérience. On explique au sujet le principe de la substitution sensorielle, le but de l'expérience qu'il va subir, son déroulement, à commencer par l'étape de calibrage. On l'avertit aussi sur les effets ressentis comme la sensation de l'activation électrique sur la langue et l'augmentation passagère de la salivation. On lui donne enfin les instructions nécessaires à l'expérience comme la tâche à réaliser, la nécessité de répondre le plus vite possible dans certaines expériences, etc.

Une fois le calibrage fait, on réalise quelques points blancs (une dizaine) jusqu'à ce que le sujet soit suffisamment familier de l'expérience. Les quelques mesures qui sont faites alors ne sont pas prises en considération.

4.3. Protocole d'hygiène

La matrice et sa nappe sont au contact de la muqueuse buccale. Elles doivent être désinfectées (Fig. 4.5). Les matrices sont donc plongées avant et après chaque utilisation par un sujet différent, durant 15 minutes au minimum, dans une solution diluée de désinfectant. Elles sont ensuite rincées à l'eau claire puis séchées.

Le détergent utilisé, l'*Ampholysine 4000*®, est un bactéricide (prédésinfectant) des dispositifs médico-chirurgicaux et des endoscopes ayant pour but d'abaisser l'inoculum des micro-organismes, de faciliter le nettoyage ultérieur mais également de protéger le personnel et l'environnement.

4.3.1. Composition du désinfectant

Solution limpide composée d'ammonium quaternaire, de biguanide, de tensio-actif non ionique et d'un inhibiteur de corrosion. Sa couleur varie du bleu au beige rosé, ce qui n'altère pas la qualité du produit.

4.3.2. Efficacité du désinfectant

- Bactéricide, fongicide et virucide (HIV et VHB).
- pH neutre à la concentration d'utilisation (0.5%) : 7,5 +/- 0,5.

4.3.3. Modalités pratiques d'utilisation

- Placer un peu d'eau froide dans un bac
- Ajouter 25 ml d'Ampholysine Plus® et compléter avec de l'eau froide jusqu'à obtenir 5 litres de solution.
- Immerger complètement les instruments.
- Couvrir les bacs.
- Temps de contact impératif : 15 minutes.
- Après ce temps, rincer les matrices abondamment avec de l'eau claire (sinon un goût de savon persiste signifiant qu'il reste du produit) et essuyer à l'aide d'un chiffon doux.
- Renouveler le bain après chaque nouvel acte et après chaque patient.

4.4. Conservation

La solution diluée à 0.5% peut se conserver 48 h dans un bidon fermé, étiqueté en mentionnant le nom du produit, la concentration et la date limite d'utilisation.

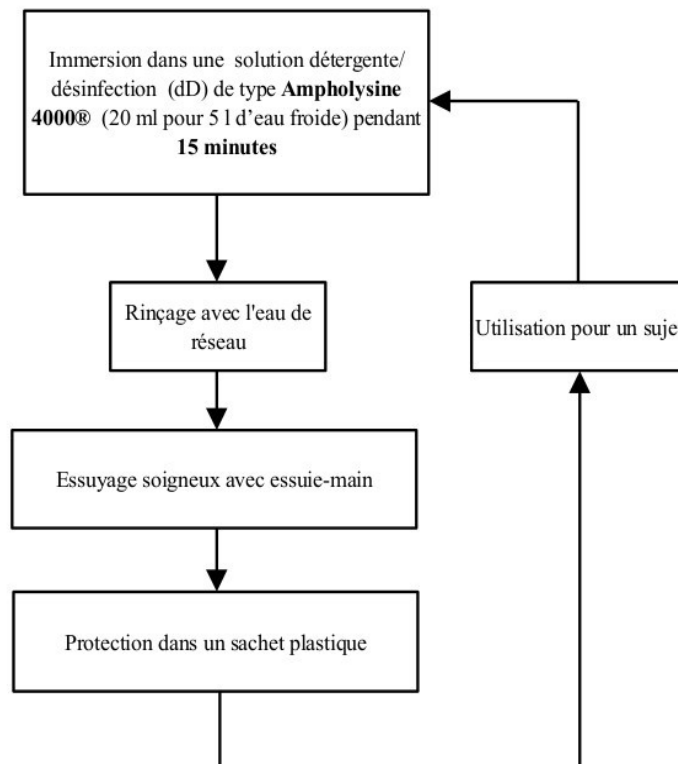


Figure 4.5. Le processus d'hygiène suivi avec la matrice de stimulation après chaque utilisation.

4.5. Protection du personnel et précautions d'emploi

- Port de gants de ménage à manchettes longues, lunettes, tablier plastique de protection
- En cas de contact avec la peau ou l'œil, se laver immédiatement et abondamment avec de l'eau et consulter la Médecine du travail et du risque professionnel.
- Conserver le produit dans son emballage d'origine, à l'abri de la lumière
- Compatibles avec: caoutchouc, plastiques (polyuréthane, polyéthylène, PVC), silicones, métaux inoxydables, verres et céramiques, endoscopes, hydroxyde de sodium (procédure soude : inactivation prion).

- Incompatible avec les autres détergents et l'eau de javel (dégagement de vapeurs toxiques).

5 Traitement des données

5.1. Mesures - Erreur d'angle et temps de réponse.

La plupart des expériences concernent la perception de l'orientation de lignes tracées dans le plan (voir Fig. 3.7, 3.9 et, plus loin 6.1 et 6.20). Dans le cadre de ces expériences, les sujets doivent percevoir l'orientation d'une ligne et sont amenés à répondre, souvent le plus rapidement possible, sous la forme d'un tracé avec un stylet graphique sur une tablette graphique (voir Fig. 3.3) pour indiquer l'orientation perçue. La réponse, sous forme angulaire, est un cumul de 2 erreurs : une erreur de perception par rapport au signal et une erreur ajoutée lors de la restitution (tracée sur la tablette graphique).

La donnée analysée est la valeur de cette erreur, c'est à dire l'écart entre la réponse attendue (angle θ_{ref} fait par la ligne avec l'horizontale) et la réponse effective faite par le sujet θ_{obs} .

Dans le traitement de nos données, nous partons de l'hypothèse, raisonnable, que la distribution des erreurs de perception d'angles sur le compact périodique $[-\pi, \pi]$ suit une distribution proche d'une loi gaussienne centrée sur une erreur de référence. Par exemple, si l'on demande à un sujet d'indiquer l'orientation d'une série de lignes orientées dans un plan, il le fera avec une certaine erreur, pour chaque orientation. Cette erreur sera principalement faible, centrée sur 0, et sa dispersion sera liée à la facilité avec laquelle il percevra (sensibilité) et restituera (précision du tracé de la réponse sur la tablette graphique) cette orientation.

Si plusieurs réponses sont possibles, comme c'est le cas dans le cadre des expériences de contradictions d'informations où 2 lignes sont tracées, l'une à un angle $\theta_{\text{ref1}} = \theta$ sur une modalité de perception (vision) et l'autre à un angle $\theta_{\text{ref2}} = \theta_{\text{ref1}} + \varphi$ sur une autre modalité de perception (TDU), l'angle φ représentant le degré de contradiction entre les 2 modalités, on doit alors s'attendre à trouver 2 réponses possibles. On devra alors *a priori* observer dans ce cas un mélange de 2 gaussiennes, l'une centrée sur un écart de $0^\circ = \theta_{\text{obs}} - \theta_{\text{ref1}}$ et l'autre sur un écart de $\varphi^\circ = \theta_{\text{obs}} - \theta_{\text{ref2}}$.

Ces hypothèses autorisent l'utilisation de certains tests statistiques et permettent surtout l'utilisation de traitements appropriés des données comme la recherche de clusters dans des mélanges gaussiens (voir 5.4).

En outre, les temps de réponses sont également enregistrés, souvent à partir du début de la stimulation ; parfois à partir du moment précis auquel la stimulation change (comme dans l'expérience 2).

La demande faite aux sujets étant une réponse rapide, les points correspondant à des réponses trop longues sont éliminés au-delà de 5 secondes.

5.2. Analyse de variance et Tests de comparaison

Afin d'éviter au maximum une perte de puissance des tests statistiques (analyse de variance et tests *a posteriori*) nous avons restreint nos analyses à des tests univariés ou ne dépendant que d'un faible nombre de facteurs. De même, nous avons cherché à restreindre au maximum le nombre de modalités que peuvent prendre chaque facteur.

La recherche d'un facteur commence par une étude ANOVA (types II et III) : en fonction du nombre de facteurs et de leur modalités on scinde la variance de la variable à expliquer. Cette variance est scindée en variance intra-groupes (résiduelle) et en variance factorielle (inter-groupes). Ce sont ces variances (intra et inter) qui sont comparées dans l'étude d'ANOVA pour signifier dans quelle mesure les différents facteurs permettent, en fonction de leur modalités, d'expliquer une discrimination significative en plusieurs groupes.

Les tests *a posteriori* sont des tests souvent non paramétriques comme le test de Mann, Withney et Wilcoxon sur les rangs, ou le test de Mood sur la médiane.

5.3. Recherche de réponses gaussiennes

Dès lors que l'on admet l'hypothèse de distribution gaussienne des écarts d'angles autour d'une réponse de référence, on peut utiliser une méthode paramétrique pour rechercher des clusters (groupes de valeurs) gaussiens dans une distribution mesurée pouvant correspondre à un mélange de distributions gaussiennes.

Pour cela, nous utilisons une méthode fondée sur la méthode EM (Expectation - Maximization) appliquée à la recherche de mélanges gaussiens. Le principe est de rechercher un modèle générateur formé d'un mélange de gaussiennes, c'est-à-dire une distribution mixte approximant au mieux les données. Ainsi, cette méthode vise à rechercher un mélange de K gaussiennes de paramètres (P_k, μ_k, σ_k) (désignant respectivement la proportion de la gaussienne k dans le mélange des K gaussiennes, de telle manière que le modèle générateur atteigne un maximum de vraisemblance. L'algorithme est divisé en 2 phases, les phases E (Expectation) et M (Maximization) se succédant l'une après l'autre, jusqu'à ce qu'un modèle satisfaisant un maximum de vraisemblance (en l'occurrence une maximisation du BIC, le bayesian information criterion) soit trouvé. À partir de conditions initiales supposant la coexistence de K clusters gaussiens de paramètres initialement aléatoires dans le mélange, on calcule à l'étape E la vraisemblance puis, durant la phase M, à partir des calculs des probabilités jointes (qu'un point soit dans un cluster k et dans un certain élément de volume de la gaussienne k) réalisés en phase E, on détermine par un processus de dérivation partielle, de meilleurs paramètres (moyenne, l'écart-type puis proportion) pour chaque cluster. Cette méthode est complètement décrite dans [Tomasi, 2004].

Sous R , il existe une bibliothèque logicielle dédiée, nommé *mclust*. Néanmoins, cette méthode ne fonctionne pas bien pour la détermination de mélanges gaussiens sur un compact périodique comme nos densités pour des valeurs angulaires d'orientations comprises entre $-\pi$ et π , ces 2 angles étant égaux si l'on ne considère que l'orientation (et pas le sens) des lignes tracées. Nous avons donc du réécrire une fonction, sous R , capable de prendre en compte le fait qu'une gaussienne sur un tel ensemble de valeurs périodiques continue de part et d'autre des bornes (fictives) de cet espace, comme sur un anneau ou un tore. Le principe est sensiblement le même que celui décrit par C. Tomasi, à ceci près qu'on calcule chaque gaussienne k sur un compact fermé centré sur sa moyenne et de largeur 2π , donc sur $[\mu_k - \pi, \mu_k + \pi]$. La réécriture a été faite par N. Glade et son code est disponible à cette URL : <https://sites.google.com/site/nicolasglade/research/PCGME.R?attredirects=0>.

Voici un exemple illustrant cette méthode. La figure 5.2 montre, sous la forme d'un histogramme, le résultat d'un tirage aléatoire de 342 valeurs réparties par 2 tirages gaussiens (modèle générateur) de paramètres respectifs $G_1(P_1=0.614, \mu_1=-4, \sigma_1=25.47)$ et $G_2(P_2=0.386, \mu_2=84.22, \sigma_2=15.78)$ (voir tableau 5.1).

Nous avons effectué une recherche de mélange de 2 clusters gaussiens sur ce jeu de données. Les résultats trouvés montrent la performance de cette méthode. On trouve 2 gaussiennes ayant des paramètres très proches de ceux du modèle générateur (Tab. 5.1 et Fig. 5.2).

	Cluster	Paramètres		
		Théoriques	Attendus	Calculés
Proportion	1	210	210	201
	2	132	132	140
Moyenne	1	-3	-4	-4.24
	2	85	84.22	83.70
Déviat ion standard	1	23	25.47	23.35
	2	15	15.78	17.85

Tableau 5.1. Le tableau contient les valeurs de paramètres (moyenne écart type et proportion) de 2 gaussiennes avec de gauche à droite : Valeurs théoriques des clusters gaussiens utilisés comme modèle générateur (paramètres théoriques). Valeurs attendues générées par le modèle générateur (ici on connaît l'appartenance des valeurs à l'un ou l'autre des 2 clusters; ce sont ces valeurs que notre traitement doit retrouver. Valeurs calculées avec notre méthode.

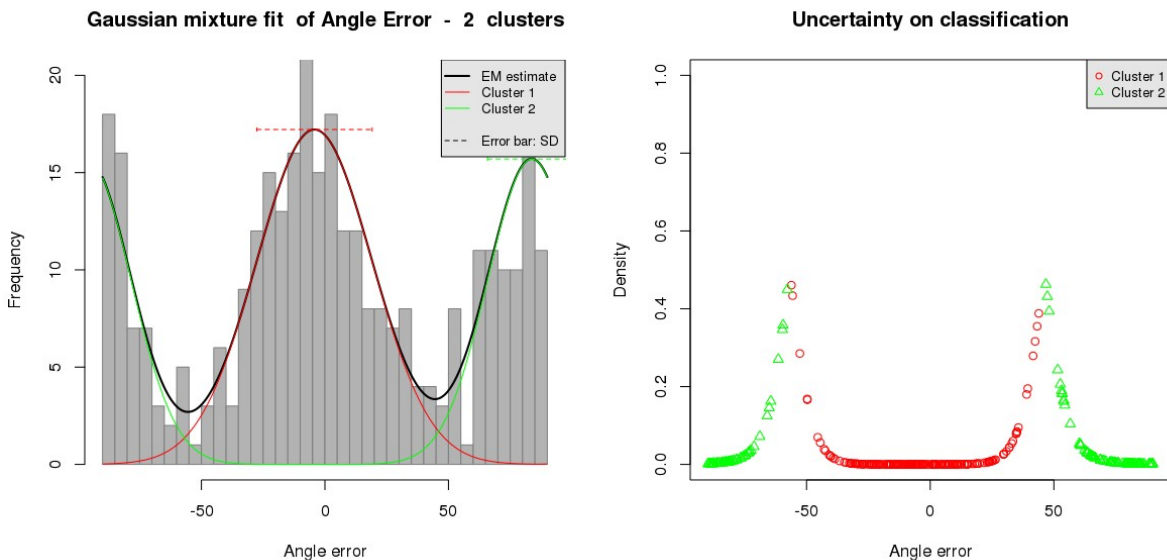


Figure. 5.2. Représentation d'un modèle de jeu de donnée factice (à gauche) l'histogramme comporte un mélange de deux gaussiennes (à droite) diagramme de l'incertitude sur la classification.

5.4. Qualité des modèles

Dans l'exemple précédent, nous avons recherché un mélange de 2 gaussiennes. Nous aurions pu rechercher une seule gaussienne ou au contraire un nombre plus important (>2) de clusters gaussiens. On parlera alors de modèle à 1 gaussienne, 2 gaussiennes, etc. Pour comparer la qualité de ces modèles (leur efficacité à représenter le jeu de données), nous avons systématiquement comparé leurs BIC et AIC. Le BIC ou Bayesian Information Criterion est un score fondé sur le calcul de la vraisemblance et pondéré par le nombre de paramètres libres (proportion, moyenne et écart-type multipliés par le nombre de clusters). Il doit être le plus élevé possible, c'est-à-dire qu'on retient le modèle de BIC le plus élevé. L'AIC ou Akaike Information Criterion est un score fondé sur l'entropie (log de l'erreur) et sur un coût au nombre de paramètres libres comme dans le BIC. Dans les 2 cas ces paramètres permettent de sélectionner les meilleurs modèles. Nous les avons systématiquement calculés tous les 2 et nous nous sommes assurés que ces modèles convergeaient.

Dans l'exemple précédent, le meilleur modèle pour représenter les valeurs données figure 5.2, est un modèle à 2 gaussiennes. Les autres modèles (1 gaussienne ou >2 gaussiennes) ne sont pas de bons modèles.

Le bootstrap, présenté ici comme moyen de tester la robustesse des modèles de mélanges gaussiens peut aussi bien entendu être utilisé pour tous les tests de comparaisons de moyenne ou pour les ANOVA.

5.5. Bootstrap

Afin de nous assurer de la robustesse de nos traitements statistiques, nous avons souvent bootstrappé nos échantillons. Le bootstrap consiste en un grand nombre d'échantillonnages avec remise de même effectif sur lesquels on effectue autant de traitements. Si le traitement est robuste, alors son résultat moyen converge vers le résultat que l'on obtient avec l'échantillon d'origine et l'écart-type des traitements est faible.

Nous avons utilisé le bootstrap principalement pour évaluer la robustesse des estimés de mélanges gaussiens.

5.6. Comparaison des modèles générateurs

Comme nous le verrons au chapitre 7, l'utilisation de la recherche de mélanges gaussiens (de clusters gaussiens) sur les données des différentes expériences consistant à déterminer l'orientation de lignes tracées sur le TDU, permet de lever l'ambiguïté de compréhension de l'orientation par les individus menant à un phénomène d'inversion spatiale. Dans ce cas, il est envisageable d'utiliser des données simulées (ou redistribuées) à partir d'un modèle générateur (voir le point 5.3) composé d'un mélange de gaussiennes pour effectuer les tests statistiques.

Supposons que l'on s'intéresse à la comparaison de 2 distributions d'erreurs d'angles avec ou sans présence d'un facteur pouvant affecter ces distributions (par exemple, des saccades du signal). On dispose donc de 2 jeux de données pour lesquels une certaine proportion d'individus a pu inverser le repère spatial et répondre un angle tenant compte de cette inversion. Pour générer un jeu de données « corrigées » à partir du modèle générateur gaussien on procède comme suit. D'abord on identifie dans chaque jeu de données les 2 clusters gaussiens (on peut montrer grâce à une quantification par l'AIC et le BIC que le meilleur modèle de mélange gaussien dans ces données comporte 2 clusters systématiquement) supposés correspondre aux réponses non inversées et aux réponses inversées (Fig. 5.4 gauche). Les paramètres utilisés sont les paramètres obtenus par bootstrap sur le jeu de données initial (Fig. 5.3).

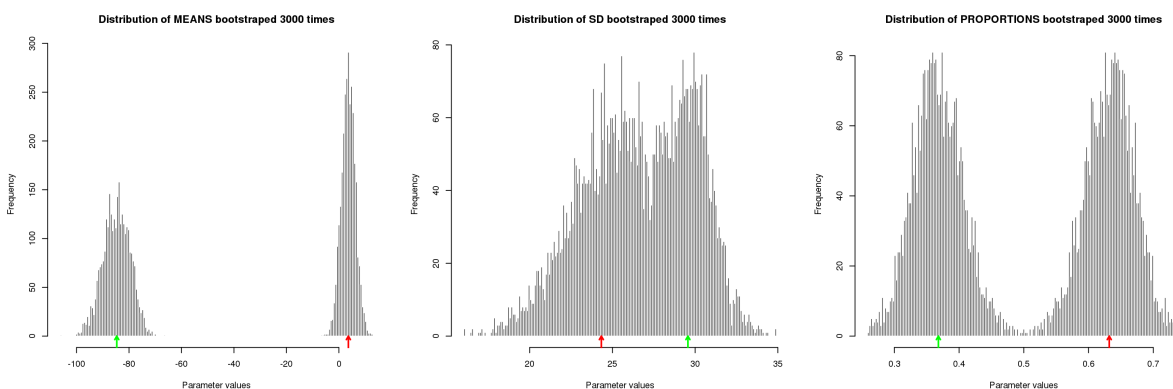


Figure 5.3. Recherche des paramètres des 2 clusters gaussiens par bootstrap (ici 3000 itérations). A gauche les moyennes, au milieu les écarts types, à droite les proportions. Les flèches rouge et verte indiquent les valeurs retenues pour chacun des paramètres et pour chaque cluster.

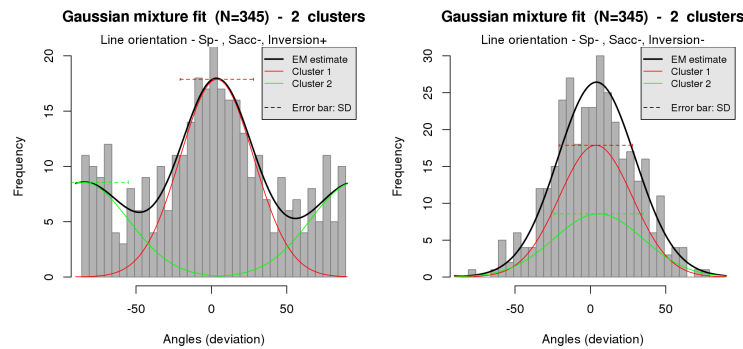


Figure 5.4. A gauche : clusters gaussiens identifiés dans un jeu de données (ici les données contrôle de l'expérience décrite section 6.1: des mesures de perception de l'orientation d'une ligne). Les courbes verte et rouge désignent les 2 clusters trouvés. La courbe noire est la somme des 2 gaussiennes et modélise le signal complet. A droite les 2 clusters recentrés sur une erreur nulle (le cluster vert a été décalé de 90°). L'histogramme correspond à un jeu de données simulé à partir du modèle corrigé (courbe noire).

Connaissant leurs paramètres (proportion, moyenne et écart type), on génère un jeu de données en faisant des tirages gaussiens selon le premier cluster gaussien non inversé et selon le second, après redressement, c'est à dire en décalant sa moyenne de 90° . On obtient ainsi un jeu de données proche de celui qu'on pourrait espérer si aucune inversion n'avait lieu (Fig. 5.4 droite). Ces jeux de données corrigés peuvent alors être utilisés pour comparer les distributions corrigées. Le bootstrap (on génère B couples de jeux de données pour effectuer B comparaisons) est recommandé pour s'assurer de la robustesse des conclusions que l'on tire sur ces données générées à partir d'un modèle du jeu de données initial.

Ce procédé permet de tirer des enseignements mais n'est pas exempt de critiques. En l'occurrence, on ne peut ignorer ce fait : on travaille sur des données simulées à partir d'un modèle, certes obtenu à partir du jeu de données initial, mais fondé sur des hypothèses (le fait qu'il n'existe que 2 clusters dans ce jeu de données initial et qu'ils soient gaussiens) elle-mêmes critiquables.

III - Expériences

6 Perception d'informations statiques

Comme cela a été décrit à plusieurs endroits précédemment, la stimulation continue de récepteurs vibro-tactiles comme sur la langue ou la peau ou de récepteurs photoniques comme dans la rétine provoque un phénomène de fading : au fur et à mesure de la stimulation, l'organe se désensibilise (ajustement du seuil de perception) et perd peu à peu la perception du stimulus. C'est un inconvénient majeur inhérent à la façon dont fonctionnent actuellement tous les dispositifs de substitution sensorielle et de rétines artificielles, lorsque l'on souhaite faire percevoir à un sujet des images (signaux) statiques (immobiles dans le champ de perception).

Dans le cas de la vision normale, le problème de fading qu'on peut mettre facilement en évidence avec le test de Troxler, est résolu grâce au mécanisme de mouvements des yeux. Nous avons émis l'hypothèse que l'application de mécanismes semblables à ceux du mouvement des yeux, serait probablement bénéfique dans le cadre de la stimulation électro-tactile sur la langue.

Au cours de la série d'expériences qui suit, nous voulons évaluer si et dans quelle mesure la présence de saccades ou de scintillement du signal électro-tactile permet de mieux percevoir et de maintenir la perception sur la langue de scènes statiques. En l'absence de ces facteurs, on s'attend à ce que cette information ne soit plus perçue correctement.

6.1. Expérience 1 - Qualité de perception de l'orientation d'une ligne 2D

6.1.1. Objectif

Dans cette expérience on cherche à savoir si la présence de saccades ou de scintillement dans un signal image améliore sa perception. L'expérience consiste, pour le sujet, à reconnaître le tracé d'une ligne sur la matrice du TDU (tracée avec une égale intensité sur toute la largeur de la matrice avec une orientation aléatoire) et à en déterminer l'orientation. L'hypothèse faite est que la qualité (précision angulaire) et la facilité (temps de

réponse) de perception de l'orientation augmente en présence de ces facteurs. On se fonde ici sur l'idée que si le signal est mieux perçu, à cause de légers tremblements, alors le sujet aura plus de facilité à le comprendre. Nous verrons que cette hypothèse est mal formulée : il est en effet difficile d'améliorer la qualité d'une image en la faisant trembler. En revanche, on peut aider à maintenir la perception de la ligne (image) avec ces facteurs. C'est ce que fait l'expérience 2.

6.1.2. Protocole

On soumet à l'utilisateur, dans une modalité de perception donnée (visuelle ou TDU), le tracé d'une droite d'égale intensité en tout point (Fig. 6.1). Celui-ci doit répondre le plus rapidement et le plus précisément possible en dessinant la ligne avec un stylet sur une tablette graphique.

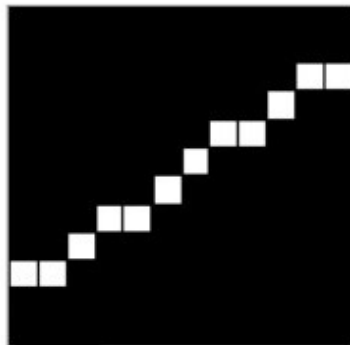


Figure 6.1. Tracé d'une ligne d'égale intensité à 35°

On réalise des séries d'environ 140 mesures¹ par utilisateur et par session (140 mesures sont faites, mais certaines sont invalidées *a posteriori* lorsqu'elles sont aberrantes, par exemple lorsqu'un sujet répond en 40 secondes). Tous les utilisateurs font au moins 2 sessions, certains jusqu'à 4. Chaque session se déroule un jour différent. Chaque session est précédée et suivie d'une étape de calibrage.

¹ NB: En pratique, nous avons fait 210 mesures (générant un total de 6930 mesures sur 14 sujets) car nous avons mesuré au cours de la même expérience la restitution du signal selon 3 modalités : modalités visuelle, TDU, mais aussi bimodalité Visuel+TDU. Les mesures en bimodalité sont présentées au chapitre 8.

Les séries de 140 mesures sont faites avec une variation des paramètres (scintillement, saccades, modalité) randomisés par blocs de 10 mesures. Elles se font dans les 2 modalités¹ (visuelle ou TDU), avec ou sans scintillement (avec des valeurs de probabilité de scintillement de 0.0, 0.3, 0.6, 1.0, ce qui représente aussi la fraction de temps pendant laquelle le scintillement se fait ; son intensité est fixée à 100%), avec ou sans saccades (avec 0.0, 0.3, 0.6, 1.0, comme probabilités de saccades ; leur intensité est fixée à 100%), saccades et scintillement n'étant jamais testés en même temps. Nous avons finalement recueilli un total de 4619 mesures valides chez 14 sujets, 7 hommes et 7 femmes jeunes (entre 22 et 36 ans).

Les données sont stockées dans un fichier dans lequel on indique l'identifiant du sujet (ID) qui donne accès à son sexe, à son âge et sa profession, l'angle d'entrée (InputAng), l'angle mesuré (OutputAng), la valeur de la fréquence du scintillement (Sparkle) et des saccades (Saccades), la modalité de restitution de la scène (VisuMode: V=visuelle ou T=TDU), le temps de réponse (en ms), le numéro de la session et enfin l'écart absolu (ErrAng). Tous les angles (ErrAng) sont compris entre -90° et 90° (ils peuvent être aussi convertis en angles compris entre 0° et 90°). Le tableau ci-dessous montre quelques résultats à titre d'exemple (Tab. 6.2).

	ID	InputAng	OutputAng	Sparkle	Saccades	VisuMode	Duration	Session	ErrAng
1	2	105	-58	0	0	V	1170	1	17
2	2	14	29	0	0	V	1220	1	15
...
2131	5	22	-57	0	0	T	6130	1	-79
2132	5	250	-38	0	0	T	790	1	72
...

Tableau 6.2. Quelques valeurs saisies au cours de l'expérience 1

Parmi les données saisies, nous avons aussi acquis des renseignements sur les sujets : âge, sexe, niveau d'expertise du maniement du périphérique de saisie (souris, clavier, ...) et de repérage dans l'espace (lié à la profession : biologiste, informaticien ...). Nous ne présenterons pas d'analyse des résultats au regard de ces variables dans ce manuscrit car il n'est pas évident que les effets liés à l'expertise ou au sexe soient significatifs dans l'échantillon dont nous disposons.

Enfin, nous avons demandé à chaque personne de répondre à un questionnaire (voir Annexes). Il leur était demandé, *a posteriori*, d'attribuer une note entre 0 et 4 aux questions suivantes² :

- *Évaluez la facilité de définir la direction de la ligne en mode Visuel*
- *Évaluez la facilité de définir la direction de la ligne en mode TDU*
- *Pensez vous que votre calibrage était adapté ?*
- *Évaluez l'ergonomie générale du dispositif pour le test que vous venez de faire.*

En outre, il leur était demandé s'ils étaient parvenus à « visualiser » une ligne en mode TDU (réponse : oui ou non); si oui, avec ou sans saccades (et intensité); avec ou sans scintillement (intensité); enfin, quelles orientations étaient les plus aisées à visualiser (horizontales, verticales, diagonales).

6.1.3. Résultats

Erreurs d'angles.

La figure 6.3 montre sous la forme d'histogrammes les résultats bruts des réponses sous forme d'erreurs d'angles (affichées entre 0 et 90°) mesurés dans le cas de perception visuelle ou sur le TDU, et en fonction de l'absence ou de la présence de scintillement ou de saccades. A première vue, comme on peut s'y attendre, avec le TDU, les erreurs augmentent. Il n'est pas non plus évident d'après ces résultats que la présence de saccades ou de scintillement, modérés ou non (on ne montre pas ici les résultats selon les niveaux de saccades ou de scintillement), augmentent sensiblement la qualité de perception dans la modalité TDU. Ceci est discuté dans le paragraphe 6.1.4 consacré à l'analyse statistique de ces résultats.

2 Comme cela a été mentionné dans la note de bas de page N°1, l'expérience a été faite aussi en bimodalité,. Nous posions donc également la question *Évaluez la facilité de définir la direction de la ligne en bimodal (Visuel+TDU)*

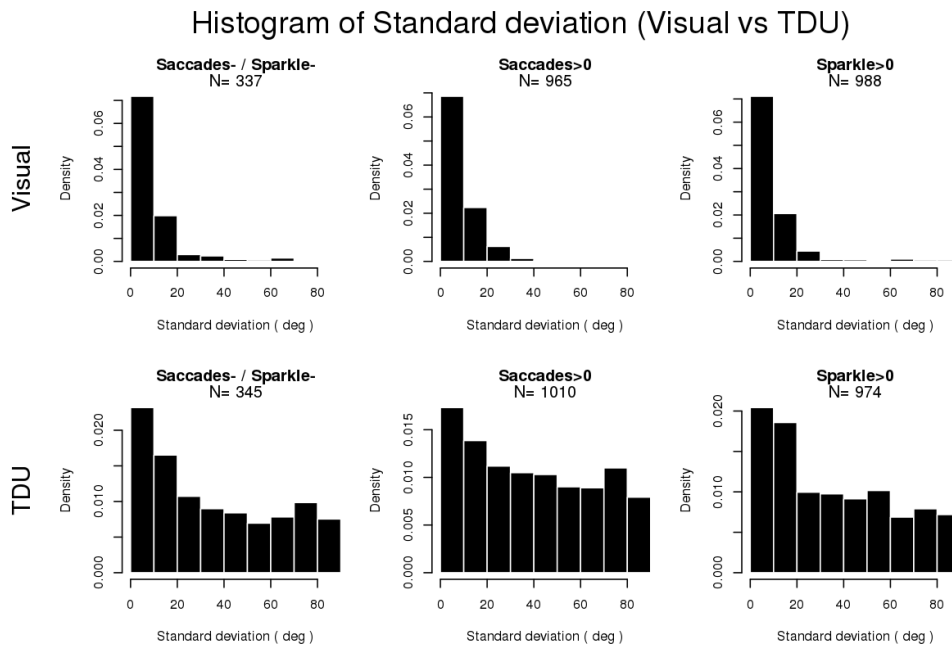


Figure 6.3. Histogrammes des erreurs d'angles (standard deviations) dans les modalités visuelles ou TDU et selon l'absence ou la présence de scintillement ou de saccades (tous niveaux confondus : 0.3, 0.6 et 1.0). L'effectif N est indiqué pour chaque expérience

Dans la modalité TDU, on voit que le nombre de faibles erreurs diminue des 2/3 environ tandis qu'augmentent les erreurs supérieures à 20°. Celle-ci semblent à peu près uniformément réparties. Nous verrons dans la section suivante (6.1.4 Traitement statistique et interprétation) que ce n'est pas réellement le cas. On peut cependant étudier comment se distribuent ces erreurs en fonction de l'angle de stimulation (input angle). C'est ce que représentent les figures 6.4 et 6.5. On remarque ainsi un phénomène singulier : une variation périodique des erreurs en fonction de l'angle de stimulation. La figure 6.4 correspond aux résultats bruts : on y compare les erreurs dans les modalités Visuelle ou TDU, selon les différents traitements (absence ou présence de saccades ou de scintillement). La figure 6.5 est obtenue après traitement de ces données brutes. Le traitement est réalisé avec un filtre médian appliqué avec une fenêtre mobile de 15°. La figure 6.5 de gauche prend en compte toutes les erreurs (comprises entre 0° et 90°) tandis que la figure de droite ne traite que les erreurs supérieures à 20°. Ces variations présentent des maxima pour les angles de stimulation de 45°, 135°, 225° et 275°, c'est-à-dire les angles qui correspondent à des orientations diagonales des lignes. Les erreurs sont par contre faibles dans le cas des lignes verticales et horizontales. On voit que, dans la modalité visuelle, il n'y a pas de variation périodique (les variations importantes pour cette modalité dans la figure 6.5 de droite viennent du fait que l'erreur

médiane n'est calculée que sur un très faible nombre de valeurs d'erreur, parfois élevées, dans cette modalité). En revanche, on voit que ces variations périodiques sont bien présentes dans les petites comme dans les grandes erreurs dans la modalité TDU.

Ces observations confirment les impressions des sujets concernant leurs sensations vis-à-vis des perceptions en fonction des orientations verticales, horizontales ou diagonales : le questionnaire montre en effet une mauvaise appréciation pour les lignes diagonales (et horizontales) en comparaison des lignes verticales (Fig. 6.6). On voit cependant qu'il y a une grande inter-variabilité des réponses entre les sujets (Fig. 6.7) et on s'aperçoit de la présence potentielle d'un effet provoquant un biais d'interprétation sur l'importance des erreurs d'angle : l'inversion de la spatialité chez certains sujets et pour certains angles. Ceci sera discuté dans une section ultérieure.

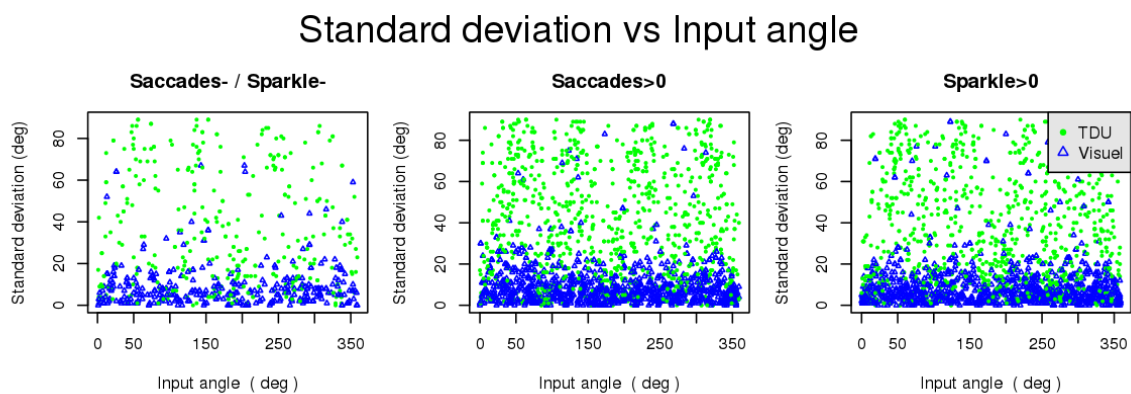


Figure 6.4. Nuages de points des erreurs de réponse sur les orientations (angle standard deviation) en fonction des angles de stimulation (input angles), de la modalité de perception (visuelle ou TDU) et de la présence ou non de saccades ou de scintillement. Dans la modalité TDU, on observe une variation périodique des erreurs en fonction du stimulus, avec des maxima sur les diagonales.

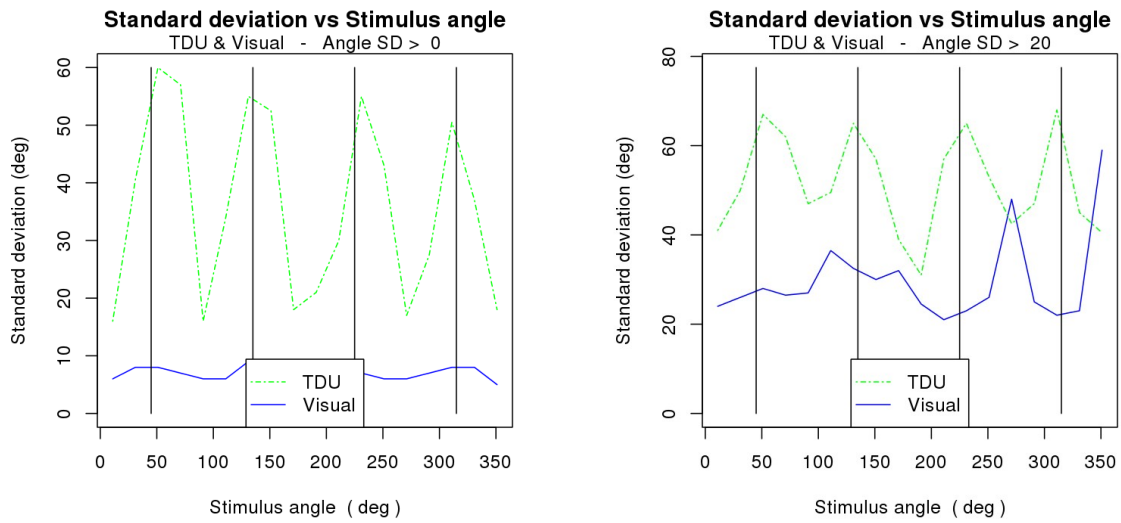


Figure 6.5. Médiane mobile (filtre médian avec une fenêtre de 15°) des erreurs de réponse sur les orientations (angle standard deviation) en fonction des angles de stimulation (input angles) et de la modalité de perception (visuelle ou TDU), toutes conditions (scintillement, saccades ou non) confondues. A gauche, on prend en compte toutes les réponses entre 0° et 90° ; à droite, on ne prend que les erreurs supérieures à 20° (donc erreurs comprises entre 20° et 90°). Ces courbes font ressortir la variation des erreurs de réponses en fonction de l'angle de stimulation.

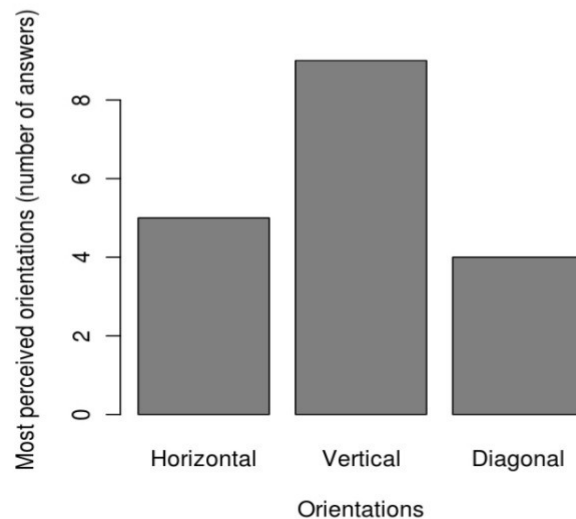
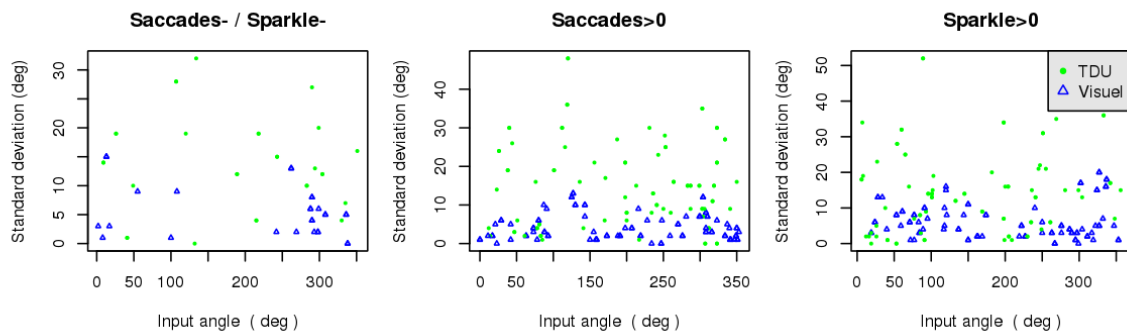


Figure 6.6. Résultats du questionnaire sur la sensation des sujets vis-à-vis de la perception des orientations des lignes. La question posée aux sujets est « quelle est la direction des lignes la plus facile à percevoir en utilisant le TDU ? ». Les lignes verticales sont dites les mieux perçues au contraire des lignes diagonales et horizontales.

Standard deviation vs Input angle - user 4



Standard deviation vs Input angle - user 5

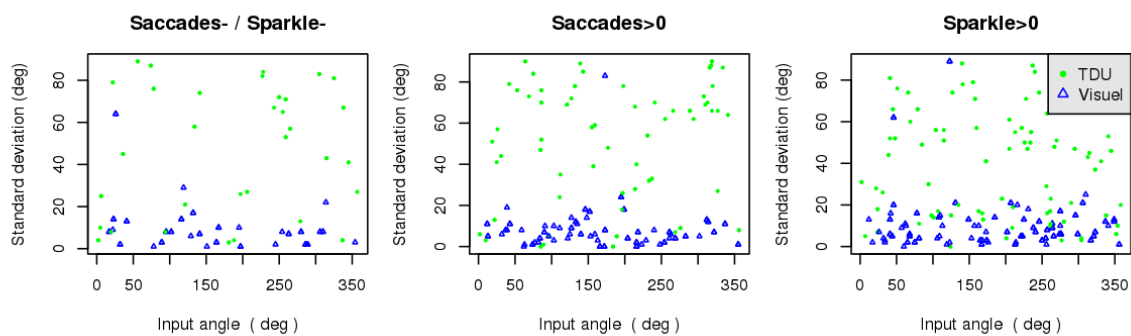


Figure 6.7. Nuages de points, pour 2 sujets (sujets 4 et 5) des erreurs de réponse sur les orientations (angle standard deviation) en fonction des angles de stimulation (input angles), de la modalité de perception (visuelle ou TDU) et de la présence ou non de saccades ou de scintillement. On observe toujours la variation périodique des erreurs d'angle dans la modalité TDU. On constate cependant une grande inter-variabilité des réponses dans la modalité TDU entre les 2 sujets. En particulier, le sujet 5 semble avoir des réponses soit proches de 0° , soit au contraire proches de 90° . Nous verrons qu'il s'agit vraisemblablement d'un effet d'inversion spatiale.

Temps de réponse.

De la même manière que pour la figure 6.3, la figure 6.8 montre les temps de réponse correspondants. On rappelle que la consigne était de répondre le plus rapidement possible. Ces temps de réponse sont mesurés comme le temps passé entre le début de la stimulation et le début de la réponse, au moment où le stylet touche la tablette graphique pour commencer la réponse. Une première analyse visuelle, montre une augmentation des temps de réponse entre les modalités Visuelle et TDU. On voit cependant qu'avec le TDU, la très grande majorité des réponses se situe dans des temps courts comme pour la modalité visuelle. Les histogrammes ne semblent pas particulièrement différer si l'on ne tient compte que des niveaux modérés de scintillement et de saccades. Enfin, et c'est le plus important, cette fois, il semble bien y avoir

un effet significatif favorable des saccades comme du scintillement sur les temps de réponse (diminution des temps de réponses).

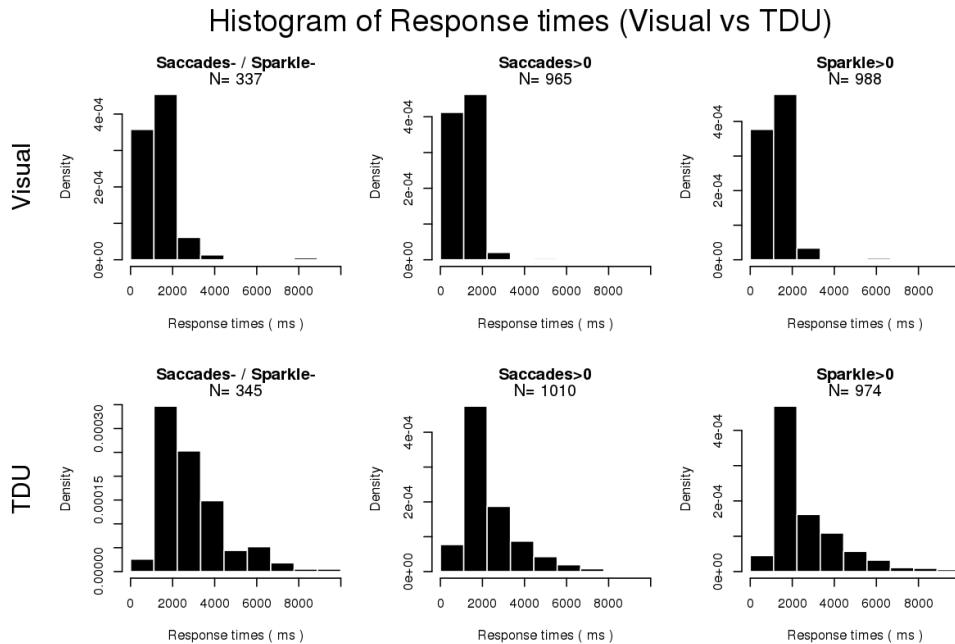


Figure 6.8. Histogrammes des temps de réponses (en ms) dans les modalités visuelles ou TDU et selon l'absence ou la présence de scintillement ou de saccades (tous niveaux confondus : 0.3, 0.6 et 1.0). L'effectif N est indiqué pour chaque expérience.

Response times vs Input angle

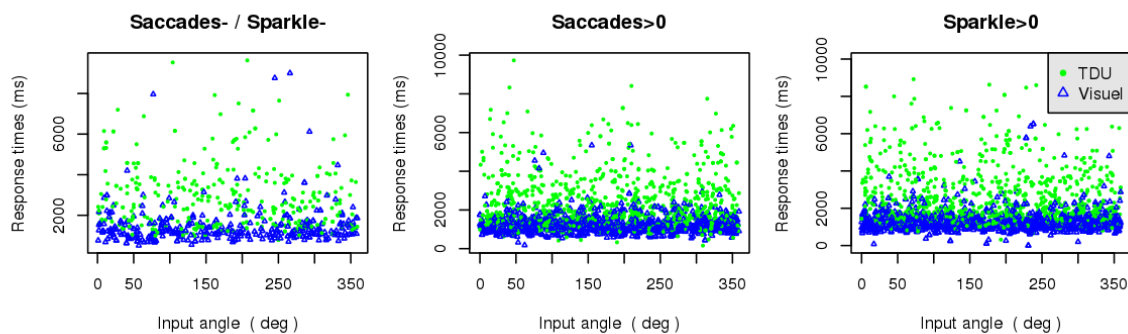


Figure 6.9. Nuage de points des temps de réponse en fonction de l'angle de stimulation et en fonction de la présence de saccades ou de scintillement. Il ne semble pas y avoir d'effet de l'angle de stimulation. En revanche, en présence de saccades ou de scintillement, la moyenne semble se décaler vers des temps de réponse plus courts.

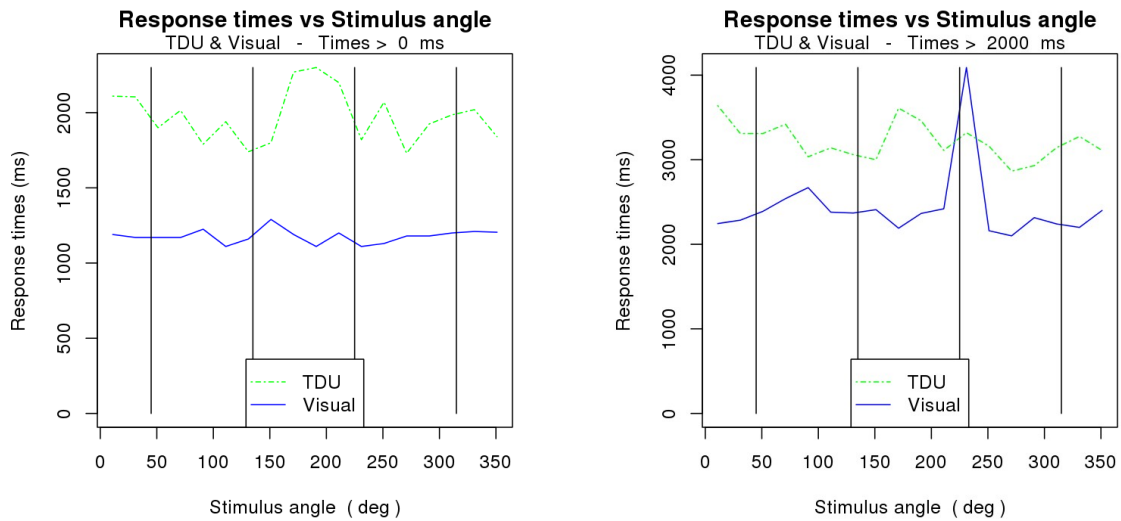
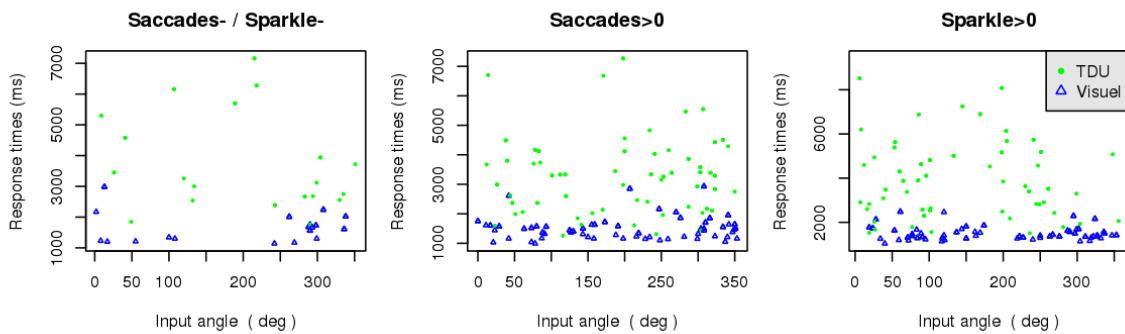


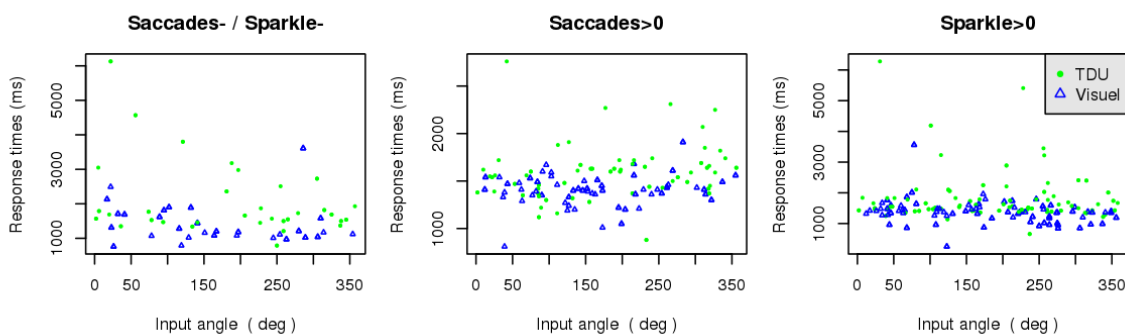
Figure 6.10. Médiane mobile (filtre médian avec une fenêtre de 15°) des temps de réponse en fonction des angles de stimulation (input angles) et de la modalité de perception (visuelle ou TDU). A gauche, on prend en compte tous les temps de réponse ; à droite, on ne prend que les temps de réponse supérieurs à 2 secondes. Ces courbes montrent qu'au contraire des erreurs d'angles, on n'observe pas de variations (significatives) des temps de réponse en fonction de l'angle de stimulation. On observe aussi qu'en moyenne les temps de réponse sont plus élevés de 1 seconde dans la modalité TDU.

De la même manière que pour les erreurs d'angles, on peut s'intéresser à l'effet de l'angle de stimulation sur les temps de réponse. Les figures 6.9 et 6.10 montrent qu'il ne semble pas y avoir de relation entre ces variables. La figure 6.10 montre en outre que les temps de réponse passent d'une moyenne d'environ 1200 ms à 2000 ms entre les modalités visuelle et TDU. Cependant, une analyse plus poussée des données chez chaque sujet montre une grande inter-variabilité entre les sujets en ce qui concerne les temps de réponse : certains ont des temps de réponse rapide quelles que soient les modalités de perceptions à la différence d'autres dont les temps de réponse sont très affectés dans la modalité TDU.

Response times vs Input angle - user 4



Response times vs Input angle - user 5



Response times vs Input angle - user 7

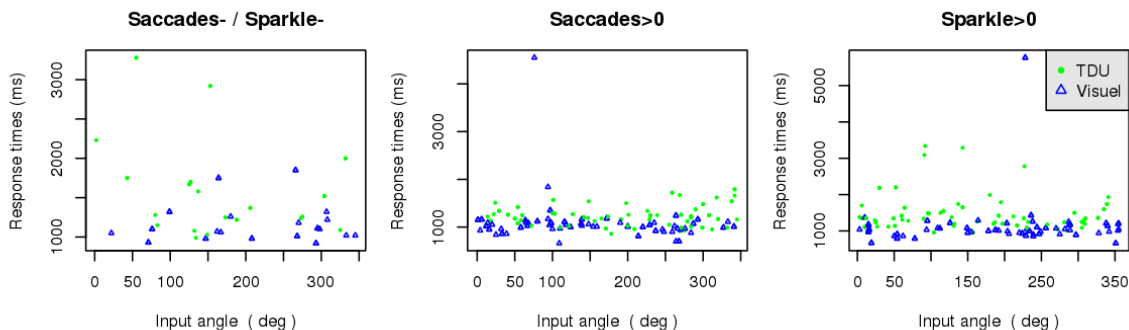
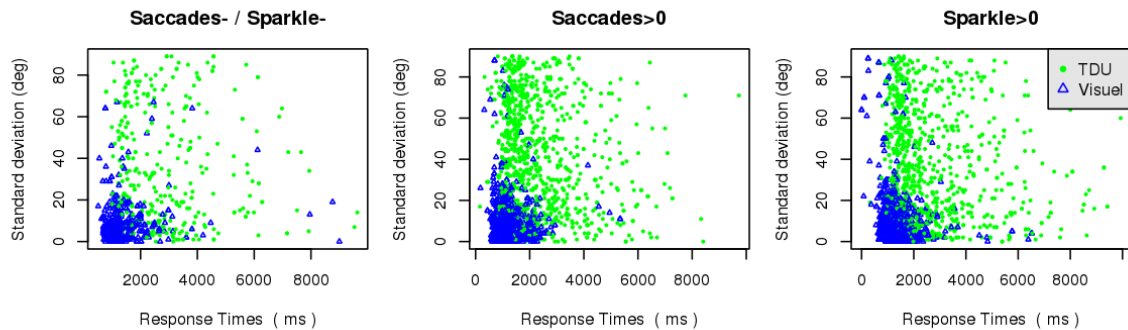


Figure 6.11. Inter-variabilité des temps de réponse (en fonction de l'angle de stimulation) entre les sujets. Certains sujets font des réponses rapides avec le TDU, presque aussi rapides que dans la modalité visuelle (sujets 5 et 7), d'autres (sujet 4) au contraire voient leur temps de réponse très affectés dans la modalité TDU. On voit aussi dans le cas du sujet 5 que la présence de saccades a affecté les temps de réponse non seulement avec le TDU, mais aussi dans le cas de la vision normale. Il est possible que la présence de tremblements de l'image ait détourné légèrement son attention et ralenti ses réponses.

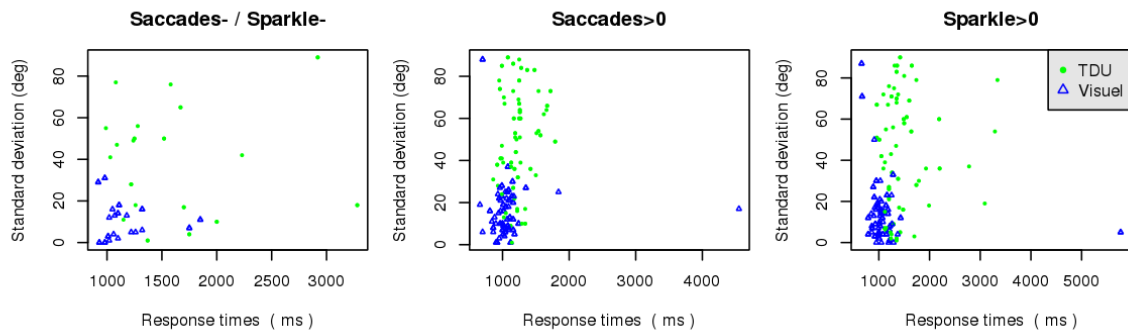
Relation entre erreurs d'angles et temps de réponses.

Nous nous sommes enfin posé la question d'une relation possible entre les temps de réponse et les erreurs de perception. La question est de savoir si lorsqu'un sujet peine à déterminer un angle, il met plus de temps également à répondre.

Standard deviation vs Response Times



Standard deviation vs Response times - user 7



Standard deviation vs Response times - user 9

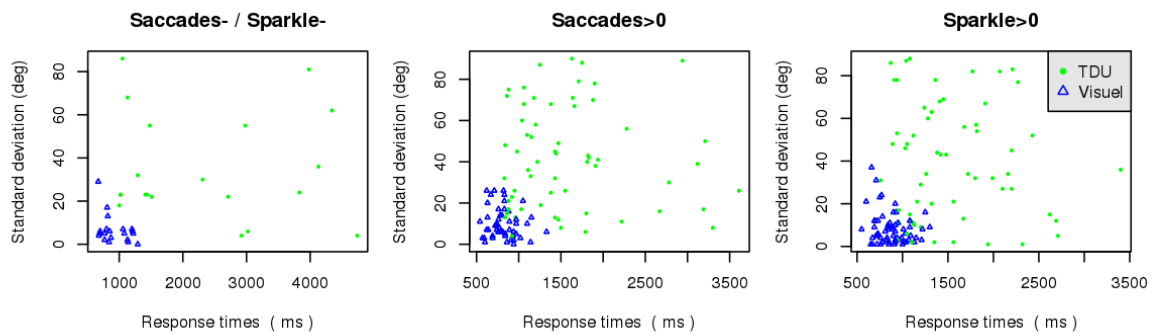


Figure 6.12. Relation entre le temps de réponse et l'erreur d'angle. En haut les données pour tous les sujets ; au milieu et en bas, les données des sujets 7 et 9 (inter-variabilité). Il ne semble pas y avoir de corrélation entre ces 2 mesures : les temps de réponse ne semblent pas augmenter avec les erreurs de perception. Cependant, si chez certains sujets (ex: sujet 7) on n'observe clairement aucune corrélation, chez d'autres (ex: sujet 9) on peut suspecter une certaine corrélation positive. Le tableau 6.13 infirme cette dernière affirmation.

Subject	1	2	3	4	5	6	7	8	9	10	11	12	13	14
cor	0.03	0.06	-0.04	0.12	0.04	0.02	0.02	0.02	-0.07	0.01	-0.03	0.03	-0.09	-0.004
p value	0.71	0.46	0.54	0.09	0.49	0.85	0.75	0.82	0.41	0.92	0.70	0.72	0.30	0.96

Tableau 6.13. Corrélations entre le temps de réponse et l'erreur d'angle. Les valeurs de corrélation sont très faibles et les p-valeurs des tests de corrélation ne montrent aucune significativité de ces corrélations.

6.1.4. Traitement statistique et Interprétation sur l'Effet des Saccades et du Scintillement.

6.1.4.1. Interprétation en première instance

Dans une première série de résultats publiés [Chekhchoukh 2012], nous avons montré que le scintillement de l'image (ligne tracée sur le TDU ou sur le moniteur) améliorerait sensiblement la précision des réponses sur les angles perçus (au contraire des saccades qui dégradait cette perception). De plus, le scintillement comme les saccades sont un facteur essentiel affectant les temps de réponse. C'est ce que montrent les graphiques 6.14 et 6.16 respectivement ainsi que le tableau récapitulatif 6.15 des tests d'ANOVA II et des tests *a posteriori* (ici des tests de rang Wilcoxon - Mahn et Withney, mais les résultats de tests de Student sur les moyennes convergent).

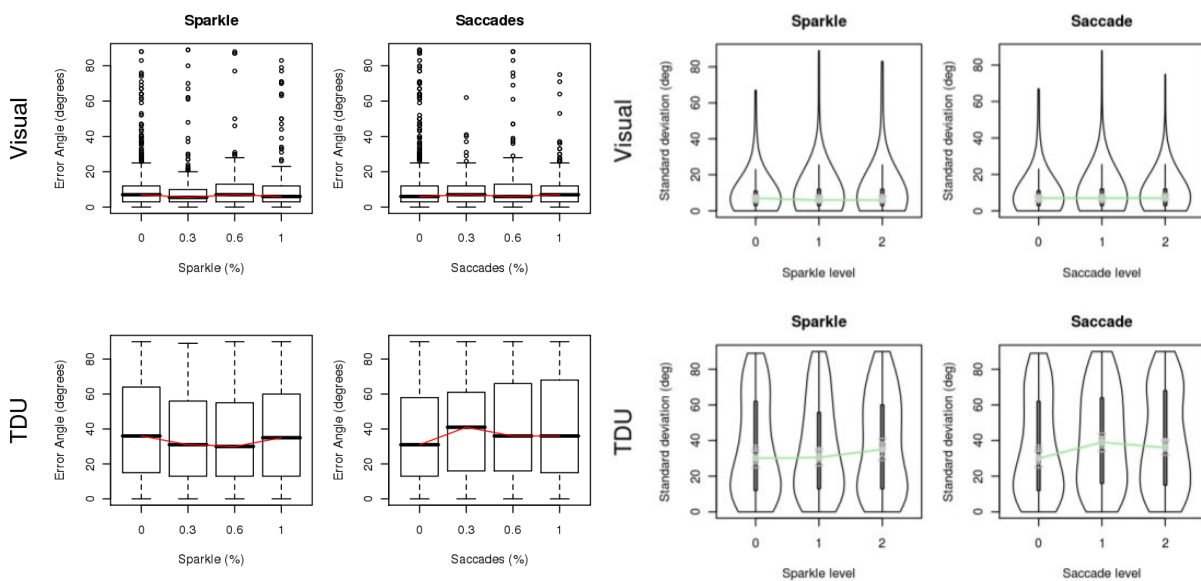


Figure 6.14. Effet du scintillement ou des saccades de l'image sur la perception des angles. A gauche, tous les niveaux de fréquence d'application de ces altérations du signal sont présentés. A droite, les niveaux modérés (1) correspondent à des fréquences de 30% et de 60% et les niveaux élevés (2) correspondent à une fréquence de 100 %. De plus, l'utilisation de *violin plots* permet de mieux comprendre comment sont distribuées les erreurs. Les courbes permettent de mieux visualiser les variations des moyennes.

On voit en particulier sur les boîtes à moustaches de la figure 6.14 que les niveaux modérés de scintillement réduisent les erreurs moyennes de perception des angles. Les niveaux forts de scintillement n'ont pas d'effet et la présence de saccades augmente l'erreur. Les *violin plots* tracés en fonction de niveaux (nul, modéré ou fort) de scintillement ou de

saccades écrasent (visuellement) l'effet du scintillement. Cet effet d'amélioration, bien présent comme on le voit dans le tableau 6.15 récapitulant l'analyse statistique de cette expérience, reste très faible. En outre, on distingue sur les *violin plots* un 2^{ème} mode (dans les grandes erreurs, vers 90°) déjà observé sur les histogrammes (Fig. 6.3) et vraisemblablement dû à l'inversion des repères spatiaux pour certains angles chez certains sujets (Figs. 6.4 à 6.7).

<i>Models and Factors</i> (and alternative hypotheses)	<i>ANOVA</i>		<i>Test</i> p-value
	F-statistic	p-value	
AngleSD~Sparkle*Visu			
<i>Sparkle</i>	F(2,6921) = 3.64	2.62 10 ⁻²	
<i>Visu</i>	F(2,6921) = 929	≈ 0	
<i>Sparkle:Visu</i>	F(4,6921) = 2.37	4.99 10 ⁻²	
AngleSD~Sparkle (TDU only)	F(2,2326) = 4.65	9.6 10 ⁻³	
<i>H₁ : AnglesSD_{moderate} < AnglesSD_{nosparkle}</i>			4.7 10 ⁻³
<i>H₁ : AnglesSD_{strong} < AnglesSD_{nosparkle}</i>			23.0 10 ⁻²
AnglesSD~Saccades*Visu			
<i>Saccades</i>	F(2,6921) = 9.3	9.5 10 ⁻⁵	
<i>Visu</i>	F(2,6921) = 930	≈ 0	
<i>Saccades:Visu</i>	F(4,6921) = 4.53	1.17 10 ⁻³	
AnglesSD~Saccades (TDU only)	F(2,2326) = 7.3	7.2 10 ⁻⁴	
<i>H₁ : AnglesSD_{moderate} > AnglesSD_{nosaccades}</i>			3.0 10 ⁻⁴
<i>H₁ : AnglesSD_{strong} > AnglesSD_{nosaccades}</i>			3.4 10 ⁻²
ResponseTime~Sparkle*Visu			
<i>Sparkle</i>	F(2,6921) = 4.59	1.02 10 ⁻²	
<i>Visu</i>	F(2,6921) = 673	≈ 0	
<i>Sparkle:Visu</i>	F(4,6921) = 1.6	17.2 10 ⁻²	
ResponseTime~Saccades*Visu			
<i>Saccades</i>	F(2,6921) = 33.5	≈ 0	
<i>Visu</i>	F(2,6921) = 682	≈ 0	
<i>Saccades:Visu</i>	F(4,6921) = 4.84	6.7 10 ⁻⁴	

Tableau 6.15. Analyse de variance et tests *a posteriori* (orientés) sur les erreurs d'angles et les temps de réponse en fonction des niveaux (nul, modéré ou fort) de scintillement ou de saccades. Le tableau donne les F-statistiques des tests sur les variances (ANOVA) et les p-valeurs associées à ces F-tests. Toujours pour l'ANOVA, il indique les degrés de liberté associés : degrés sur le nombre de modalités de perception et sur le nombre de mesures. Le nombre de modalité égal à 3 (degrés de liberté égal à 2) vient du fait que ces analyses de variance ont été réalisées également dans le cadre de la bimodalité visuelle+TDU. Les tests de rang associés sont orientés pour montrer la diminution des erreurs dans le cas du scintillement modéré et l'augmentation des erreurs en présence de saccades.

La diminution des temps de réponse est visible sur les *violin plots* présentés figure 6.16. La présence de scintillement comme de saccades permet de diminuer ces temps de réponses, ce que l'on observait déjà sur la figure 6.8.

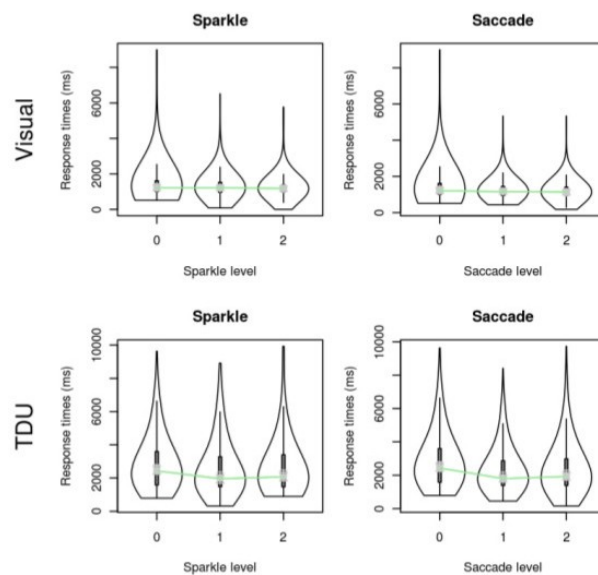


Figure 6.16. Effet du scintillement ou des saccades de l'image sur les temps de réponse.

Si ces résultats sont sans ambiguïté quand on se fie aux résultats des tests de variance et sur les rangs, il reste que lorsqu'on regarde les distributions (et les p-valeurs), les effets d'augmentation de la précision de perception des angles restent faibles au regard des effectifs (nombre de mesures) élevés que comporte cette étude. De plus, lorsque cette étude a été publiée, nous ne tenions pas compte d'un biais important dont nous n'avions pas connaissance : l'inversion du repère spatial par certains sujets et pour certains angles de stimulation particulièrement. Il semble donc qu'il faille proposer une nouvelle interprétation des résultats de cette première expérience après un nouveau traitement tenant compte de ce biais.

6.1.4.2. *Interprétation en deuxième instance*

À ce stade, nous invitons le lecteur à faire une incursion au chapitre 7.2 « Contradictions d'informations - inversion spatiale » et à revoir le chapitre 5, en particulier les points 5.3 à 5.6 concernant la recherche de mélanges gaussiens, de modèles et l'utilisation de modèles générateurs gaussiens pour réaliser les tests statistiques.

Si l'on regarde la forme des distributions des histogrammes 6.3, ou leur équivalent sous la forme de violin plots de la figure 6.14 B, mais aussi les nuages de points des figures 6.4 et 6.7 (sujet N°5), on s'aperçoit de la présence possible d'un mode dans les grandes erreurs. Nous avons finalement compris que ces erreurs importantes n'en étaient pas

forcément et pouvaient résulter d'un effet d'inversion haut-bas / avant-arrière du repère spatial des sujets (voir les sections 5.6 et 9.1). En particulier, les angles correspondants aux diagonales (45° ...) se retrouvent souvent inversés à 90°. Nous avons traité les résultats comme décrit dans la section 5.6 de façon à corriger (partiellement) cette inversion. La figure 6.17 montre les distributions théoriques (mélanges gaussiens) obtenues pour les conditions de contrôle, de scintillement modéré, de scintillement fort, de saccades modérées et de saccades importantes.

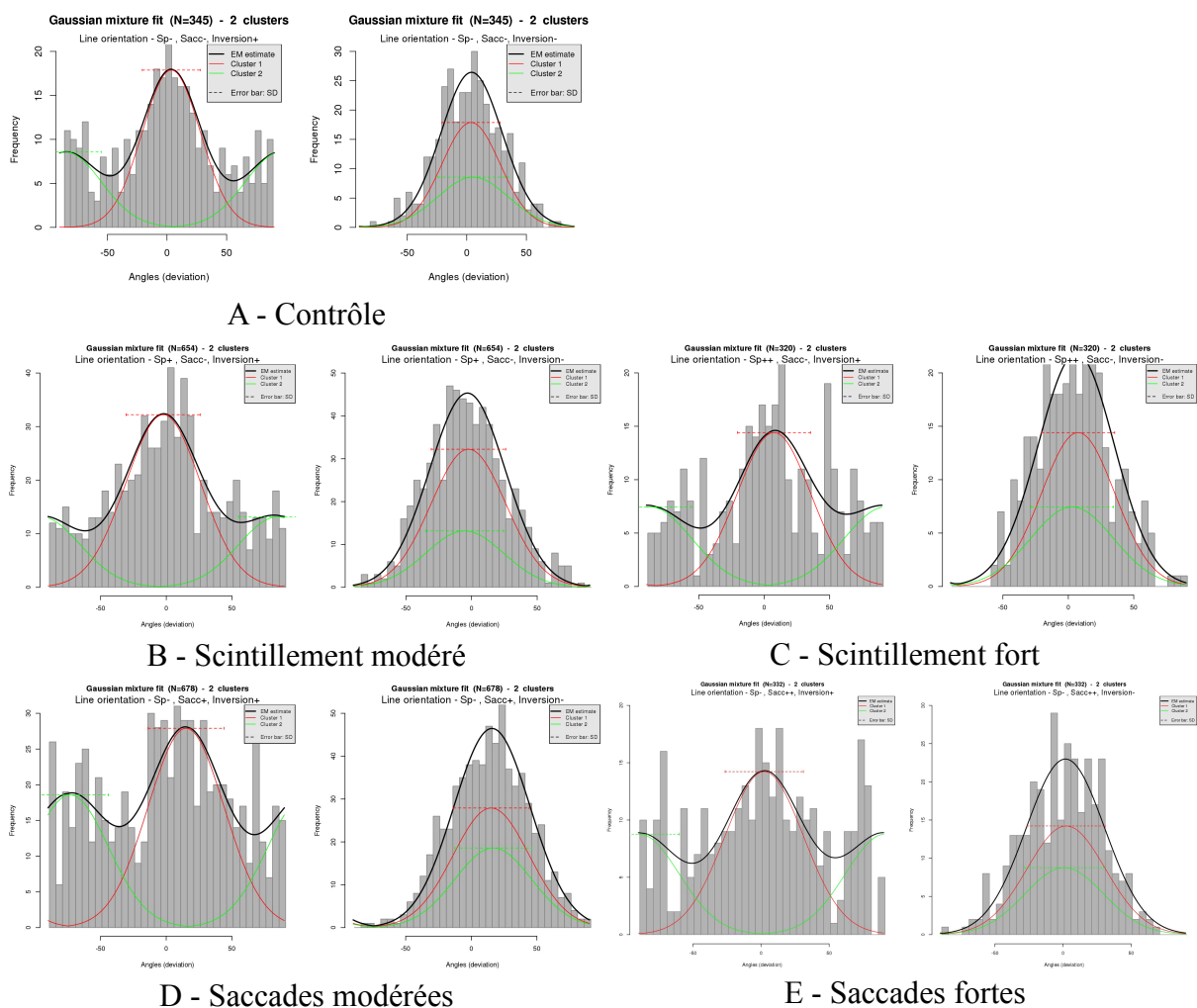


Figure 6.17. Traitement de l'inversion pour les erreurs d'angles dans différentes conditions de saccades et de scintillement (A à E). Chaque graphique montre à gauche les données d'origine et à droite les données traitées.

Le tableau 6.18 donne les valeurs des paramètres de ces distributions gaussiennes ainsi que la moyenne et l'écart type du mélange gaussien corrigé (valeur qui est utilisée pour comparer l'étalement des erreurs d'angles entre les différentes conditions).

Conditions		Moyenne (degrés)			Proportion (%)		Ecart Type (degrés)		
Scintillement	Saccades	G1	G2	G1+G2	G1	G2	G1	G2	G1+G2
-	-	3.6	-84.7	4.2	63.2	36.8	24.3	29.6	26.2
Modéré	-	-2.3	84.8	-3.1	70.1	29.9	28.4	29.7	28.5
Fort	-	7.5	-87.1	5.7	63.2	36.8	28.0	31.5	29.0
-	Modérées	15.1	-73.3	15.3	59.7	40.3	29.0	29.3	28.4
-	Fortes	2.35	-89.1	1.78	61.7	38.3	28.7	29.0	28.58

Tableau 6.18. Valeurs des paramètres des 2 gaussiennes G1 et G2 déterminées pour chaque condition. On indique aussi la moyenne et l'écart type de la distribution théorique obtenue avec le mélange G1+G2 des deux gaussiennes recentrées.

En ce qui concerne l'amélioration de la perception des orientations, étant donné ce traitement particulier (et critiquable) des données, les résultats sont clairs : ni les saccades, ni le scintillement ne diminuent les erreurs, ce que montrent les augmentations des écart types théoriques (de G1+G2). Des tests de comparaisons de variances faits entre les données contrôle corrigées et les données corrigées dans les différentes conditions de scintillement ou de saccades ne montrent aucune différence significative entre les erreurs en condition contrôle et en conditions de saccades ou de scintillement.

Si une différence est à noter dans ce tableau, c'est peut-être la diminution des inversions en conditions de scintillement modéré. Il est difficile néanmoins de dire si cette différence est significative, et encore plus de savoir à quoi elle serait due.

Cette première expérience est riche en enseignements: tous ceux décrits dans cette section 6.1 (en particulier, on montre que la présence de scintillement ou de saccades diminue les temps de réponse dans la modalité TDU ; on s'aperçoit aussi de l'effet d'inversion et on trouve un moyen possible pour la corriger) et tous les résultats concernant la bimodalité (section 7). Il apparaît toutefois qu'elle ne nous a pas permis d'identifier clairement en quoi des saccades ou un scintillement du signal pouvait améliorer un dispositif de substitution sensorielle. Nous avons vu aussi que la présence de saccades ou de scintillement diminuait les temps de réponse, mais il est probable que cette diminution ne soit pas due à une décision de réponse de la part du sujet sur la base d'une bonne perception de l'orientation du stimulus, mais simplement à une décision de réponse à cause de la sollicitation de réponse que provoque le signal vibrant (scintillement ou saccade).

La question était en effet mal posée dès le début : ajouter du bruit (scintillement ou saccades) sur un signal de très basse résolution (12x12 taxels) peut difficilement améliorer la qualité de perception des orientations des lignes tracées sur le TDU. Les saccades oculaires ne font pas mieux ! Elles permettent en revanche de ne pas perdre le signal (phénomène de fading) lorsque celui-ci est maintenu constant. Une autre expérience a donc été faite pour voir si les saccades ou le scintillement d'un signal permettent de le percevoir même après une longue période de stimulation.

6.2. Expérience 2 - Perception de l'orientation d'une ligne persistante

6.2.1. Objectif

Dans l'expérience précédente, on a vu que la présence de saccades ou de scintillement diminuait les temps de réponse. Nous avons cependant vu que cette première expérience n'avait pas un bon design pour mettre en évidence l'effet que des mécanismes inspirés de la vision, scintillement et saccades, pourraient avoir. Les saccades oculaires permettent en effet de maintenir la perception d'un signal et, en cela, de continuer à le percevoir correctement même après une longue stimulation, donc d'améliorer sa perception sur le long terme.

Le but de la présente expérience est de mettre en valeur l'influence bénéfique des deux mécanismes (micro-saccades et scintillement) sur le maintien de la perception pendant des périodes de stimulation importantes. Lors de la première expérience (section 6.1) le sujet devait répondre le plus vite possible après le début de la stimulation. Ses réponses se faisaient dans un temps court de l'ordre de 1 à 2 secondes, c'est-à-dire bien avant que le problème de fading n'apparaisse. Cette fois, on force des durées de stimulation longues, comprises entre 2 et 10 secondes, de façon à forcer l'apparition du fading, donc la perte progressive de la perception des orientations, et mettre en évidence l'importance des saccades ou du scintillement pour maintenir celle-ci.

6.2.2. Protocole

Au cours de cette expérience, on commence par afficher deux lignes d'intensité égale et croisées (Fig. 6.19) pendant un temps T , l'angle entre les deux lignes étant égal à 90° et la période d'affichage pouvant être de 2, 5 ou 10 secondes. A la fin de cette période, un signal

sonore (un « bip » fort) est émis, on fait disparaître une des deux lignes et on demande à l'utilisateur de dessiner la ligne qui est encore affichée sur sa langue (Fig. 6.20). Comme dans la première expérience, le sujet utilise le stylet pour répondre sur la tablette graphique.

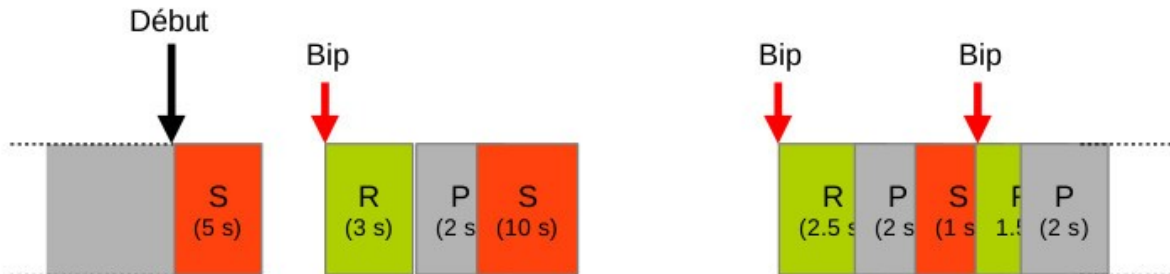


Figure 6.19. Exemple de séquence avec 3 stimulations (S) fixes (2, 5 ou 10 secondes) suivies de leurs temps de réponse (R) variables de 2 à 3 secondes environ, et de périodes de repos (P) fixes de 2 secondes.

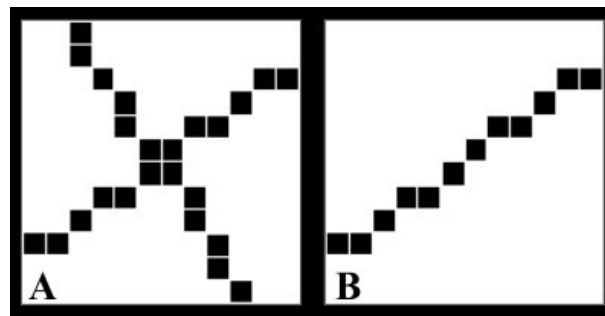


Figure 6.20. Stimuli tels qu'ils apparaissent aux sujets (A) Pattern de stimulation: deux lignes croisées. (B) Après une durée de 2, 5 ou 10 secondes, une des deux lignes disparaît. On demande au sujet de dessiner sur une tablette graphique la ligne qui reste affichée (réponse) le plus vite possible.

Les facteurs que l'on fait varier sont (i) la durée d'affichage des lignes croisées (3 valeurs de temps possibles : 2, 5 et 10 secondes), (ii) la présence ou l'absence de saccades ou du scintillement (3 conditions : contrôle, présence de scintillement modéré, présence de saccades modérées), et (iii) l'angle des tracés qui varie aléatoirement à chaque mesure entre 0° et 360° (10 tirages d'angles par sujet et par condition).

L'hypothèse est que l'affichage des lignes croisées pendant une longue durée (5 ou 10 secondes) va saturer la langue si ce signal ne scintille pas ou ne tremble pas (saccades). Si le scintillement ou des saccades empêchent la saturation de la langue, alors le sujet devrait être en mesure d'indiquer correctement quelle ligne continue d'être tracée parmi les deux lignes tracées initialement.

Sept sujets sains jeunes (âgés entre 21 et 33 ans) ont participé volontairement à cette étude. Ils étaient tous naïfs quant à l'objectif de l'expérience, et utilisaient le TDU pour la

première fois. Ils avaient les yeux bandés avec un loup pendant l'expérience. Une familiarisation avec les outils de la plate-forme et une explication du protocole est faite juste avant qu'on commence le scénario de l'expérience.

Pour chaque facteur (durée d'affichage et procédé de modification du signal) on mesure 10 valeurs. On mesure ainsi 3 temps x 3 modifications x 10 angles = 90 mesures par sujet pendant une séance. On réalise 2 séances par sujet, ce qui fait un total de 180 mesures par sujet à la fin de l'étude. Cela permet d'obtenir 10 mesures x 2 séances x 7 personnes = 140 points par facteur et par durée. Sur un total potentiel de 1260 points (7x90x2), nous avons fait en pratique 1219 mesures (certaines étaient invalides).

Estimation de la durée d'une session. Le nombre de mesures que nous souhaitons acquérir pour chaque sujet est suffisamment important pour que l'analyse statistique des données soit pertinente (140 points par facteur et par durée sur les trois sessions). Le nombre de mesures par sujet et par session est également élevé (90 mesures) mais cela ne veut pas dire que les sujets ont passé beaucoup de temps en expérience. Pour chaque session et pour chaque sujet, on a 90 mesures qui durent chacune [(10 secondes x 10 mesures) + (5 secondes x 10 mesures) + (2 secondes x 10 mesures)] x 3 modifications = 8.5 minutes de stimulation pour toute la session sans inclure les périodes de repos et les temps de réponse variables entre chaque point de mesure. Les mesures s'enchaînent avec une pause de 2 secondes après chaque mesure. De plus, un repos d'au moins 30 secondes est fait toutes les 30 mesures. On a donc $90/30 = 3$ repos de 30 secondes, soit 1.5 minute de repos auxquelles il faut ajouter 90×2 secondes de pause entre chaque point = 3 minutes de temps intermédiaires, soit un total de 4.5 minutes de temps de repos. Le temps de réponse occupe 2 secondes en moyenne pour chaque mesure soit 90×2 secondes = 3 minutes de temps de réponse. Le temps total de l'expérience est donc égal à : Temps de stimulation + Temps de repos + Temps de réponse = $8.5 + 4.5 + 3 = 16$ minutes. Cette expérience durera donc en moyenne 15 à 20 minutes par sujet par session. Ceci est supportable par l'ensemble des sujets : sains, malvoyants ou aveugles.

A noter tout de même qu'*a posteriori*, après l'achèvement de cette série d'expériences, on a remarqué que le temps théorique de l'expérience n'était pas exactement respecté pour quelques sujets. Pour certains, lors de la première session, on a eu des durées de l'ordre de 25

minutes (en adéquation avec la prévision de 16 minutes sur la base de 2 secondes de temps de réponse) ; pour d'autres nous avons eu des durées plus longues allant jusqu'à 42 minutes (réponses bien plus longues, plus de pauses). De plus, ces durées ont réduit lors de la deuxième session, ce qui est cohérent avec le fait que les sujets savaient alors utiliser le TDU.

6.2.3. Résultats

Dans cette expérience, on regarde l'effet du scintillement ou des saccades dans le maintien de la perception de l'orientation d'une ligne. Pendant la première phase de stimulation qui dure 2, 5 ou 10 secondes, le sujet perçoit 2 lignes croisées et peut en déterminer les orientations respectives, mais il ne peut savoir laquelle persistera après la stimulation avec les 2 lignes. Si le temps de stimulation est court, on espère que le fading ne sera pas encore présent ; ainsi le sujet devrait détecter la bonne orientation facilement. Si au contraire le temps de stimulation est long, on peut supposer que l'effet de fading (désensibilisation des récepteurs de la langue) sera important et gênera, voire empêchera complètement la détection de la ligne restante dont le signal ne sera pas perçu ou perçu avec une trop grande imprécision.

La figure 6.21 montre sous la forme de violin plots (graphiques de densité à noyau avec ici une largeur de bande de 10°) la répartition des erreurs d'angles en fonction de la durée de stimulation avec les lignes croisées (2s, 5s ou 10s) et de la présence ou de l'absence de scintillement ou de saccades (fréquence de 50%). Le graphique indique également les modes principaux estimés par l'estimateur de mode du noyau de Parzen (valeur qui maximise l'estimateur de densité du noyau). De plus, comme décrit dans la section 5.6 et utilisé précédemment dans la section 6.1.4.2, nous avons cherché à corriger le phénomène d'inversion du repère spatial qui provoque des erreurs d'angles décalées vers les grandes erreurs (le décalage le plus fréquent étant de 90°). La figure montre donc ainsi les moyennes des 2 gaussiennes détectées pour chaque distribution. Ces sont les moyennes des valeurs moyennes obtenues par bootstrap sur 3000 tirages par distribution. Ces paramètres (modes principal et moyennes des gaussiennes estimées) indiquent donc là où se concentre l'erreur. La médiane, sa variation entre le contrôle et la condition étudiée, ainsi que les 1^{er} quartile (Q25) et 3^{ème} quartile (Q75) sont également indiqués.

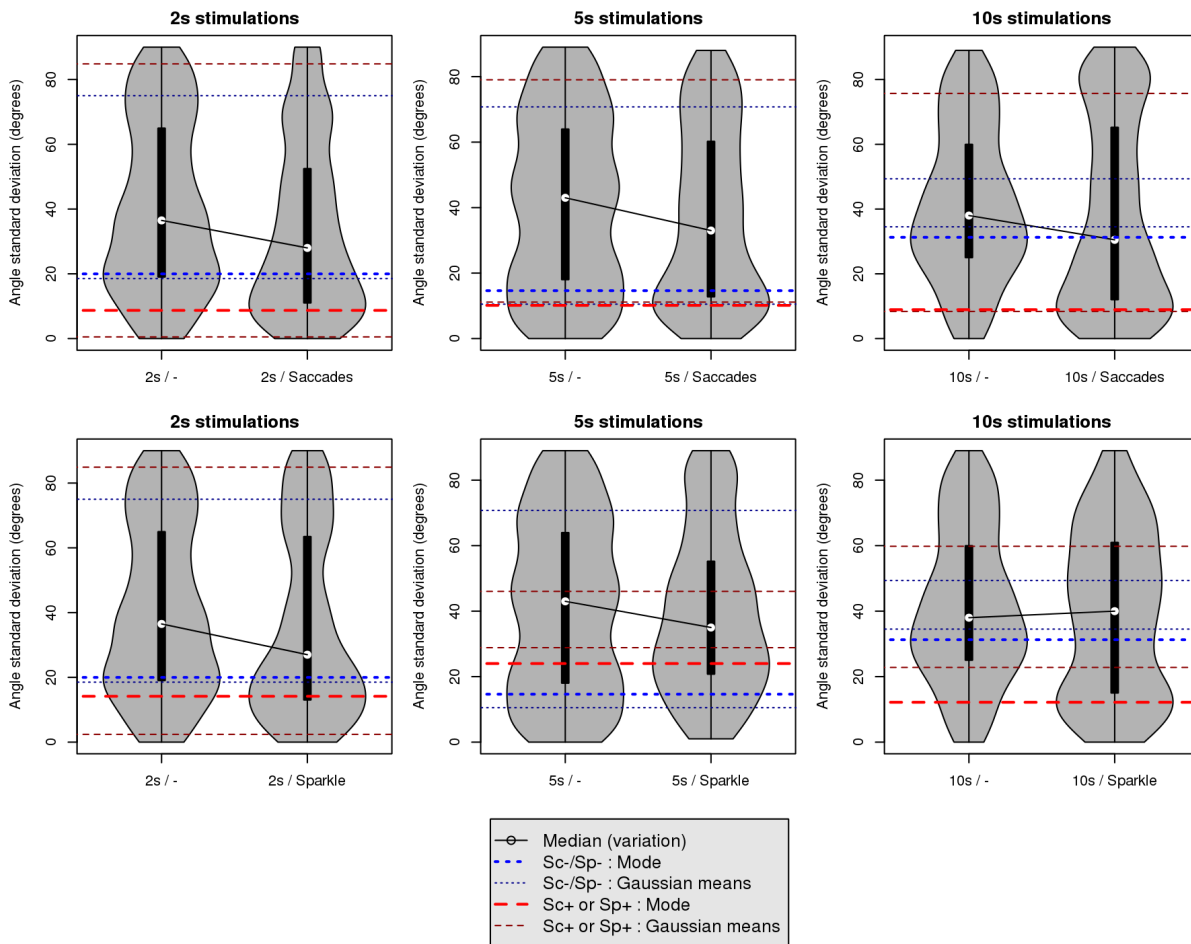


Figure 6.21. Distribution des erreurs d'angles en fonction de la durée de stimulation avec les lignes croisées et en fonction de la présence ou de l'absence de saccades. Les conditions contrôle correspondantes sont reportées sur chaque graphique. Les distributions sont tracées avec des violin plots pour montrer la forme des distributions et la présence de modes multiples. Les modes principaux sont repérés par des lignes pointillées épaisses tandis que les moyennes des 2 gaussiennes détectées le sont par des lignes pointillées fines. La médiane, sa variation, ainsi que les 1^{er} quartile (Q25) et 3^{ème} quartile (Q75) sont également indiqués.

Sans tenir compte des grandes erreurs partiellement imputables au phénomène d'inversion, on voit clairement que l'erreur augmente fortement dans les conditions de contrôle. À 2 secondes, le mode principal apparaît autour de 20° alors qu'il est de 35° pour 10 secondes de stimulation. Pour 5 secondes de stimulation, le graphique est moins clair et montre une distribution assez étalée des erreurs. Le mode principal et les moyennes des 2 gaussiennes de cette distribution contrôle ne sont ainsi pas bien estimés. En présence de saccades, on voit clairement aussi que le mode principal se maintient autour de 10° (ce qui est déjà plus faible que celui des conditions contrôle à 2 secondes). Avec du scintillement, le résultat est semblable : le mode principal est maintenu vers les petites erreurs (10°) pour des temps de stimulation de 2s ou de 10s. À 5 secondes de stimulation, c'est un peu moins clair,

mais l'erreur reste faible avec un mode principal aux alentours de 20°. L'observation de la médiane montre aussi un maintien vers les petites erreurs en condition de saccades ou de scintillement (sauf pour les conditions de scintillement à 10s de stimulation). La médiane n'est cependant pas un descripteur complètement fiable de ces distributions d'erreur d'angle car celle-ci incluent des erreurs importantes dues à l'inversion qui tirent la médiane vers les grandes valeurs.

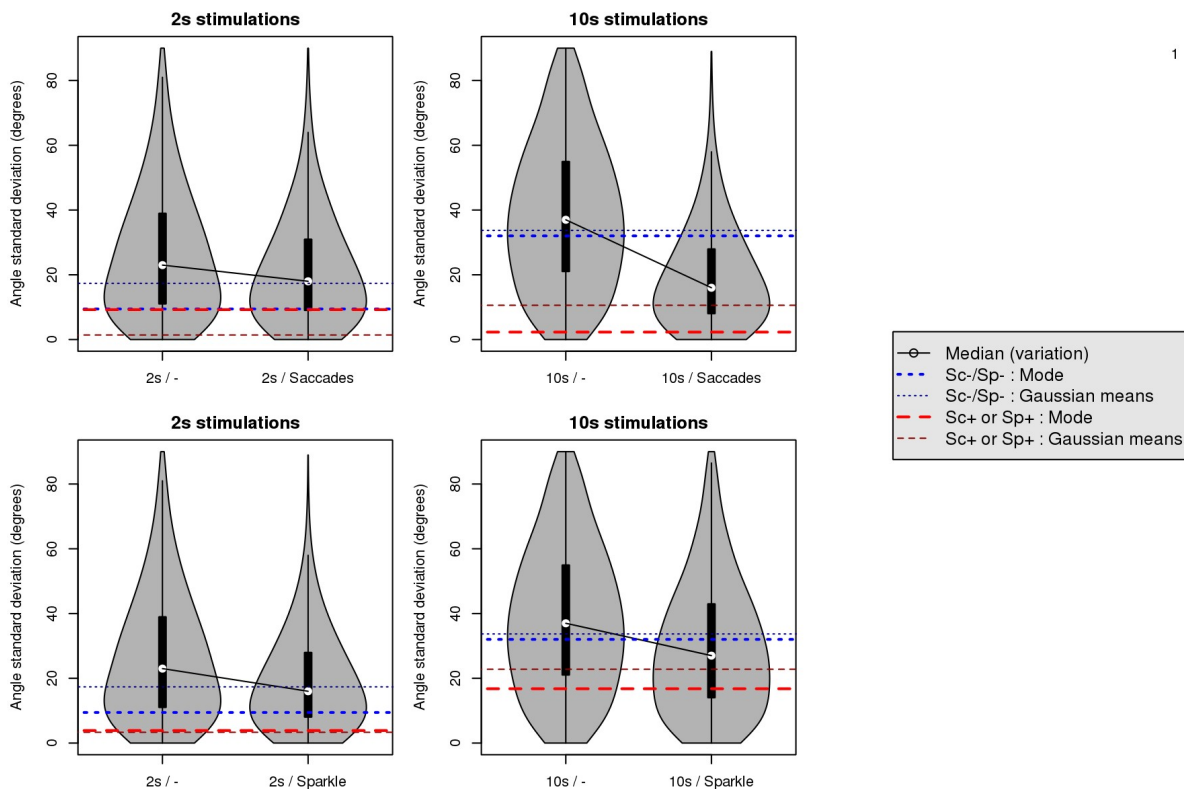


Figure 6.22. Distribution des erreurs d'angles générées à partir d'un modèle de mélange gaussien corrigeant le phénomène d'inversion du repère spatial. Ces distributions sont tracées en fonction de la durée de stimulation (courtes de 2s et longues de 10s) avec les lignes croisées et en fonction de la présence ou de l'absence de saccades. Les conditions contrôle correspondantes (générées également) sont reportées sur chaque graphique. Les distributions sont tracées avec des violin plots pour montrer la forme des distributions modélisées. Les modes sont repérés par des lignes pointillées épaisses tandis que les moyennes des distributions modélisées le sont par des lignes pointillées fines.

Si maintenant on s'appuie sur le traitement de l'inversion proposée en section 5.6 et déjà utilisé précédemment, donc en acceptant l'approximation que les réponses (erreurs d'angles) correspondent à un mélange de 2 gaussiennes, l'une centrée sur les réponses sans inversion, l'autre correspondant aux réponses quand le sujet inverse son repère spatial, on obtient grâce au modèle générateur corrigé la figure 6.22. Dans cette figure, on ne montre que les résultats obtenus pour les stimulations courtes (2 secondes) pour lesquelles on ne s'attend

pas à ce que le fading pose problème, et pour les stimulations longues (10 secondes) pour lesquelles on est sûrs que la langue du sujet a perdu toute sensibilité vis-à-vis de la perception de l'orientation de la ligne restante. On voit qu'à 2 secondes de stimulation, que l'on soit en condition de contrôle, de saccades ou de scintillement, les erreurs sont faibles avec une densité maximale (modes et moyennes) des données comprise entre 5 et 15°. Après 10 secondes de stimulation, en conditions de contrôle, l'erreur a considérablement augmenté : le mode et la moyenne se situent autour de 35° (38° et 33° respectivement) tandis que tant avec les saccades (surtout) et le scintillement, l'erreur est maintenue faible, aux alentours de 10° en conditions de saccades et aux alentours de 20° en conditions de scintillement.

Peu de tests statistiques s'offrent à nous pour associer une p-value (un niveau de significativité) à ces expériences, bien que les résultats bruts soient assez parlants d'eux-mêmes. Ces distributions présentent en effet plusieurs modes importants, au moins 2 correspondant aux réponses avec ou sans inversion, et leurs valeurs sont ainsi très étalées de 0 à 90° malgré un effectif assez faible par distribution (entre 130 et 140 valeurs). Cela affecte considérablement les résultats de tout test statistique qui, rappelons-le mesure systématiquement la significativité d'une variation (variance). Plus ennuyeux encore, cette variance factorielle supposée doit pour autant être significativement supérieure à la variance résiduelle intrinsèque à chaque distribution. Du fait de ces 2 modes de part et d'autre du spectre des erreurs d'angles (0° d'une part et 90° d'autre part), la variance résiduelle est très grande dans nos mesures.

Un paramètre (de position) assez intéressant (bien qu'affecté par le mode de l'inversion) pour décrire naïvement les distributions, est la médiane. Le test associé est le test de Mood. Pour maximiser la puissance du test, nous avons considéré uniquement les stimulations à 2s et à 10s. Une autre possibilité de comparaison est une comparaison globale des rangs des distributions par un test de Wilcoxon, Mann et Withney. Dans ces conditions, nous obtenons le tableau suivant (Tab. 6.23). Toutefois, il ne faut pas oublier que, théoriquement, ces tests a posteriori n'ont aucune valeur s'ils ne présupposent pas que la variance factorielle observée est supérieure à la variance résiduelle mesurée.

Les résultats sont peu interprétables pour les raisons évoquées ci-dessus.

Echantillon 1	Echantillon 2	ANOVA II	Test sur la médiane (Mood)	Test sur les rangs (Wilcoxon)
2s, contrôle	2s, saccades	0.0497	0.323	0.0311
2s, contrôle	2s, scintillement	0.135	0.0942	0.0696
10s, contrôle	10s, saccades	0.276	$2.51 \cdot 10^{-5}$	0.118
10s, contrôle	10s, scintillement	0.413	0.091	0.406

Tableau 6.23. Analyse statistique des résultats bruts. Le tableau contient les p-valeurs des ANOVA II (non séquentielle), tests sur la médiane et sur les rangs.

Si maintenant on se fie aux tests statistiques (ANOVA et test a posteriori sur les rangs) faits sur des données simulées à partir du modèle de mélange gaussien corrigé pour l'inversion, les résultats sont sans appel (Tab. 6.24) : on observe systématiquement un effet de la présence de saccades ou de scintillement (au seuil de 5%), mais ce qui est frappant est la chute importante des p-valeurs (augmentation de la significativité) à 10 secondes de stimulation par rapport à celles obtenues pour 2 secondes de stimulation. Ces résultats (p-valeurs) ont été obtenus par bootstrap sur des échantillons simulés de 140 valeurs (même effectif que celui des échantillons réels). A chaque itération du bootstrap, on réalise l'étude statistique, ANOVA et test de Wilcoxon, Mahn et Withney pour lesquels on enregistre les p-valeurs. Le tableau 6.24 indique les médiane et mode des distributions de ces p-valeurs (celles-qui ressortent le plus fréquemment).

Echantillon 1	Echantillon 2	ANOVA II		Test sur les rangs (Wilcoxon)	
		p-value (median)	p-value (mode)	p-value (median)	p-value (mode)
2s, contrôle	2s, saccades	0.0276	0.00198	0.0532	0.00391
2s, contrôle	2s, scintillement	0.000534	0.000119	0.0029	0.00037
10s, contrôle	10s, saccades	$9.18 \cdot 10^{-15}$	$8.64 \cdot 10^{-16}$	$2.81 \cdot 10^{-13}$	$3.83 \cdot 10^{-14}$
10s, contrôle	10s, scintillement	0.00265	0.000191	0.00342	0.000477

Tableau 6.24. Analyse statistique sur les valeurs simulées (d'effectif 140). Le tableau contient les p-valeurs (médiane et mode de la distribution des p-valeurs obtenues par bootstrap) des ANOVA II (non séquentielle) et des tests sur les rangs.

7 Multi-modalité et Contradictions d'informations

7.1. Contexte général

Notre perception (sensorielle) est multi-modale: dans la vie courante, percevons les informations sur notre environnement par l'intermédiaire de plusieurs sens. Dans la plupart des situations on se sert effectivement de plus d'un sens pour accomplir une tâche. Par exemple, pour conduire une voiture, on regarde la route, les voitures, les piétons et les panneaux de signalisation ; en parallèle on entend la voix du GPS qui nous guide et qui nous suggère de tourner vers une direction ou de sortir d'une voie d'autoroute et la radio qui annonce plein de mauvaises nouvelles passionnantes. Pour arriver à destination on se sert de la vision et l'audition en même temps : ces deux modes de perception sont ainsi fusionnés. Ajoutons à cela nos pensées et autres rêveries, ainsi que notre propre connaissance de l'environnement, du chemin à parcourir, etc. Nos pensées conscientes et nos expériences s'ajoutent aussi à cette multi-perception, souvent de façon harmonieuse (notre connaissance du trajet concorde avec les indications du GPS ; la musique passant radio ne nous dérange pas dans notre conduite), parfois de façon contradictoire (le GPS indique un chemin que je ne connais pas ; le téléphone sonne et cela me perturbe dans ma conduite, surtout si je réponds).

Parfois, et pour la même tâche, on se retrouve dans une situation de conflit entre les modalités de perception. Pour illustrer la situation, restons sur l'exemple de la conduite ; le stationnement en particulier. Dans ce cas de figure, la vision et l'audition collaborent pour bien se garer. Dans ce cas, on peut énumérer trois possibilités : (i) suivre la méthode classique fondée uniquement sur la vision dans les rétroviseurs et à travers le pare-brise arrière ; (ii) en faisant confiance aux capteurs de recul avant et arrière et écouter le bip sonore qui nous indique la distance par rapport aux obstacles (bordures, voitures, ...) ; (iii) alterner entre les deux modes (vision + audition) en essayant de les faire concilier ce qui est idéal mais en pratique bien plus lent. Les personnes habituées vont suivre la troisième stratégie. Dans ce cas il y a souvent prédominance soit de la vision sur l'audition (courant) soit l'inverse (si l'on fait confiance au système de recul aveuglement) ; la fusion n'est pas immédiate et requiert un apprentissage. D'autre part, il se peut que le conducteur ait à gérer des conflits entre les deux

modes de perception. Cela arrive quand l'information sonore nous indique qu'on est trop près d'un obstacle et que notre regard confirme l'inverse. À ce moment, la décision d'avancer ou de reculer est délicate parce qu'elle est influencée par deux informations contradictoires. La distance entre la voiture et un obstacle quelconque peut être mal estimée par les deux sens de perception. D'une part le son qui provient des capteurs de distance et l'information visuelle d'une autre part. L'estimation de la distance est étroitement liée à la précision des capteurs (qu'on estime grande *a posteriori* (pas forcément lors de l'action)), ainsi qu'à notre champ de vision et notre acuité visuelle (qu'on estime souvent sans faille : nous nous fions à nos sens, notre expérience de l'environnement ...). Cet exemple anodin dont le risque se limite souvent à un peu de tôle froissée est considérablement accru en a2ro et astronautique notamment. De très nombreux crashes d'avion sont dus à une contradiction entre les sens et l'expérience des pilotes d'une part et les informations des instruments de mesure (altimétrie, vitesse ...) de l'appareil : le pilote automatique se fie uniquement à ces mesures ; les pilotes dans certaines conditions (brouillard, fort vent de face...) pensent ces instruments faussés et débranchent le pilote automatique. Ne parlons pas des contradictions entre les repères spatiaux fournis par ces instruments et ceux d'un pilote désorienté, malmené par les mouvements brusques de son appareil.

Il existe d'autres exemples de contradiction, en particulier entre l'audition et la vision. Cet exemple se rencontre fréquemment lorsqu'on regarde une série américaine doublée sur une chaîne française : il y a contradiction, surtout si le doublage n'est pas bon, entre la voix du doubleur francophone et les lèvres de l'acteur américain.

Les travaux de recherche autour de la contradiction inter-sensorielle ont commencé il y a longtemps, notamment en ce qui concerne l'étude de contradiction entre la vision et l'audition. « When we hear lips and see voices » (*quand on entend les lèvres et on regarde les voix*) ; c'est le titre du papier présenté par [McGurck McDonald 1976]. Les auteurs montrent que la combinaison des deux voies de perception, visuelle et auditive, produit une information combinée des deux. Au cours d'une série d'expériences, on demande à des sujets de regarder des vidéos dans deux conditions différentes: avec ou sans le son (mute). Ces vidéos représentent une personne prononçant de façon répétitive le phonème « ba » ou le phonème « ga ». De même, des bandes audio séparées sont écoutées avec ou sans la vidéo. Enfin, les vidéos peuvent aussi être vues avec une bande son qui ne correspond pas (prononciation du

phonème « ba » et écoute du son « ga »). On demande à chaque fois aux sujets ce qu'ils perçoivent. Les auteurs mettent alors en évidence un effet, maintenant connu sous le nom d'effet Mc Gurck & Mc Donald, dans lequel un processus de fusion s'opère entre les modes de perception Visuel et Auditif : la combinaison du son « ba » avec le visage d'un interlocuteur qui prononce « ga » est fréquemment perçue « da ». L'étude montre qu'en amont de la fusion phonétique il y a un processus de liaison (*binding*) qui contrôle la combinaison de l'image et le son. Ce processus peut être bloqué ou réduit en cas d'incohérence audiovisuelle (*conditional binding process*) (voir notamment la revue de littérature [J.L. Schwartz et al 2012]).

Il existe un autre exemple bien décrit de multi-perception, cette fois-ci entre la vision et le sens tactile. Quand une personne explore un objet avec la main en le regardant, le toucher et la vision fournissent tous deux des informations permettant d'estimer les propriétés de l'objet (sa forme, sa rugosité, *etc*). Pour juger de la taille, de la forme ou de la position de l'objet, la vision domine fréquemment dans ce cas de double perception visuo-haptique. Néanmoins, dans d'autres circonstances la perception est affectée par le sens haptique [Rock and Victor, 1964]. En 1991, Loomis et al ont traité les similarités d'une tâche de reconnaissance visuelle et tactile. Au cours de cette étude les sujets doivent reconnaître des images (imprimées par la technique de fac-similé en relief) d'objets de la vie quotidienne en utilisant deux modes de perception différents avec un champ de perception réduit. D'une part, dans le cas de la perception tactile, l'aspect exploratoire de l'expérience exige l'utilisation d'un doigt (l'index) en premier lieu, ensuite de deux doigts (l'index et le majeur) pour reconnaître l'objet. D'autre part, dans le cas de la perception visuelle, le sujet manipule un stylet qui permet de déplacer une fenêtre glissante limitant le champ de vue pour explorer l'image (ou deux fenêtres glissantes pour correspondre avec l'exploration à 2 doigts). Ils concluent que l'efficacité de reconnaissance tactile des images est étroitement liée à l'importance du champ de vision [Loomis et al., 1991].

Une étude de modélisation du comportement dans le cas de contradiction entre le sens haptique et la vision a été réalisée par M.O. Ernst and M.S. Banks, 2002. Les auteurs proposent une expérience de mesure de perception multimodale (Vision/Toucher) (figure 7.1). Les sujets portent des lunettes stéréoscopiques équipées de cristaux liquides observent un stimulus visuel à travers du verre réfléchissant sur lequel peut être projeté une image bruitée (avec un écran CRT situé au dessus). Ceci permet de créer une disparité binoculaire affectant

ainsi la vision stéréoscopique. En ce qui concerne la perception haptique, le protocole expérimental propose une barre virtuelle suggérée par deux dispositifs de retour de force placés sur le pouce et l'index respectivement. Afin de quantifier la combinaison des deux sens, les auteurs mesurent la variance associée au deux sens pour l'estimation de la hauteur de la barre. Ils ont ensuite élaboré un modèle bayésien intégrant l'efficacité (probabilité) de perception selon la fiabilité des informations fournies par les 2 sens, les paramètres du modèle maximisant la vraisemblance par rapport aux observations. Le modèle combine les informations visuelles et haptiques d'une manière similaire à ce que fait le système nerveux. La dominance de la vision s'observe ainsi par une valeur de la variance plus faible que celle calculée pour la perception haptique.

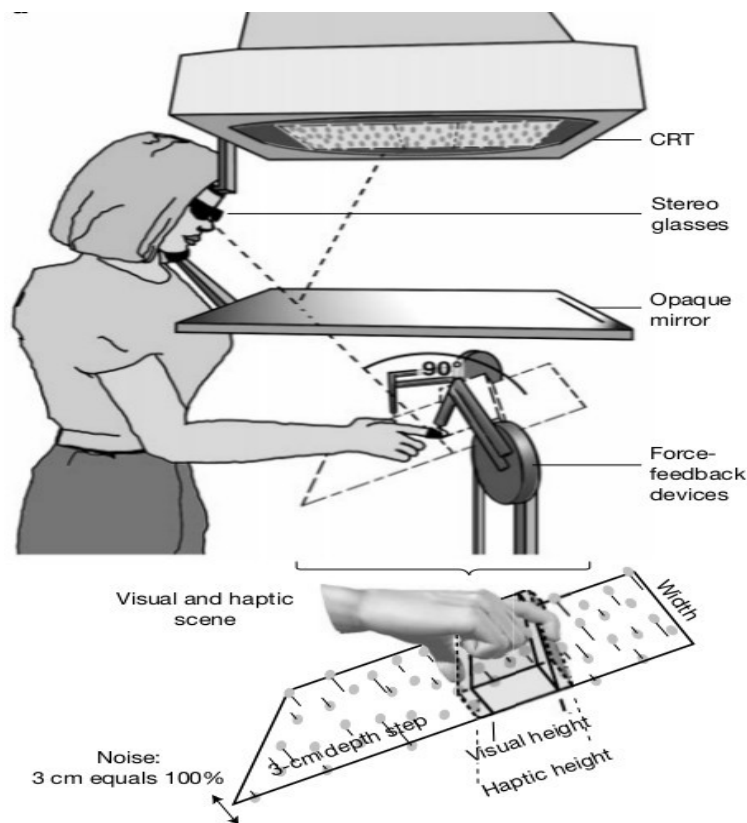


Figure 7.1. Expérience de contradiction visuo-haptique [Ernst and Banks, 2002]

Les stimuli peuvent provoquer des illusions pour un sens seul comme la vision. Cependant, les conflits de la perception multimodale impliquent facilement des illusions. Notons qu'on peut supposer l'inter-sensorialité biaisée dans le sens où il existe des modes de

perception plus dominants (comme la vision) que d'autres. Ceci est particulièrement important lorsqu'une décision est en jeu, qui plus est lorsqu'on demande une réponse rapide.

Nous n'avons pas trouvé, dans la littérature, d'étude de contradiction entre la perception tactile par électro-stimulation sur la langue et la vision. Plusieurs études rapportées dans l'état de l'art décrivent le domaine d'application de la substitution visuelle avec ce que ce mode de perception peut apporter en tant que suppléant de la vision. Cependant, l'information apportée par ce mode sensoriel peut être en contradiction (ou provoquer une surcharge cognitive) avec d'autres modes de perception (vision notamment), mais également avec notre représentation de l'environnement (dans notre mémoire).

A travers une série d'expériences décrites dans ce chapitre, on cherche ainsi à montrer que, bien que la vision soit infiniment plus efficace et précise que la perception tactile par électro-stimulation sur la langue, cette dernière a une influence sur un mode de perception dominant comme la vision en cas de bimodalité sensorielle (contradictoire ou non).

A travers les 3 sous sections qui suivent, nous allons décrire les observations et mesures que nous avons effectuées pour étudier les effets liés à la multi-modalité perceptive et aux contradictions.

7.2. Inversion spatiale : effet miroir

7.2.1. Problématique

Dans ce chapitre, nous ne décrivons pas une expérience à proprement parler, mais un phénomène que nous avons observé et qui s'est manifesté dans toutes nos expériences, un effet d'inversion spatiale.

La perception visuelle permet d'acquérir des informations spatiales: où se situent le haut, le bas ... Notre position ainsi que notre posture influencent cette perception spatiale. Quand on est debout devant un miroir, les directions gauche et droite sont inversées. Cela pose parfois problème quand on essaye d'effectuer une tâche de précision, par exemple se raser la barbe.

Un autre exemple se rencontre dans l'apprentissage de certains instruments de musique, par exemple la guitare. Se référer au haut et au bas d'une guitare est un vrai souci pour les nouveaux élèves qui se posent des questions comme « Où est le haut (de la guitare) dans un repère horizontal (la guitare est horizontale) ? » ou « Parle-t-on du haut spatial (là où se trouve le point de plus haut de la guitare à un moment donné, point qui varie selon la position que donne l'élève à son instrument), ou du haut du manche, vers le sillet de tête ? » (Fig. 7.2).



Figure 7.2. Exemple de la guitare qui illustre le conflit du haut et du bas dans un repère horizontal.

On peut citer également le problème lié à l'utilisation des outils informatiques avec en particulier la navigation avec la souris ou sur la surface d'un touchpad (Fig. 7.3). Les enfants ont du mal à définir les directions (haut, bas, gauche et droite) quand ils bougent la souris dans tout les sens. On rencontre le même problème avec les manettes de jeux vidéo et joysticks. Pousser la souris (ou le joystick) vers l'avant fait monter la souris sur l'écran (vers le haut). On doit donc faire correspondre le repère avant-arrière du périphérique d'entrée avec celui haut-bas du périphérique de sortie (écran). Pour cela (comme pour la guitare, et probablement pour tout problème requérant une mise en adéquation de 2 repères différents, y compris le TDU), un apprentissage est requis. On peut poser la question (qui a du être répondu, en particulier dans le cadre de recherches en navigation astro et aéronautique) des difficultés rencontrées dans des cas où le sens de l'équilibre est perturbé (apesanteur, ...). On remarque par contre qu'avec les tablettes tactiles les repères sont plus clairs et l'utilisation en est plus facile. Dans ce cas en effet, il n'y a pas existence de 2 repères différents.

Le phénomène d'inversion dans notre cas est dû à la différence des repères entre celui du moniteur (placé verticalement dans nos expériences, mais qui pourrait être placé horizontalement, avec le sujet sur le dos, dans le cas d'expériences dans un scanner IRM) et

celui du TDU (qui est en perpendiculaire à celui du moniteur). Dans cette situation, nous avons constaté que la majorité des sujets perçoivent la plupart du temps une information inversée par rapport à ce que nous attendions : notre correspondance 'naïve' était en effet celle d'un repère avant-arrière de la langue avec respectivement le repère haut-bas de l'écran. Il s'est avéré que, pour la plupart des sujets, au cours des expériences, l'avant de la langue correspondait au bas de l'écran (et respectivement l'arrière de la langue avec le haut de l'écran) (Fig. 7.4). Les sujets donnaient ainsi des réponses (angles) inversées par rapport à l'angle de stimulation. Sans apprentissage, ce choix est arbitraire et dépend probablement des capacités de repérage dans l'espace des différents sujets. Ce phénomène n'est toutefois pas systématique puisque pour certains sujets les réponses ne sont pas inversées (Figs. 7.5 et 7.6). Dans nos mesures, ce sont 2 sujets, habitués de questions de repérage dans l'espace et habitués à de nombreux périphériques informatiques qui n'inversent pas.

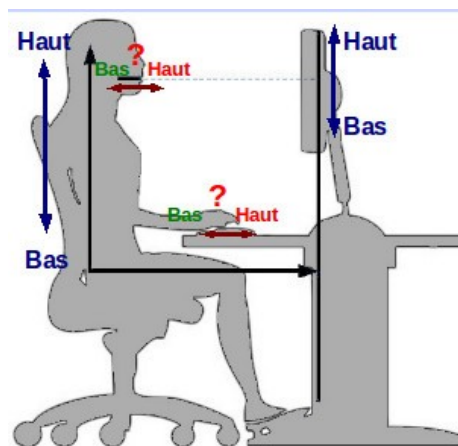


Figure 7.3. Représentation d'un sujet au cours d'une expérience. Le plan de posture de la personne est parallèle avec le plan du rendu graphique (écran). En revanche, le plan de la langue et le plan des déplacements du périphérique de saisie (souris, stylet graphique, joystick) sont orthogonaux avec celui de l'écran et de la posture.

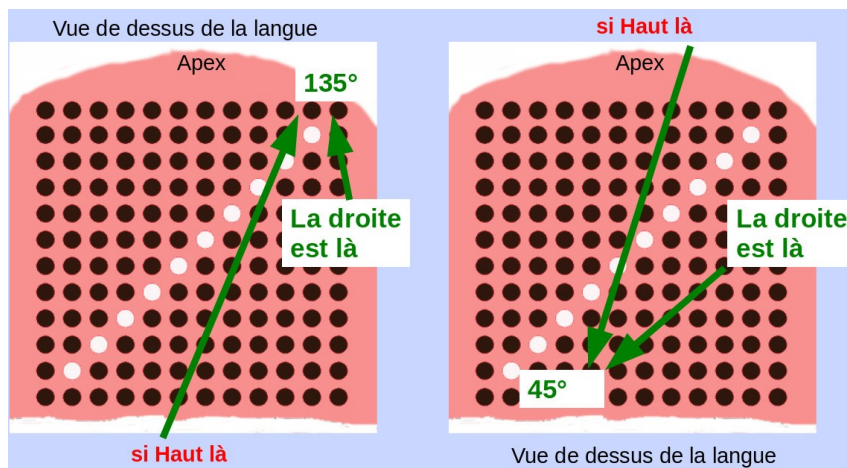


Figure 7.4. Effet d'inversion de la perception de l'orientation d'une ligne 2D. La langue est représentée par la région colorée en rose ; elle est vue de dessus. La droite réelle de la langue est indiquée (elle en change pas). En revanche, le haut de la langue, tel que défini par le sujet correspond soit à l'arrière de la langue (à gauche), soit à l'avant (à droite). Une même ligne est ainsi perçue comme étant soit à 135° (gauche) soit à 45° (droite).

Au cours de nos expériences d'orientation de la ligne 2D, ce problème s'est manifesté particulièrement en ce qui concerne les orientations diagonales. Les lignes très proches de la verticale ou de l'horizontale restent en effet "inchangées" en cas d'inversion. C'est ce qu'on observe notamment dans les figures 6.4 et 6.5. La figure 7.5 montre le diagramme des angles de réponse (et non pas des écarts par rapport à la réponse attendue) par rapport aux angles de stimulation. Dans ce diagramme, on voit clairement les réponses non inversées (plus denses et matérialisées par des lignes vertes) et les réponses inversées (moins denses et matérialisées par des lignes rouges).

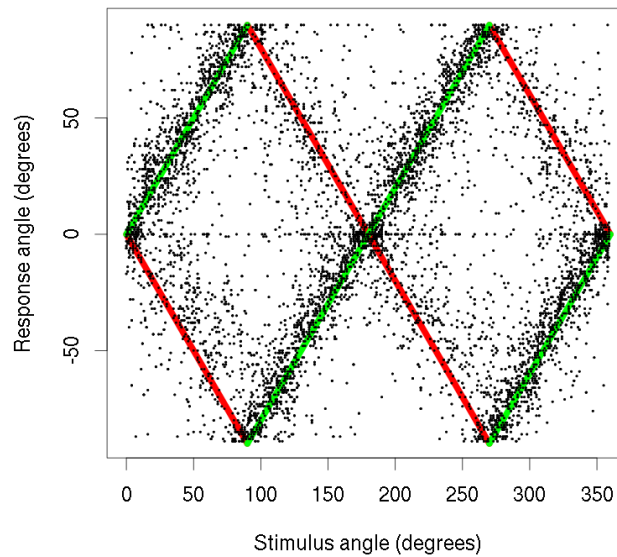


Figure 7.5. Diagramme des angles de réponses en fonction des angles de stimulation. L'ensemble de ces points provient de la première expérience (section 6.1) tous sujets, tous modes de perception (VDU, TDU, bimodal) et toutes conditions confondus. Les lignes vertes indiquent les réponses attendues non inversées, tandis que les lignes rouges correspondent aux réponses attendues en cas d'inversion. La figure 7.6 montre les données obtenues sur le VDU et sur le TDU séparément.

Comme nous l'avons vu, nous pouvons proposer un traitement pour tenter de corriger ce phénomène (voir sections 5.3 à 5.6) et nous permettre d'interpréter nos résultats « en l'absence de celui-ci ». Cependant, comme nous l'avons dit, il ne s'agit pas d'une solution miracle car elle repose sur un certain nombre d'hypothèses discutables. En particulier, rechercher 2 gaussiennes correspondant aux réponses « normales » et aux réponses « inversées » présuppose d'une part un comportement gaussien des réponses autour d'une réponse attendue (réponse moyenne théorique) et surtout le fait que l'inversion soit uniforme quel que soit l'angle de stimulation, ce qui n'est pas le cas comme le laissent voir les figures 6.4 et 6.5.

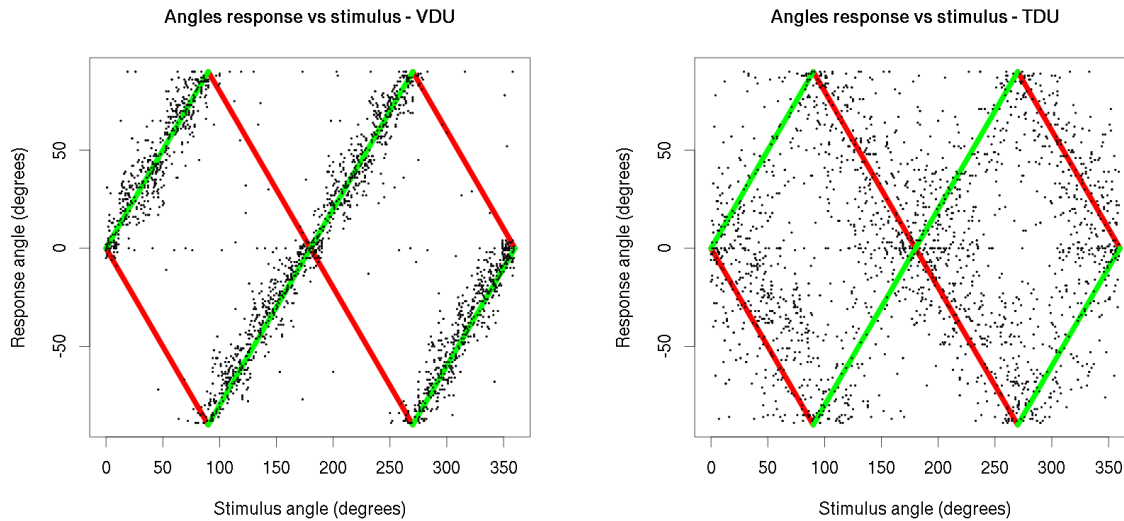


Figure 7.6. Diagramme des angles de réponses en fonction des angles de stimulation. L'ensemble de ces points provient de la première expérience (section 6.1) tous sujets et toutes conditions confondus. Les lignes vertes indiquent les réponses attendues non inversées, tandis que les lignes rouges correspondent aux réponses attendues en cas d'inversion. A gauche, le diagramme pour le VDU ; à droite pour le TDU. On voit qu'il n'y a aucun cas d'inversion dans le mode de perception visuelle. En revanche, on voit que la majorité des réponses avec le TDU sont inversées.

Il est raisonnable d'associer ce phénomène à un cas de multi-modalité et de contradiction d'information. Les sujets intègrent en effet une double information, l'une provenant de leur propre expérience et représentation de l'environnement et de leurs repères spatiaux, l'autre provenant du dispositif de stimulation.

Ce phénomène mériterait d'être davantage étudié, d'une part pour savoir pourquoi certains sujets inversent les repères et d'autres pas (rôle de facteurs liés à la représentation dans l'espace) ? Quel est l'effet de l'apprentissage pour limiter l'inversion ? Pourquoi se phénomène n'est il pas systématique ?

De plus, il serait envisageable de concevoir de nouvelles expériences pour éviter toute confusion dans l'interprétation des résultats à cause de ce phénomène. Ainsi, dans la 2eme expérience décrite, concernant la persistance (voir section 6.2), nous pourrions limiter l'étude à celle des angles de stimulation verticaux et horizontaux. Nous ne pouvons nous contenter des résultats déjà acquis en nous restreignant aux angles de stimulation très proches de la verticale ou de l'horizontale car il n'y en a pas assez. Il faudrait acquérir de nouvelles données expérimentales pour une étude statistique significative. Une autre possibilité serait de

concevoir une nouvelle expérience dans laquelle on proposerait non pas des lignes croisées, mais des lignes parallèles horizontales ou verticales suffisamment éloignées sur le TDU pour que l'on distingue facilement une ligne d'une autre ... sauf en cas de perte de la sensibilité.

7.3. Multi-modalité TDU et VDU

7.3.1. Objectif

Au cours de l'expérience N°1 visant à montrer un éventuel effet des saccades ou du scintillement dans l'amélioration de la qualité de perception de lignes orientées dans le plan du TDU, nous avons comparé les résultats en termes d'erreurs d'angles (entre les réponses et les angles de stimulation) et de temps de réponses entre le TDU et le VDU (visualisation sur un moniteur). Nous en avons profité pour mesurer les erreurs d'angles et les temps de réponse en cas de bimodalité perceptive : les sujets percevaient la ligne orientée dans la «même direction» sur le TDU et sur le VDU. L'affichage sur le VDU était semblable à l'image montrée figure 6.1. A noter cependant que parler de la « même direction » est difficile dès lors que certains sujets peuvent inverser certains angles, ceux-ci se retrouvant dès lors en contradiction entre le TDU et le VDU (bien qu'intentionnellement affichés de la même manière par l'expérimentateur qui a choisi les repères spatiaux associés à la langue et à l'écran. maintenir la perception de la ligne (image) avec ces facteurs. C'est ce que fait l'expérience 2.

Il s'agit donc dans cette expérience de comparer la qualité de perception et les temps de réponse entre 3 conditions expérimentales : VDU seul, TDU seul et bimodalité TDU+VDU. Bien que nous ayons les données pour les différentes conditions de saccades ou de scintillement, nous ne nous concentrerons pas particulièrement dessus. Nous avons en effet vu que les saccades et le scintillement n'avaient pas un rôle important dans cette expérience. De plus, il est question ici de savoir si une modalité de perception domine, en l'occurrence la vision comme on pourrait s'y attendre, ou si l'on peut s'attendre à une certaine fusion des 2 modalités de perception. Nous voulons savoir si les résultats sont très proches de ceux obtenus dans le cas de la vision seule (VDU) où si les angles et les temps de réponse sont influencés par la perception d'un signal provenant du TDU.

7.3.2. Protocole

Le protocole est identique à celui décrit dans la section 6.1.2 à ceci près que l'on mesure 70 points supplémentaires par sujet et par session correspondants aux conditions de bimodalité TDU+VDU.

7.3.3. Résultats

Erreurs d'angles.

La figure 7.7 montre sous la forme d'histogrammes les résultats bruts des réponses sous forme d'erreurs d'angles (affichées entre 0 et 90°) mesurés dans le cas de perception visuelle (VDU seul), sur le TDU seul (résultats déjà montrés, Fig. 6.3) ou en condition de bimodalité TDU+VDU, et en fonction de l'absence ou de la présence de scintillement ou de saccades. On constate qu'en conditions de bimodalité, les erreurs augmentent légèrement. Néanmoins, il semble que celles-ci correspondent à des erreurs très importantes (>45°) et doivent ainsi correspondre à des inversions. Ceci est confirmé par les figures 7.8, 7.9 et 7.10.

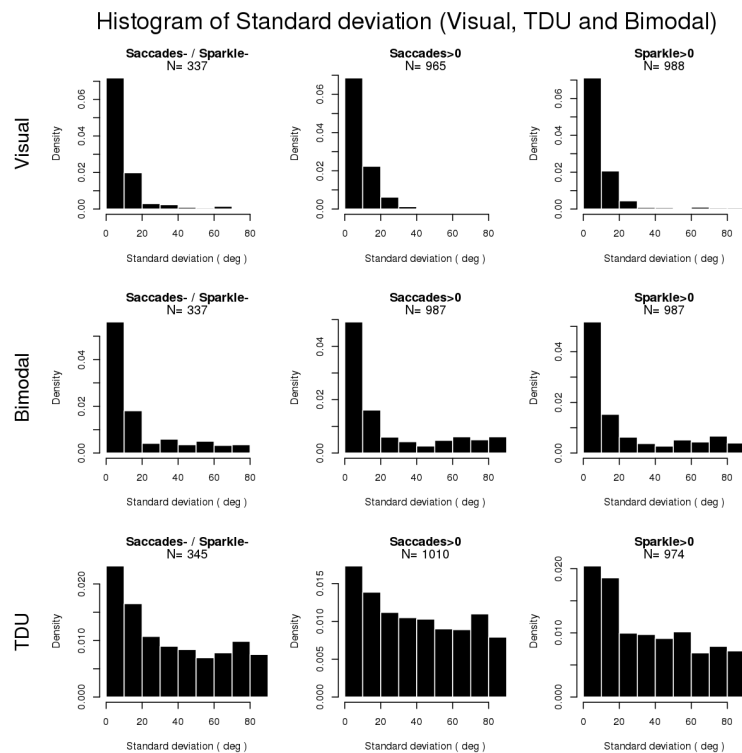


Figure 7.7. Histogrammes des erreurs d'angles (standard deviations) dans les modalités visuelles (VDU), bimodale (TDU+VDU) ou TDU et selon l'absence ou la présence de scintillement ou de saccades (tous niveaux confondus : 0.3, 0.6 et 1.0). L'effectif N est indiqué pour chaque expérience

Les figures 7.8 (résultats bruts) et 7.9 (résultats traités par une médiane mobile ; voir section 6.1) montrent en effet que, contre toute attente, en conditions de perception bimodale VDU+TDU certains sujets inversent leurs réponses pour certains angles. C'est aussi ce qui est observé sur le diagramme des réponses en fonction des angles de stimulation (Fig. 7.10) sur lequel on observe que, bien que la majorité des points ne soient pas inversés (et correspondent probablement davantage aux réponses fondées sur le référentiel du VDU), un bon nombre toutefois le sont, mais dans tous les cas, il s'agit de réponses précises, comparables en précisions à celles obtenues avec le VDU (Fig. 7.6). Cette inversion a peu de chances de provenir du VDU. Cela signifie que certains sujets ont parfois choisi le TDU comme référence de perception plutôt que le VDU (dont on pourrait supposer qu'il s'agisse, en toute situation, du mode de perception dominant). De plus, ces réponses faites « selon le TDU » sont inversées par rapport à celles qui devaient être affichées sur le VDU. Il faut donc préciser qu'au cours de cette expérience, aucun sujet ne s'est senti en situation de contradiction d'information (dans laquelle les angles sur le VDU et sur le TDU diffèrent) au contraire de l'expérience décrite dans la section suivante (section 7.4). Les sujets avaient l'impression que

les angles de stimulation étaient bien les mêmes, mais que leur précision en était affectée (Fig. 7.11). Enfin, bien que le nombre de points inversés soit assez faible, il semble au vu des figures 7.8, 7.9 et 7.10 que l'erreur d'angle faite sur les réponses inversées soit nettement plus faible que celle observée dans le cas du TDU. Cela pourrait laisser présager une fusion des informations d'angles provenant du TDU (provoquant l'inversion) et du VDU (donnant la précision).

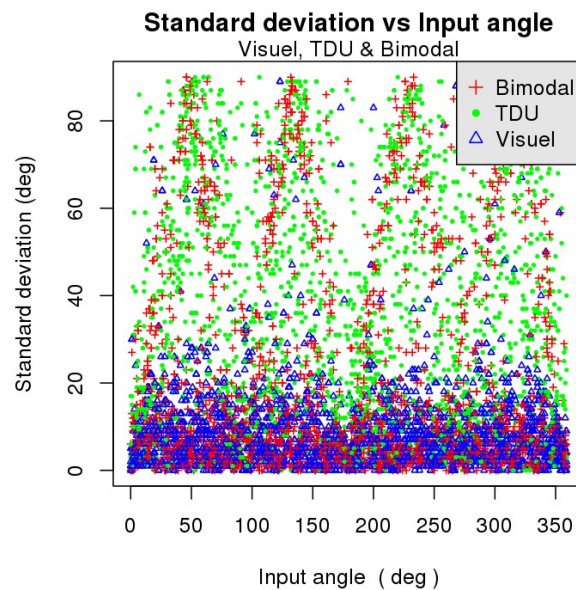


Figure 7.8. Nuages de points des erreurs de réponse sur les orientations (angle standard deviation) en fonction des angles de stimulation (input angles), de la modalité de perception (visuelle, bimodale ou TDU) toutes conditions de saccades ou de scintillement confondues. En conditions de bimodalité, on observe, comme pour le TDU, une variation périodique des erreurs en fonction du stimulus, avec des maxima sur les diagonales. En revanche, les erreurs sont moins importantes (peu de dispersion autour des oscillations), contrairement au TDU.

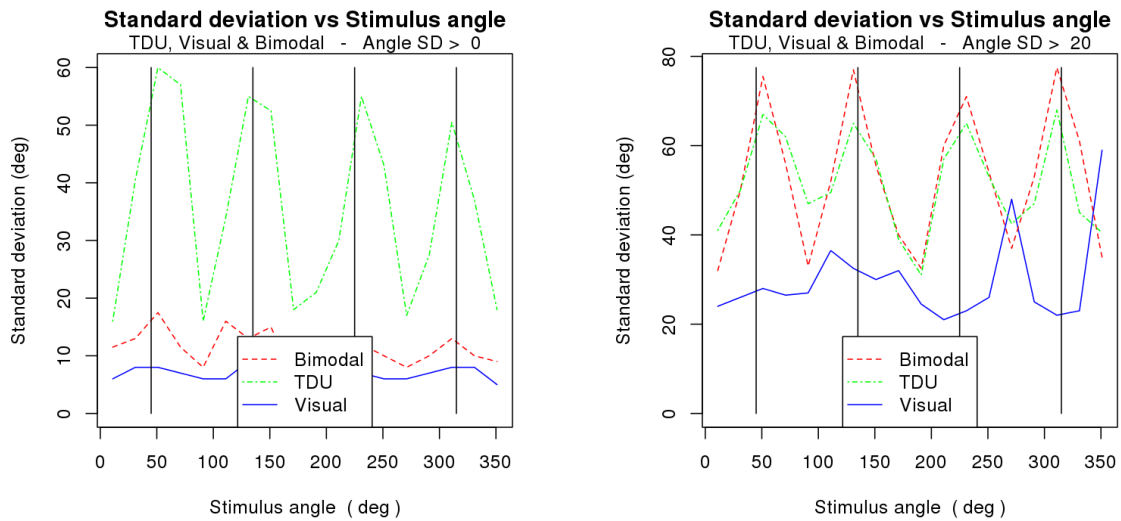


Figure 7.9. Médiane mobile (filtre médian avec une fenêtre de 15°) des erreurs de réponse sur les orientations (angle standard deviation) en fonction des angles de stimulation (input angles) et de la modalité de perception (visuelle, bimodale ou TDU), toutes conditions (scintillement, saccades ou non) confondues. A gauche, on prend en compte toutes les réponses entre 0° et 90° ; à droite, on ne prend que les erreurs supérieures à 20° (donc erreurs comprises entre 20° et 90°). Ces courbes font ressortir la variation des erreurs de réponses en fonction de l'angle de stimulation. On observe qu'en conditions de bimodalité, les oscillations n'apparaissent nettement que pour des erreurs importantes. En fait, le fait que les oscillations soient très nettes sur le graphique de droite est simplement dû au cumul de toutes les valeurs d'erreur au dessus de 20° . Dans le cas de la bimodalité, les oscillations ne semblent pas (ou peu) liées à l'erreur, mais uniquement à l'angle de stimulation (voir Fig. 7.10).

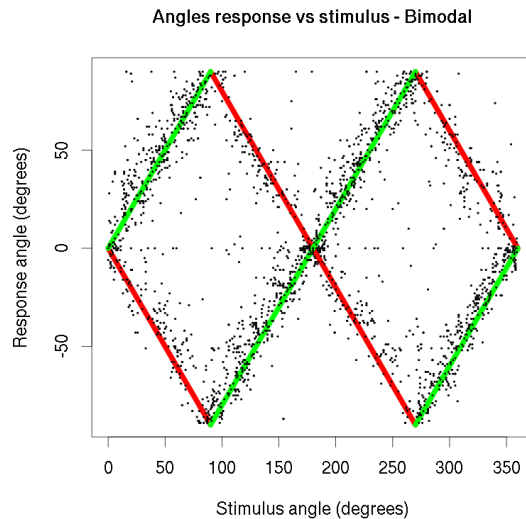


Figure 7.10. Diagramme des angles de réponses en fonction des angles de stimulation dans le cas de la perception bimodale. L'ensemble de ces points provient de la première expérience (section 6.1) tous sujets et toutes conditions (scintillement et saccades) confondus. Les lignes vertes indiquent les réponses attendues non inversées, tandis que les lignes rouges correspondent aux réponses attendues en cas d'inversion.

Ces résultats sont corroborés par ceux d'un questionnaire que nous avons fait concernant le ressenti des sujets vis-à-vis de la qualité de perception des orientations des lignes sur le VDU, le TDU et en conditions de bimodalité perceptive (figure 7.11). Comme on peut s'y attendre les notes sont maximales avec le VDU, elles sont médiocres avec le TDU, mais elles ne sont aussi que moyennes en cas de bimodalité perceptive. En perception bimodale, Les sujets avaient la sensation d'être moins précis, de moins bien être capables de répondre avec précision aux angles de stimulation. Cependant, il faut remarquer qu'en réalité, la précision des sujets n'a pas diminué par rapport à une perception avec le VDU. Le manque de précision ressenti provient probablement d'une augmentation de la charge cognitive nécessaire pour intégrer les informations provenant à la fois du VDU et du TDU. Cette augmentation de la charge cognitive peut être mise en évidence avec la mesure des temps de réponse.

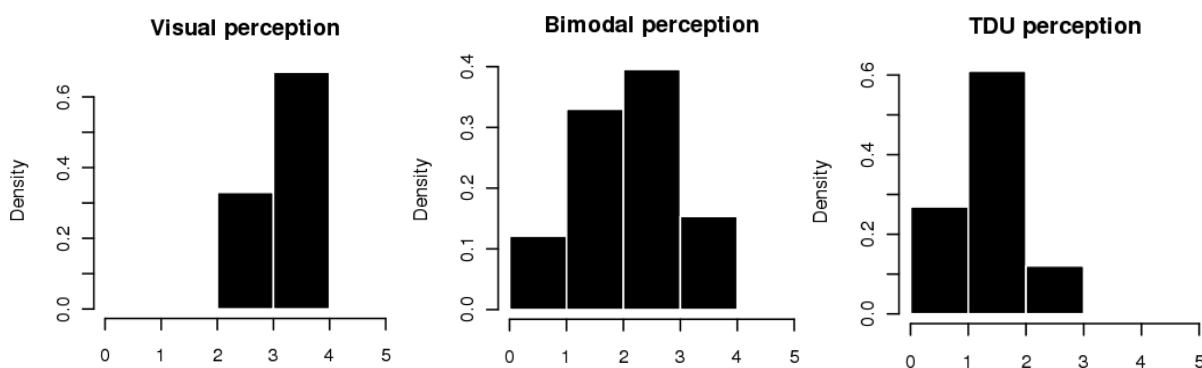


Figure 7.11. Résultats du questionnaire sur le ressenti des sujets en terme de précision de perception des angles dans les différentes modalités de perception (VDU, bimodale et TDU). Les notes sont comprises entre 0 (très mauvaise, voire aucune, perception des orientations) et 4 (excellente perception des orientations).

Temps de réponse.

En ce qui concerne les temps de réponse en situation de bimodalité, on observe (Fig. 7.12) que les sujets mettent plus de temps à répondre qu'avec le VDU et moins qu'avec le TDU. On observe également que comme pour le VDU, les temps de réponse sont légèrement plus dispersés mais se limitent principalement à un pic compris entre 1.5 et 2s (500 ms de plus qu'avec le VDU et 800 ms de moins qu'avec le TDU). On constate aussi que les saccades comme le scintillement semblent jouer un léger effet en limitant la dispersion des temps de réponse, comme pour le TDU.

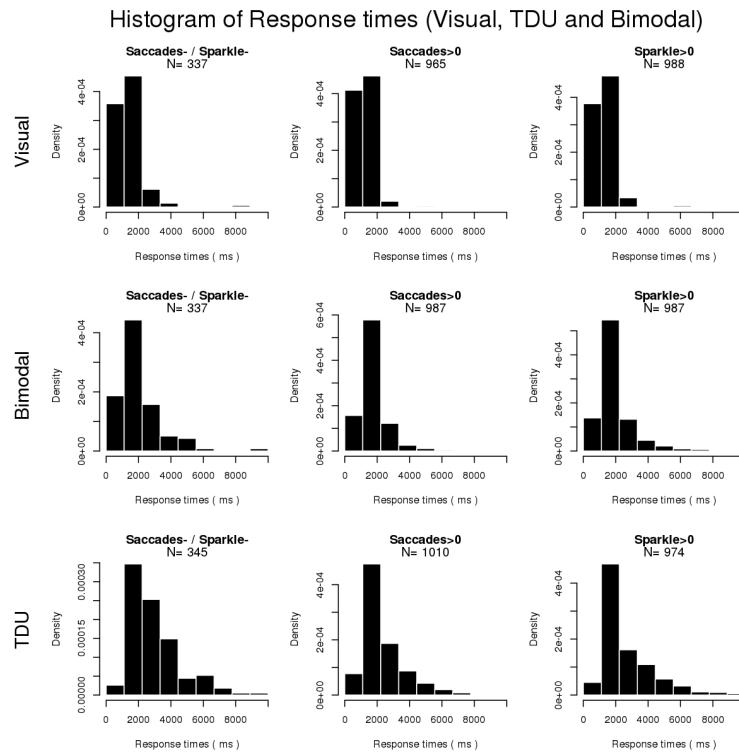


Figure 7.12. Histogrammes des temps de réponses (en ms) dans les modalités VDU, TDU ou bimodale et selon l'absence ou la présence de scintillement ou de saccades (tous niveaux confondus : 0.3, 0.6 et 1.0). L'effectif N est indiqué pour chaque expérience.

Le tableau 7.13 donne la moyenne, la médiane et la dispersion (écart-type) des temps de réponse avec le VDU, le TDU ou en bimodalité perceptive, en l'absence de saccades et de scintillement, ou en présence de l'un des deux.

	Sans saccades ou scintillement			Présence de saccades ou scintillement		
	Moyenne	Médiane	Écart type	Moyenne	Médiane	Écart type
VDU	1465.2 ms	1210 ms	949.6 ms	1267.3 ms	1180 ms	542.0 ms
Bimodal	2056.4 ms	1660 ms	1363.4 ms	1842.6 ms	1550 ms	1018.3 ms
TDU	2859.3 ms	2430 ms	1578.4 ms	2450.1 ms	1910 ms	1475.0 ms

Tableau 7.13. Paramètres (moyenne, médiane, écart type) des distributions observées des temps de réponse avec le VDU, le TDU et en bimodalité, en l'absence ou en présence de saccades ou de scintillement.

Les paramètres de position (médiane comme moyenne) ainsi que le paramètre de dispersion utilisé (écart type) confirment l'observation déjà faite dans la section 6.1 sur l'effet des saccades ou du scintillement sur les temps de réponse : ceux-ci sont plus courts et moins dispersés en présence de saccades ou de scintillement. Cette remarque peut maintenant être étendue à la bimodalité TDU+VDU. Encore une fois, une explication de leur effet serait de

« presser » les sujets à répondre, la vibration bien perceptible du signal les incitant davantage à répondre vite.

7.4. Contradictions d'informations

7.4.1. Objectif

Cette expérience vise à mettre en évidence un effet de contradiction dans les réponses données pour une expérience semblable à l'expérience de bimodalité (sensibilité à l'orientation d'une ligne en deux dimensions affichée à la fois sur le moniteur (VDU) et sur le TDU présentée section 7.3), mais dans laquelle les angles de stimulation ne sont pas les mêmes entre les deux modalités de perception : il y a donc contradiction d'information entre les deux modalités et le sujet est amené au mieux à choisir une des deux informations, au pire à combiner ces deux informations. Cette expérience vise à mettre en évidence un effet de fusion d'informations semblable à celui décrit par McGurck et McDonald [McGurk and MacDonald, 1976]. On rappelle que cet effet décrit section 7.1 est un phénomène perceptif qui montre une interférence entre l'audition et la vision lors de la perception de syllabes. Il met en évidence que la perception du langage est multimodale (http://fr.wikipedia.org/wiki/Effet_McGurk) et fait intervenir la fusion des informations auditives et visuelles. Dans notre cas, il s'agirait d'une contradiction entre 2 modalités de type visuelles, à supposer que la perception via le TDU soit de ce type. Nous pensons que cette expérience est décisive, son intérêt étant de montrer que la langue peut être un vrai organe de perception de la spatialité (voire de perception visuelle), dont l'efficacité se hisse à un niveau comparable à celui de la vision.

7.4.2. Protocole expérimental.

Au cours de cette expérience de contradiction on mesure comme dans l'expérience de bimodalité les temps de réponse et l'angle perçu. Le protocole est donc identique entre les deux expériences. La différence par rapport à l'expérience de bimodalité vient du fait qu'aléatoirement on décide que certaines informations sont en contradiction tandis que d'autres non (valeurs de référence). Ainsi, dans certains cas la stimulation par une ligne avec un certain angle se fait sans contradiction et est identique à celle qui serait faite dans l'expérience de bimodalité, tandis que dans d'autres cas on introduit un décalage fixe (de plus

ou moins 90° , 75° , 60° ou 45°) entre les lignes tracées sur le TDU et celles indiquées sur le VDU, la référence sur laquelle est calculée l'erreur de perception faite par le sujet restant toujours l'angle de la ligne tracée sur le VDU.

Nous avons réalisé 2 séries d'expériences : une première dans laquelle nous ne nous intéressions qu'aux contradictions de $\pm 90^\circ$; une deuxième dans laquelle nous avons étudié divers degrés de contradiction ($\pm 90^\circ$, 75° , 60° ou 45°).

La première série d'expériences a été faite avec 6 sujets et comportait 100 mesures par sujet. En pratique, cette expérience commençait par 10 mesures sans contradiction (comme dans l'expérience décrite en section 6.1) pour habituer le sujet à des réponses « normales », suivies de 90 mesures comportant 50 contradictions à 90° et 40 mesures sans contradiction, la nature de la stimulation (contradiction ou non) étant aléatoire au fil de l'expérience. Nous avons ainsi récolté plus de 600 mesures (précisément 601) au cours de cette série d'expériences.

La deuxième série d'acquisitions a été faite sur 2 séances à l'aide de 17 sujets, 8 femmes et 9 hommes âgés de 20 à 36 ans. Comme dans la première expérience, à chaque séance nous avons d'abord acquis 10 points sans contradiction suivis de 150 mesures mêlant aléatoirement et 70 contradictions de 0° , et 80 contradiction à $\pm 45^\circ$ (20 points : 10 points à -45° et 10 à $+45^\circ$), $\pm 60^\circ$ (20 points), $\pm 75^\circ$ (20 points) et $\pm 90^\circ$ (20 points). Nous avons ainsi recueilli environ 320 mesures par sujet au cours de cette expérience, soit près de 5500 mesures (précisément 5479).

Sur l'ensemble de ces expériences, nous avons ainsi récolté 6080 mesures comprenant 3046 points sans contradiction (0°), 984 mesures avec $\pm 90^\circ$ de contradiction, et plus de 680 mesures pour les chacune des contradictions de $\pm 75^\circ$, $\pm 60^\circ$ et $\pm 45^\circ$.

Les instructions étaient les mêmes que pour l'expérience sur la bimodalité : les sujets devaient répondre le plus rapidement possible ce qu'ils pensaient avoir perçu (l'angle de stimulation). Lors de la première session, les sujets savaient qu'ils allaient recevoir un stimulus sur le VDU et le TDU en même temps, mais ils n'étaient pas prévenus que certaines stimulations seraient contradictoires. Comme les 10 premières stimulations se faisaient sans contradiction pour chaque sujet, leur comportement était semblable à celui qu'ils avaient lors de l'expérience de bimodalité. En revanche, dès la première stimulation avec contradiction, lors de la première séance, tous les sujets ont retiré la nappe de stimulation de leur bouche et

nous ont dit qu'il y avait un problème (une ou 2 mesures supplémentaires étaient alors réalisées pour compenser celles-ci, mal faites).

Au vu des résultats de la première expérience (section 6.1) et parce que les temps de stimulation et de réponse attendus étaient courts, nous n'avons pas appliqué de saccades ni de scintillement dans ces expériences.

7.4.3. Résultats

7.4.3.1. Analyse globale des résultats

Erreurs d'angles.

Les données brutes sont présentées dans le tableau 7.14 et dans la série de figures 7.15 à 7.18. Les données d'intérêt sont d'une part l'angle de stimulation sur le VDU, les angles de stimulation sur le TDU pouvant être égaux ou différent d'un angle additionnel de $\pm 45^\circ$, $\pm 60^\circ$, $\pm 75^\circ$, $\pm 90^\circ$ et l'angle de réponse donné par le sujet et, d'autre part, les temps de réponse. Nous avons éliminé les données pour les temps de réponse supérieurs à 3 secondes.

ID	...	VDU	TDU	Add	Answer	Time
1	..	164	89	-75	-58	980
1	..	27	27	0	25	900
1	..	70	145	75	55	1000
1	..	162	117	-45	-70	850
..
10	..	18	-42	-60	-33	1500
10	..	159	159	0	-41	1260
..

Tableau 7.14. Échantillon des résultats de l'étude sur les contradictions d'angles entre le VDU et le TDU. Le tableau indique notamment la valeur *Add* de la contradiction (0° , $\pm 45^\circ$, $\pm 60^\circ$, $\pm 75^\circ$, $\pm 90^\circ$), l'angle de stimulation sur le *VDU* et sur le *TDU*. Il indique aussi la réponse en terme d'angle (compris entre -90° et 90°) et en temps de réponse (en ms).

Sur la figure 7.15, on constate que sans contradiction, la très grande majorité des données est très semblable à ce qu'on obtenait dans l'expérience présentée section 6.1 : la

plupart des erreurs sont faibles (ligne verte) et une partie est inversée (ligne rouge). Avec une contradiction de $\pm 90^\circ$ on voit clairement que les données se répartissent plutôt uniformément, avec une erreur assez faible soit sur la ligne verte indiquant la référence d'angle correspondant au VDU (0° de contradiction), soit sur les lignes bleues indiquant la référence d'angle correspondant au TDU ($+ ou - 90^\circ$ de contradiction). Le nombre de données est différent (3046 points sans contradiction contre 984 avec $\pm 90^\circ$ de contradiction), mais l'effet de la contradiction reste tout de même extrêmement marqué (au niveau des lignes bleues).

Pour les contradictions suivantes l'effet est moins visible car la dispersion plus élevée et il faut se référer aux figures 7.16, 7.17 d'une part et 7.25 à 7.33 pour mieux comprendre leur effet. Néanmoins, on voit que les points sont soit situés sur la ligne verte marquant la référence d'angle donnée par le VDU (0° de contradiction), soit dans le voisinage des lignes oranges marquant respectivement sur chaque figure les contradictions à $\pm 75^\circ$, $\pm 60^\circ$ et $\pm 45^\circ$.

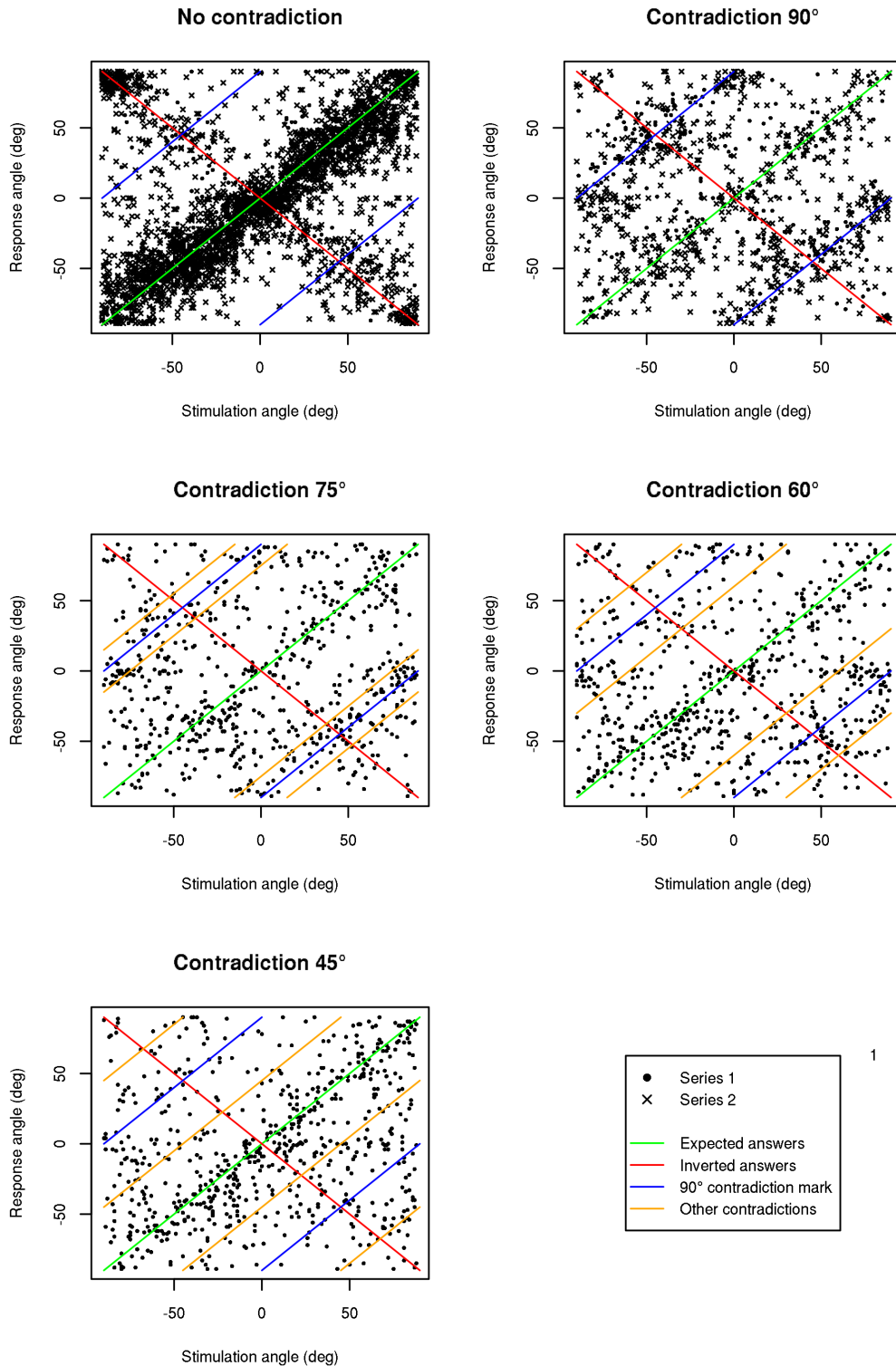


Figure 7.15. Diagramme des angles de réponses en fonction des angles de stimulation égaux (no contradiction) ou en contradiction de $\pm 90^\circ$, $\pm 75^\circ$, $\pm 60^\circ$ ou $\pm 45^\circ$ entre le VDU et le TDU. Les lignes vertes indiquent les réponses attendues non inversées correspondant à l'angle affiché sur le VDU, tandis que les lignes rouges correspondent aux réponses attendues en cas d'inversion. De plus, on indique avec une ligne bleue la position de l'angle de stimulation avec contradiction de 90° ainsi qu'en orange sur chaque figure, la position de l'angle de

stimulation avec contradiction de respectivement $\pm 75^\circ$, $\pm 60^\circ$, $\pm 45^\circ$. Les deux séries de mesures sont marquées par des points différents mais il n'y a pas de différence significative de comportement entre ces séries.

Les figures 7.16 et 7.17 apportent plus de détails sur l'effet des contradictions d'angle. Ces deux figures correspondant respectivement à un boxplot et un violin plot (indiquant en plus la forme de la densité estimée de points) indiquent la position médiane et la dispersion des données autour de l'erreur d'angle de référence de 0° correspondant à une réponse proche de l'angle de stimulation du VDU.

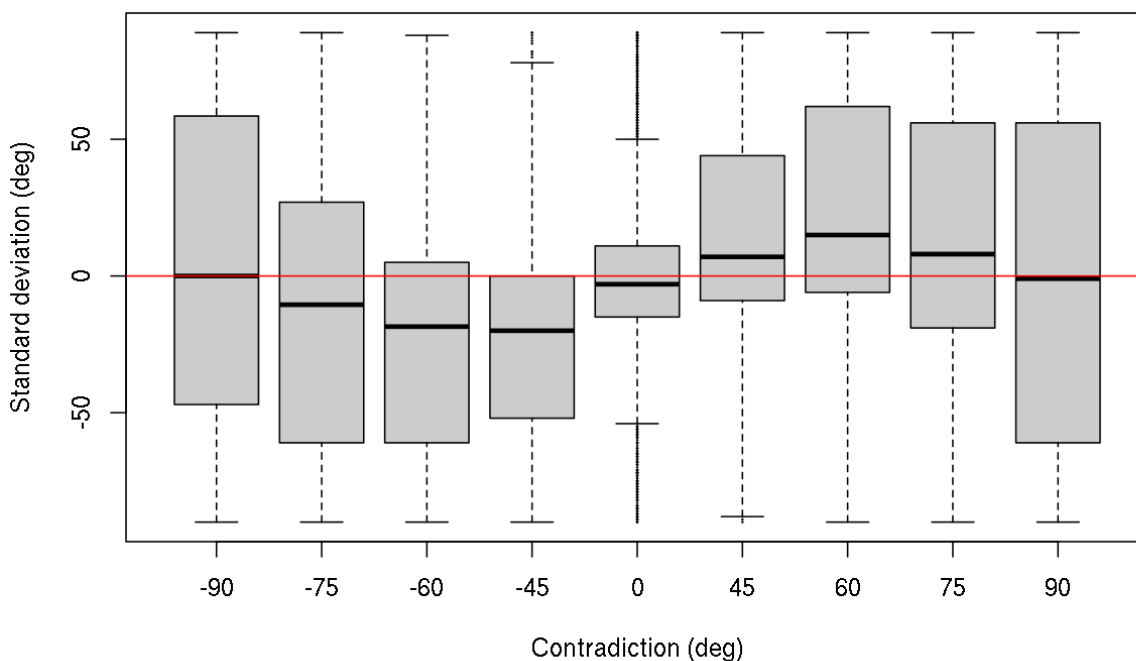


Figure 7.16. Boxplot montrant pour chaque valeur de contradiction la position médiane, les quantiles Q25 et Q75 permettant d'apprécier la dispersion des valeurs autour de la médiane.

On voit que les données sans contradiction ont peu de dispersion et que leur médiane est centrée sur 0° comme attendu. On voit bien aussi que, pour des contradictions de 45° et 60° , et de façon moins marquée pour 75° , les distributions sont excentrées et asymétriques. De plus, elles sont décalées en adéquation avec le signe de la contradiction (+ ou - l'angle ajouté pour la contradiction). A 90° de contradiction, en revanche, on voit que la médiane est à nouveau centrée sur 0 et que la dispersion est très élevée. Les réponses étant centrées soit

sur une erreur de 0° , soit sur une erreur de 90° (et $\pm 90^\circ$ étant des contradictions identiques), les réponses sont donc symétriques par rapport à une erreur de 0° , expliquant ainsi la symétrie et l'étalement important de ces distributions. Le même effet est présent dans une moindre mesure pour les contradictions à $\pm 75^\circ$ (les grandes erreurs autour de $+75^\circ$ pouvant valoir plus de 90° , donc moins de -90° , et réciproque).

Il est plus facile encore de comprendre ces résultats en regardant la figure 7.17, un violin plot qui permet de voir dans quelles régions de ces distributions se concentrent les informations (mesures).

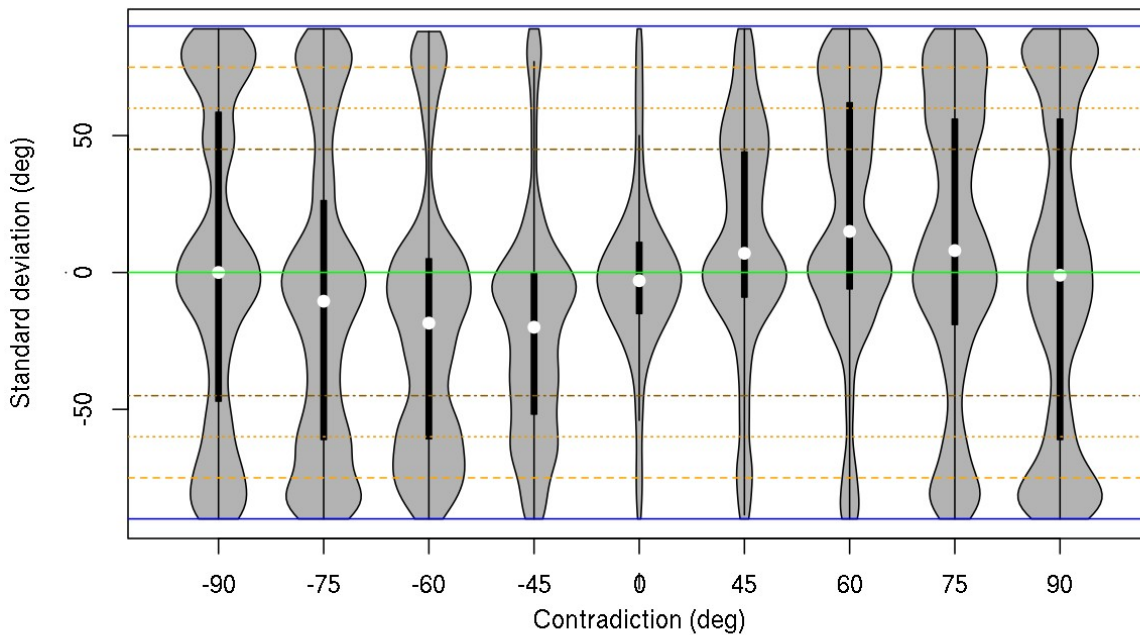


Figure 7.17. Violin plot montrant pour chaque valeur de contradiction la position médiane (point blanc), les quantiles Q25 et Q75 (barre noire) permettant d'apprécier la dispersion des valeurs autour de la médiane. Le violin plot montre en plus une estimation de densité par noyau (largeur de bande de 10°). Les lignes oranges indiquent les contradictions à $\pm 45^\circ$, $\pm 60^\circ$ et $\pm 75^\circ$ et les lignes bleues marquent les contradictions à $\pm 90^\circ$.

Cette fois, on voit clairement que les modes de ces distributions sont positionnés proches des valeurs des contradictions (et de la référence à 0°). On voit aussi que la proportion des choix de réponses est à peu près égale entre la modalité de perception de référence VDU (0° d'erreur) et la modalité TDU (en contradiction) : il y a autant de réponses

fondées sur la perception par l'intermédiaire du VDU que sur celle par l'intermédiaire du TDU. À noter que l'estimateur de densité se fonde sur un noyau gaussien et qu'en conséquence l'estimation aux limites ($+90^\circ$ et -90°) est mauvaise car non contiguë entre $+90$ et -90° . Nous n'avons pas pris le temps de recoder un meilleur estimateur de densité tenant compte de la topologie de l'espace des angles (compact périodique).

S'il apparaît clair que les contradictions ont un effet important dans la perception, il n'est pas facile d'y voir des effets de fusion (réponses intermédiaires). Dans la sous-section suivante (section 7.4.3.2), nous proposons de rechercher des modèles pour expliquer les réponses des sujets en fonction des valeurs de contradiction et tenter de mettre en évidence des fusions d'informations entre la référence et l'information contradictoire.

Temps de réponse.

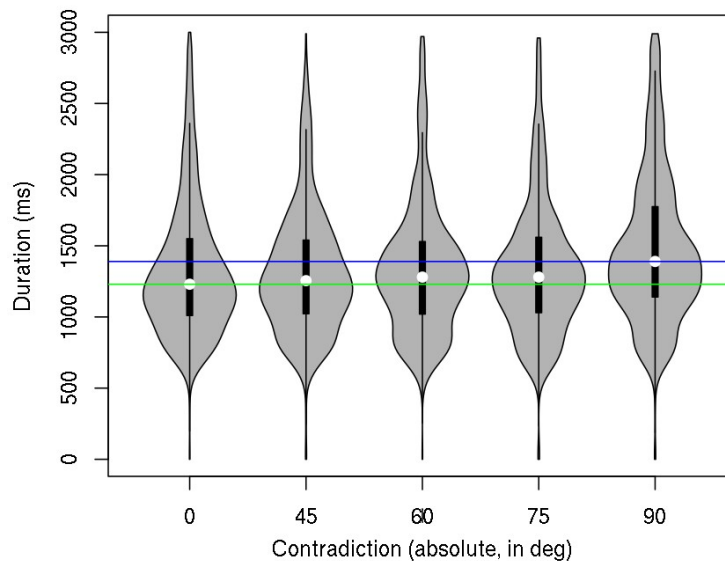


Figure 7.18. Temps de réponse en fonction de la valeur de contradiction (en valeur absolue). La ligne verte matérialise la médiane des temps de réponse sans contradiction ; la ligne bleue celle des temps de réponse avec contradiction à $\pm 90^\circ$.

Nous nous attarderons peu sur l'analyse des temps de réponse (Fig. 7.18), mais il est intéressant de constater que la présence de contradiction augmente significativement ceux-ci, surtout lorsque la contradiction est forte, donc parfaitement identifiable comme une

contradiction (un problème à intégrer et à résoudre) par les sujets. Les contradictions faibles (45°) ont des temps de réponse similaire à l'absence de contradiction, soit parce que les angles sur le TDU ne diffèrent pas suffisamment de ceux du VDU et ne sont pas forcément interprétés comme un écart significatif et perceptible, soit parce qu'il est plus facile de fournir une réponse consensus (acceptable pour représenter à la fois la référence et l'information contradictoire) quand l'écart d'angle est faible.

7.4.3.2. Analyse des erreurs d'angles - recherche de modèles

Comme décrit dans la section 5.3, nous utilisons la méthode EM adaptée à une recherche de mélange de distributions gaussiennes sur un compact périodique pour tenter de comprendre la façon dont les sujets ont répondu en fonction des valeurs de contradiction. En pratique, nous calculons l'AIC pour des modèles estimés sur les données d'origine et allant de 2 à 6 clusters gaussiens. Nous faisons de même pour des jeux de données obtenus par bootstrap (tirage avec remise des valeurs mesurées). Les calculs étant longs, nous avons fixé à 300 le nombre d'échantillons bootstrapés sur lesquels nous estimons à chaque fois un modèle à N clusters et grâce auxquels nous calculons l'AIC moyen par modèle et la variance de l'AIC par modèle. Rappelons que l'AIC, pour Akaike Information Criterion, est un score évaluant la capacité d'un modèle à représenter des données en minimisant la distance du modèle aux données et pour un nombre de paramètres libres (3 paramètres par cluster : moyenne, variance et proportion) minimal. Normalement, nous devons retenir le modèle ayant l'AIC le plus faible. Cependant, nous verrons que d'une part l'AIC minimal n'indique pas forcément le même modèle dans le cas de la recherche de modèle sur les données d'origine et par bootstrap. Le bootstrap donnant une indication sur la robustesse de l'estimation des paramètres, il est important de tenir compte des résultats ainsi obtenus. D'autre part, la variance des scores d'AIC obtenus est importante autorisant ainsi une certaine latitude dans le choix des modèles.

Toujours grâce au bootstrap, nous calculons les estimés moyens des paramètres moyenne, écart type et proportion et affichons leurs distributions. La variance de ces estimés nous renseigne aussi sur la qualité du modèle obtenu et aide ainsi à choisir un "bon modèle" : si les paramètres estimés sont très robustes, cela indique vraisemblablement un bon modèle, même si celui-ci n'a pas forcément le meilleur score d'AIC.

Dans l'ensemble de nos mesures, nous devons nous attendre à un minimum de 2 clusters gaussiens, dans le cas de l'expérience sans contradiction : un pour les réponses majoritaires, l'autre pour les cas d'inversion observables dans ce cas sans contradiction. Dans les cas de contradiction, on s'attendra au moins à 2 clusters gaussiens importants : un pour les réponses relatives aux angles du VDU, la référence, donc centré sur 0° d'erreur, et un pour les réponses relatives au TDU, donc centrées sur les valeurs de contradiction. De plus, on peut s'attendre aussi à des clusters additionnels en cas de réponses consensus, fusions des modalités de perception VDU et TDU. C'est ce que nous recherchons dans cette étude.

Pas de contradiction.

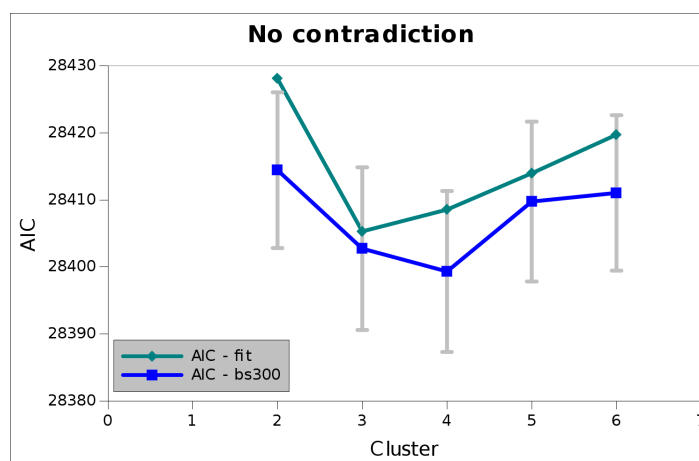


Figure 7.19. Scores d'AIC pour les mesures sans contradiction. Valeurs de l'AIC en fonction du nombre de clusters gaussiens utilisés. La courbe verte représente la recherche de modèle sur les valeurs initiales ; la courbe bleue celle obtenue par bootstrap (300 sets bootstrapés). Les barres d'erreur représentent l'intervalle de confiance à 95% sur les valeurs d'AIC obtenues par bootstrap pour chaque modèle (pour chaque nombre de clusters gaussiens).

Comme nous venons de l'écrire, dans le cas sans contradiction, on peut s'attendre à trouver 2 gaussiennes correspondant à 2 comportements moyens et leurs variances respectives, à savoir le comportement majoritaire des réponses centrées sur 0° (perception VDU et TDU), et le comportement d'inversion (qui, rappelons le, n'est probablement pas gaussien, cette approximation étant une facilité que nous avons choisie). Les meilleurs scores d'AIC (Fig. 7.19) sont obtenus pour des modèles à 3 ou 4 gaussiennes. Cependant, il faut remarquer que l'intervalle de confiance de l'AIC bootstrapé est très large, rendant difficile le

choix d'un modèle parmi les 5 déterminés. Si l'on regarde les résultats à 2 gaussiennes (ce qu'on attend *a priori*) et 3 gaussiennes (Figs. 7.20 et 7.21) on voit qu'effectivement, dans tous les cas, le cluster principal correspond aux données centrées sur 0° . À 2 clusters, on obtient une gaussienne centrée sur 90° qui correspond vraisemblablement aux inversions. Au-delà de 2 clusters (3 ou 4) on obtient de nombreuses gaussiennes de faible ampleur qui résolvent finement les valeurs éloignées de 0° . Ces gaussiennes ont peu de sens (il est difficile d'y trouver une explication) et on ne doit pas oublier que, d'après la figure 7.19, l'estimation de l'AIC est peu fiable et qu'il est tout aussi justifié de retenir le modèle à 2 clusters gaussiens. L'observation d'une (ou plusieurs) gaussienne(s) mettant à jour le phénomène d'inversion n'est pas anodine ici. En effet, il ne faut pas oublier que les sujets perçoivent exactement la même information d'angle à la fois sur le TDU et le VDU. On pourrait dès lors s'attendre à ce qu'ils fixent leur choix sur le VDU, plus intuitif pour eux et d'une lecture plus aisée. Comme dans l'expérience de bimodalité (section 7.3) on s'aperçoit cependant que l'information du TDU est prise en compte et qu'elle induit parfois un comportement d'inversion.

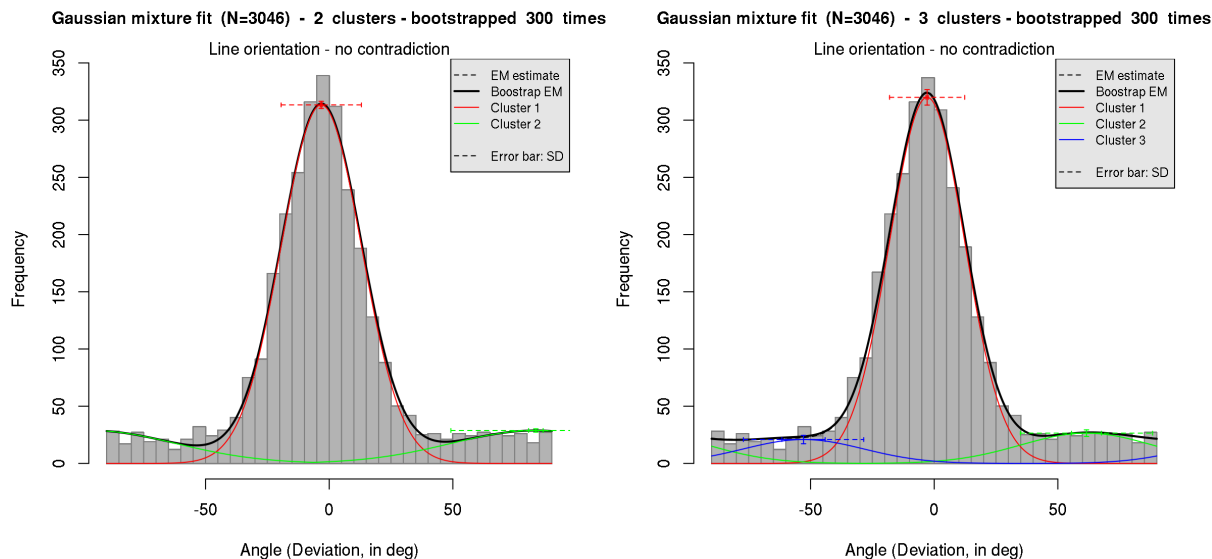
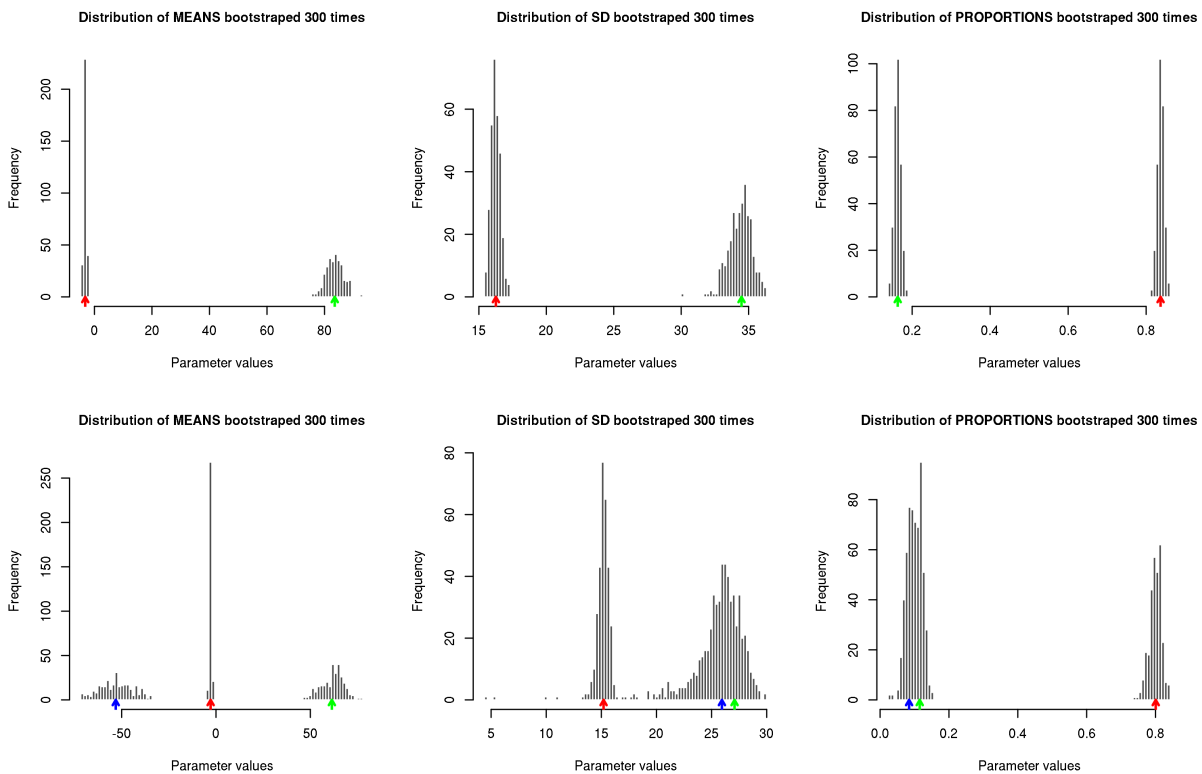


Figure 7.20. Histogrammes des angles de réponse mesurés et leur approximation par des sommes de gaussiennes. Pour chaque approximation, la courbe noire en pointillés, quasiment confondue avec la courbe noire en trait plein, représente le modèle obtenu par mélange des gaussiennes (courbe de couleur). La courbe noire en trait plein est celle qui a été obtenue par bootstrap (courbe robuste). Les barres d'erreur horizontales en pointillés représentent l'écart type des distributions gaussiennes. Les barres d'erreur horizontales en trait plein représentent l'écart type des moyennes obtenues par bootstrap (donc l'incertitude sur leurs positions respectives). Les barres d'erreur verticales en trait plein représentent l'écart type des proportions obtenues par bootstrap (donc l'incertitude sur leur importance). (A gauche) approximation par un mélange de 2 gaussiennes. (A droite) approximation par un mélange de 3 gaussiennes.



Nb Gaussiennes	Moyennes	Écart Types	Proportions
2 clusters	-3.2 ± 0.4	16.2 ± 0.3	0.84 ± 0.008
	83.6 ± 2.9	34.4 ± 0.9	0.16 ± 0.008
3 clusters	-52.8 ± 0.5	15.2 ± 0.4	0.8 ± 0.02
	-2.9 ± 6.1	26.7 ± 1.4	0.12 ± 0.01
	61.6 ± 8.7	24.3 ± 3.2	0.08 ± 0.01

Figure 7.21. Distributions des valeurs de moyenne (gauche), écart type (milieu) et proportion (droite) obtenues par bootstrap sur les données d'origine et recherche de mélange gaussien. Les flèches indiquent la moyenne de ces valeurs (qui sont reportées dans le tableau : moyennes et écart types des moyennes, écart types et proportions obtenus par bootstrap). (Graphique du haut) modèle à 2 gaussiennes. (Graphique de bas) modèle à 3 gaussiennes. Ce sont ces valeurs qui sont utilisées pour tracer les graphiques de la figure 7.20.

Contradiction de 90°.

Les résultats obtenus avec la contradiction à 90° sont les plus simples à analyser, contrairement à ceux des contradictions à 75, 60 et 45°. Cela est dû au fait que les réponses attendues pour le VDU et pour le TDU sont orthogonales et ainsi les plus éloignées possibles.

Le meilleur modèle expliquant les angles de réponses est un modèle à 4 gaussiennes (Figs. 7.21 à 7.24). La première constatation est que les sujets ont choisi comme référence

d'angle pour leur réponse le VDU comme le TDU à parts égales (mêmes proportions de près de 40% des réponses). On observe en effet 2 gaussiennes centrées respectivement sur des valeurs très proches de 0° et 90° . Il n'était pas évident d'obtenir ce résultat très net. On pouvait en effet s'attendre à ce que dans le cas de contradiction, la référence donnée par le VDU soit prioritaire dans le choix des sujets. Ce n'est pas le cas et ceci signifie que le niveau de perception « visuelle » par l'intermédiaire de la langue peut être considérée comme égal à celui de la vision, confirmant ainsi l'affirmation de P. Bach-y-Rita concernant l'indépendance de la fonction de perception vis à vis de l'organe.

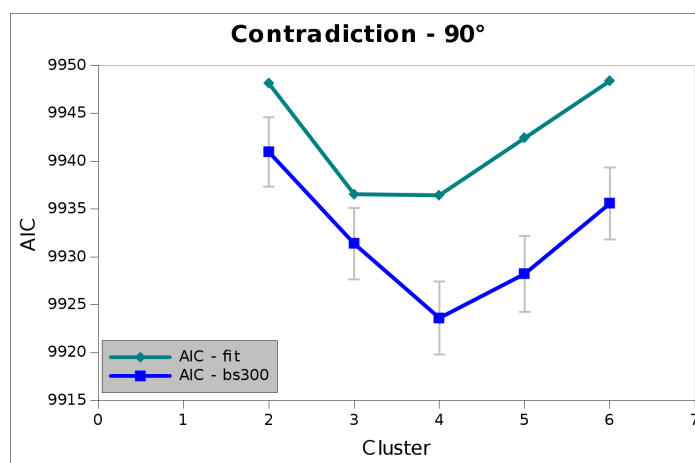


Figure 7.22. Scores d'AIC pour les mesures avec 90° de contradiction. Valeurs de l'AIC en fonction du nombre de clusters gaussiens utilisés. (voir Fig. 7.19)

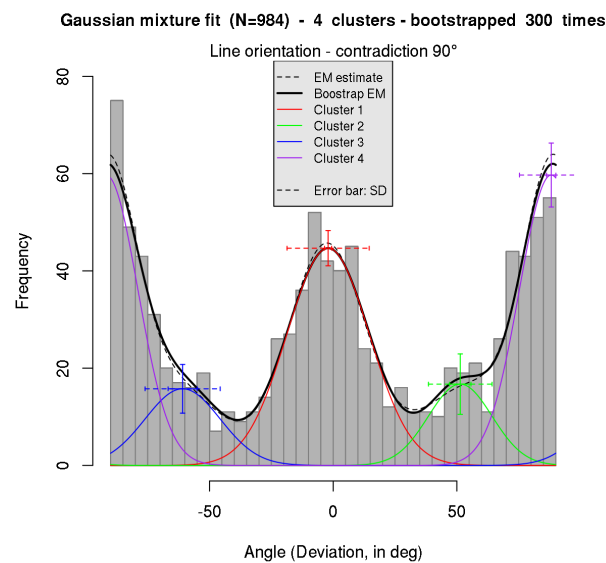
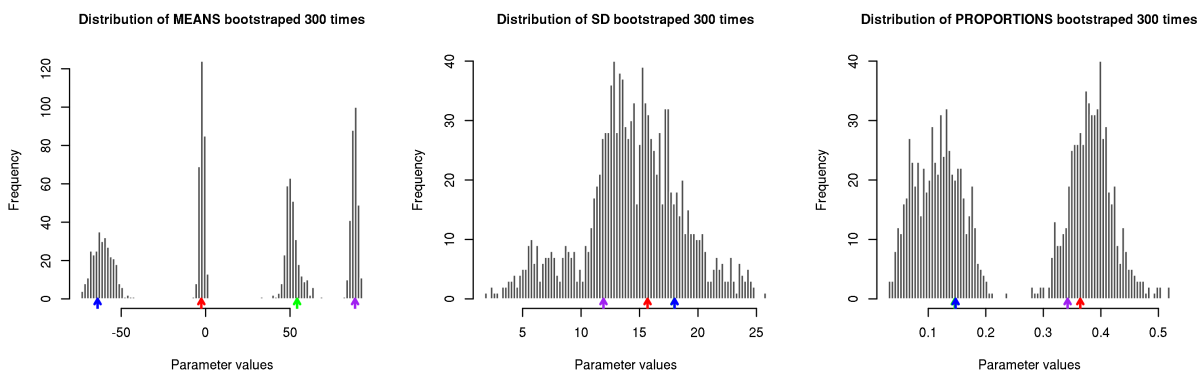


Figure 7.23. Histogrammes des angles de réponse mesurés et leur approximation par des sommes de gaussiennes (voir Fig. 7.20). Il s'agit du meilleur modèle pour cette distribution, un modèle à 4 gaussiennes.

On observe en outre 2 gaussiennes supplémentaires et intermédiaires entre ces 2 gaussiennes principales, légèrement plus proches de la gaussienne des réponses « TDU » que de celle des réponses « VDU ». Ces gaussiennes sont très robustes : d'une part l'AIC est très bas pour ce modèle par rapport aux autres ; d'autre part les écarts types sur leurs paramètres (moyenne, écart type et proportion) sont peu élevés indiquant une bonne robustesse de leur estimation. Ces gaussiennes ne sont pas dues au hasard.

Nous pensons que ces réponses intermédiaires sont le résultat de fusions d'informations d'angles provenant du VDU et du TDU. Ce phénomène serait comparable à celui observé par Mc Gurck et Mc Donald dans le cas de fusions d'informations auditives et visuelles contradictoires.



Nb Gaussiennes	Moyennes	Écarts Types	Proportions
4 clusters	-2.1 ± 1.4	16.6 ± 1.8	0.38 ± 0.03
	51.4 ± 4.3	12.9 ± 5.7	0.11 ± 0.04
	-60.9 ± 5.9	15.2 ± 4.9	0.12 ± 0.04
	88.2 ± 1.9	12.8 ± 1.2	0.39 ± 0.04

Figure 7.24. Distributions des valeurs de moyenne (gauche), écart type (milieu) et proportion (droite) obtenues par bootstrap sur les données d'origine et recherche de mélange gaussien. Les flèches indiquent la moyenne de ces valeurs (qui sont reportées dans le tableau : moyennes et écart types des moyennes, écart types et proportions obtenus par bootstrap). Le graphique correspond au modèle à 4 gaussiennes. Ce sont ces valeurs qui sont utilisées pour tracer le graphique de la figure 7.23.

Notons qu'en toute rigueur nous devrions observer 4 gaussiennes correspondant à des fusions d'information, en plus des 2 gaussiennes des réponses principales (et sans compter les réponses inversées dont la figure 7.20 montre qu'elles sont minoritaires). En effet, on peut s'attendre à ce que les personnes qui choisissent plutôt la référence TDU mais qui sont

influencées par le VDU répondent légèrement à côté de la valeur de référence du TDU (par exemple 20 à 30° de moins ou de plus que 90°), tandis que les personnes qui choisissent davantage la référence visuelle (VDU) mais qui sont influencées par le TDU répondent légèrement à côté de la valeur de référence du VDU (par exemple 20 à 30° de plus ou de moins que 0°). Il est aisé de comprendre l'analogie que nous faisons avec la déviation d'un conducteur par rapport à une trajectoire routière au niveau d'une bifurcation, lorsque la personne qui choisi est indécise entre aller à droite et à gauche. Si elle a plutôt choisi d'aller à gauche mais qu'elle a un doute (que la direction de droite la tente), elle déviara légèrement de sa trajectoire (souhaitée) de gauche vers la trajectoire de droite (et rétablira en toute logique sa trajectoire pour rester sur la route) ; si au contraire elle a plutôt choisi d'aller à droite mais qu'elle est influencée par la trajectoire de gauche, elle déviara légèrement depuis la droite vers la gauche. À ces réponses de fusion, il faut ajouter celles des personnes qui ne sont pas influencées : les trajectoires de gauche et de droite séparées. Si l'on observe bien la gaussienne du VDU (centrée sur -2.1°), on s'aperçoit que, contre toute attente, elle est plus large (écart type de 16.6° pour le VDU contre 12.8° pour le TDU) que celle du TDU. Il est possible que cette gaussienne du VDU soit en fait une combinaison de réponses perturbées par la perception des angles du TDU. La quantité de données dont nous disposons ici ne permet pas de mettre ceci en évidence.

Bien entendu, dans cet exemple du conducteur comme dans notre expérience, tout n'est pas aussi tranché : certains sujets seront plus ou moins influencés par l'autre mode de perception et dévieront en conséquence par rapport à ce qui est attendu, produisant non pas un nombre fini de réponses fusionnées, mais autant qu'il y a de sujets et de points d'expérience.

Nous allons voir que, bien que ce ne soit pas aussi net, les expériences suivantes tendent à montrer l'existence de comportements de fusion des différentes perceptions.

Contradiction de 75°.

Les figures 7.25 à 7.27 montrent les résultats pour les contradictions à 75°. Cette expérience de contradiction est une des plus difficiles à interpréter. Contrairement à toutes les autres expériences décrites dans ce chapitre, la courbe du score d'AIC en fonction du nombre de gaussiennes utilisées pour approximer les données n'est d'une part pas stable (en ce qui

concerne la courbe moyenne obtenue par bootstrap ; figures 7.25 courbe bleue avec barres d'erreur) et d'autre part, la courbe d'AIC calculée sur les données originales (non bootstrappées ; figures 7.25 courbe verte) ne suit pas la courbe moyenne d'AIC obtenue par bootstrap et ne fait qu'augmenter. Cela signifie d'une part que l'algorithme de recherche de mélange gaussien trouve plusieurs solutions possibles (2, 4 et même éventuellement 7 clusters) ; comme il n'y a clairement pas unicité de la solution il est difficile de se fier à un modèle ou à un autre. Le bootstrap se comporte comme une sorte de recuit simulé dans lequel en changeant un peu les données, on accède à d'autres solutions. Le fait que les 2 courbes d'AIC ne se suivent pas indique un manque de robustesse important sur toute analyse faite sur les données d'origine.

Si l'on se fie à la courbe d'AIC obtenue sur les données bootstrappées, il ressort tout de même que les meilleurs modèles (score d'AIC) sont les modèles à 2 et 4 clusters gaussiens (le meilleur par bootstrap étant le modèle à 4 gaussiennes). Il semble également que le modèle à 7 clusters puisse être retenu car, bien qu'ayant un score plus élevé, il correspond tout de même à un net minimum de la courbe malgré son nombre élevé de paramètres (21 paramètres libres). N'oublions pas que l'augmentation du nombre de paramètres libres (moyenne, écart type et proportion de chaque gaussienne) coûte très cher au score d'AIC. Ainsi, dans le modèle à 7 gaussiennes, bien qu'on minimise considérablement l'écart des données au modèle avec 7 gaussiennes, l'AIC remonte en flèche.

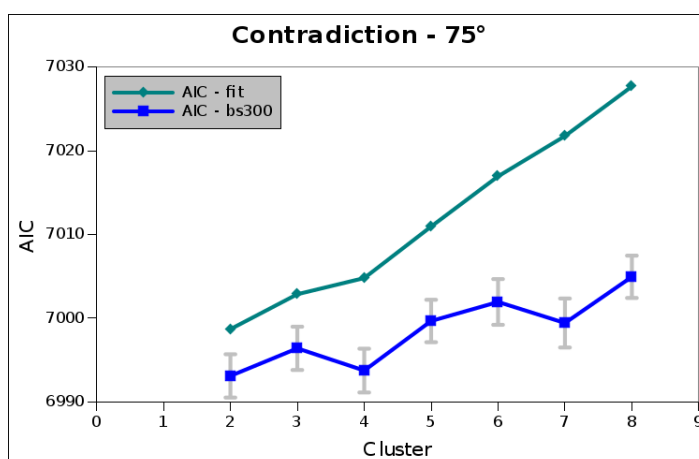


Figure 7.25. Scores d'AIC pour les mesures avec 75° de contradiction. Valeurs de l'AIC en fonction du nombre de clusters gaussiens utilisés. (voir Fig. 7.19)

De façon comparable à l'expérience de contradiction à 90°, nous devons nous attendre à observer au moins 2 réponses principales centrées sur le VDU d'une part (0°) et le TDU

d'autre part. Ce qui est plus compliqué dans le cas des contradictions différentes de 90° , ici de $\pm 75^\circ$ c'est qu'on est censés observer cette réponse TDU centrée d'une part sur 75° et sur -75° d'autre part, c'est-à-dire 2 angles espacés de seulement 30° . Ainsi, il y a toutes les chances pour que ces 2 réponses soient quelque peu confondues aux alentours de 90° d'erreur et qu'on ne puisse facilement les distinguer. Cette remarque sera valable pour les contradictions à 60° et à 45° , bien que ce problème soit plus prononcé pour la contradiction à 75° .

Visuellement, on observe effectivement 2 réponses principales, l'une centrée sur une erreur proche de 0° (VDU), l'autre dans les grandes erreurs, proche de 90° (TDU). On voit cependant que la distribution est étendue dans les grandes erreurs et présente plusieurs modes aux alentours des 75° . L'analyse automatique par la méthode EM de recherche de mélange gaussien fait ressortir ces 3 gaussiennes principales dans une approximation à 3 gaussiennes (certaines incluant donc des réponses de fusion) avec des moyennes respectives centrées sur -3.6° , 64.8° , -78.7° . La recherche de modes de fusion de perception est plus complexe dans ce cas. Cependant, le modèle à 4 clusters, meilleur modèle d'après le score d'AIC, montre bien que des modes gaussiens intermédiaires existent. Ce modèle à 4 clusters ne correspond cependant pas vraiment à ce que l'on attend, à savoir une réponse centrée sur 0° , 2 réponses centrées sur -75° et 75° , et des réponses de fusion entre 0° et 75° (1 ou 2 gaussiennes) d'une part et entre 0° et -75° (1 ou 2 gaussiennes) d'autre part, soit un modèle à 5 gaussiennes ou un modèle à modèle à 7 gaussiennes.

Les figures 7.26 et 7.27 montrent ces modèles à 5 et 7 gaussiennes et donnent leurs caractéristiques. Ces deux modèles sont cohérents avec ce que l'on attend :

Le modèle à 5 clusters montre en effet les gaussiennes principales centrées sur -1.8° (VDU) d'une part et -81.7° et 77.4° (TDU) d'autre part (ce qui est assez proche des 0° et $\pm 75^\circ$ attendus, surtout si l'on considère l'écart-type des moyennes de l'ordre de 5°) et couvrant respectivement 28% des effectifs pour la réponse VDU et 46% pour le TDU. Les réponses intermédiaires correspondent à 2 gaussiennes -31.3 et 37.3° et couvrant 25% de l'effectif.

Plus le nombre de gaussiennes augmente, plus il devient difficile de les interpréter, néanmoins, il semble que ce que trouve le modèle à 7 gaussiennes reste assez cohérent avec 3 gaussiennes principales centrées sur -1.74° (comme dans le modèle à 5) et couvrant près de 20% des effectifs pour la réponse VDU, et pour le TDU -85.8° et 80.5° avec $\pm 5^\circ$ d'écart type sur ces moyennes, donc assez proches encore de $\pm 75^\circ$, et couvrant près de 30% des

effectifs. Les réponses intermédiaires se trouvent d'une part à proximité des réponses TDU, à -63° et 59.7° (donc proches de $\pm 60^\circ$ avec une distance de 15° par rapport aux réponses TDU) et d'autre part de part et d'autre et à proximité de la réponse VDU, soit à -22.2° et 14.5° (avec plus de 5° d'erreur), donc proches de $\pm 15^\circ$, distance à la réponse VDU. Notons aussi la faible variabilité des estimés de ces moyennes par bootstrap dans la figure 7.27 (bas).

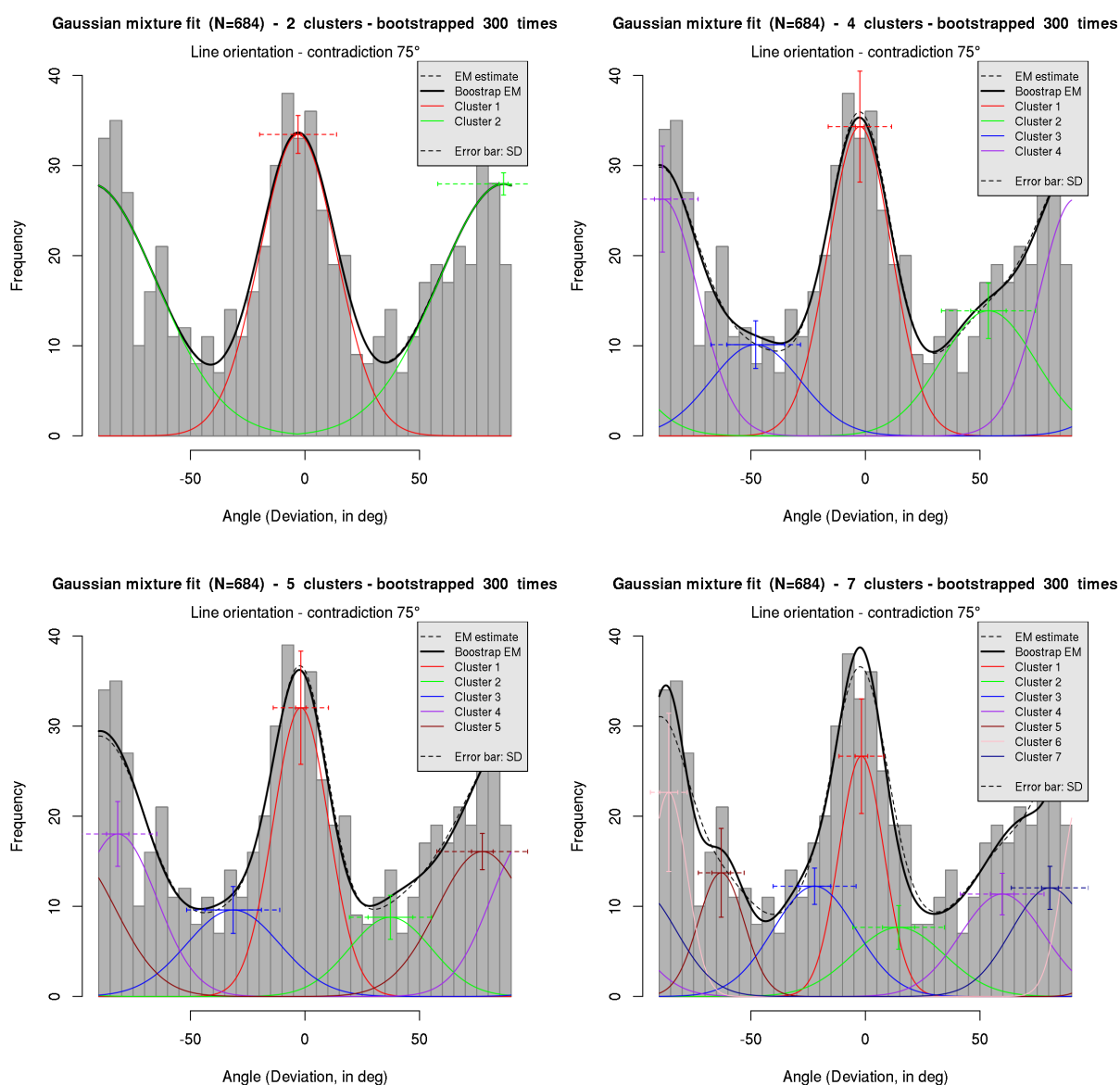
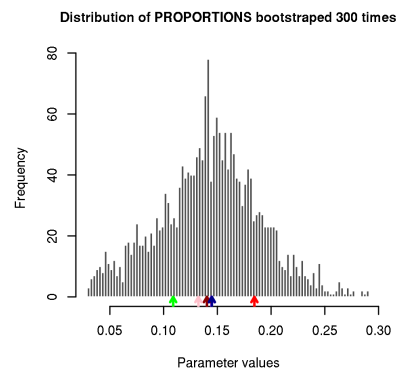
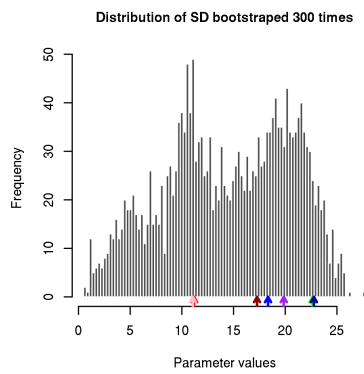
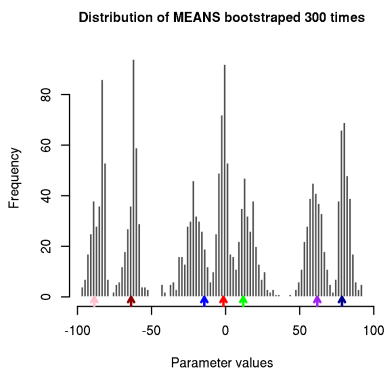
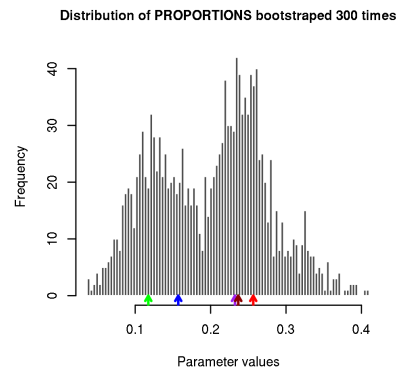
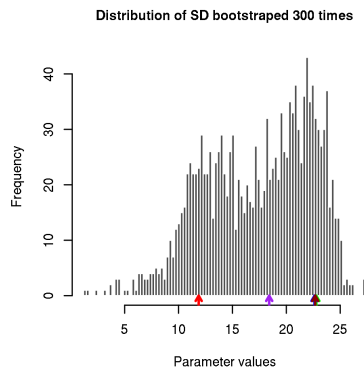
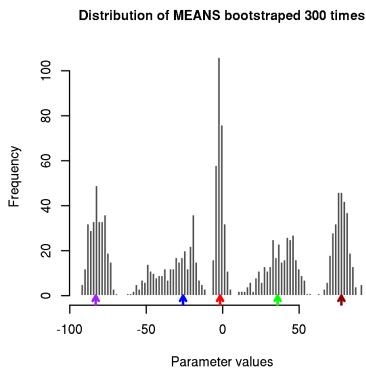
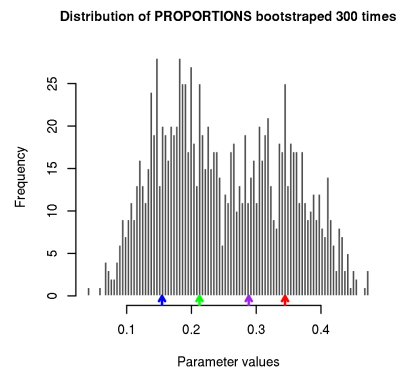
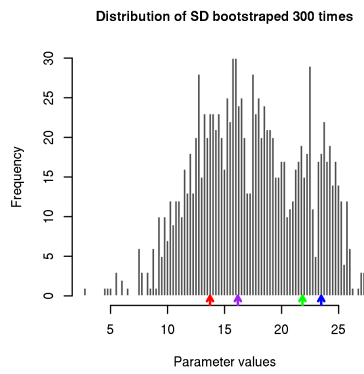
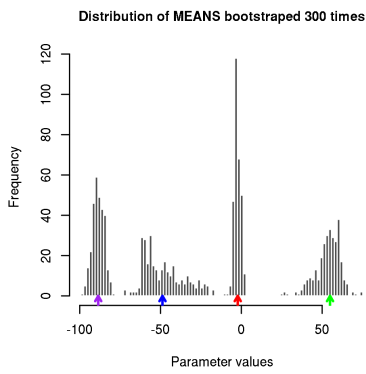
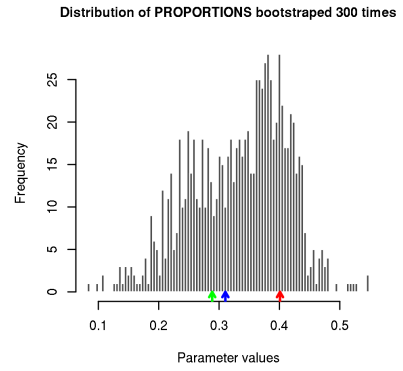
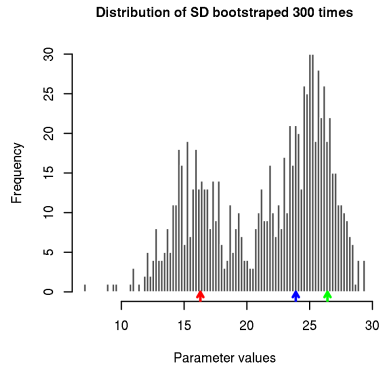
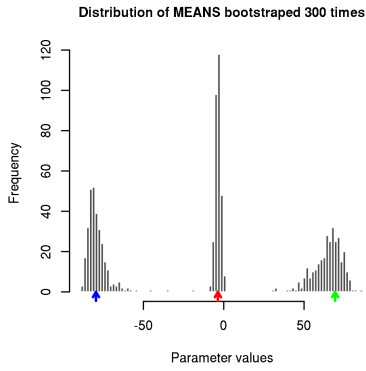


Figure 7.26. Histogrammes des angles de réponse mesurés et leur approximation par des sommes de gaussiennes (voir Fig. 7.20). Il s'agit des 4 meilleurs modèles pour cette distribution, d'après les scores d'AIC : un modèle à 2 gaussiennes et un modèle à 4 gaussiennes qui n'ont que peu de sens, ainsi qu'un modèle à 5 gaussiennes et un plus raffiné à 7 gaussiennes, cohérents avec ce qui est attendu.



Nb Gaussiennes	Moyennes	Écart Types	Proportions
3 clusters	-3.6 ± 1.6	15.9 ± 2.0	0.39 ± 0.03
	64.8 ± 8.9	24.4 ± 3.0	0.26 ± 0.07
	-78.7 ± 7.2	24.1 ± 2.8	0.34 ± 0.06
4 clusters	-2.4 ± 1.9	13.8 ± 2.5	0.35 ± 0.06
	53.7 ± 7.7	20.5 ± 4.0	0.21 ± 0.05
	-47.9 ± 12.6	19.6 ± 4.6	0.15 ± 0.04
	-88.5 ± 3.6	15.5 ± 3.6	0.30 ± 0.07
5 clusters	-1.8 ± 2.3	12.1 ± 2.0	0.28 ± 0.06
	37.3 ± 9.8	17.9 ± 4.7	0.12 ± 0.03
	-31.3 ± 12.3	20.3 ± 4.4	0.14 ± 0.04
	-81.7 ± 4.9	17.0 ± 4.8	0.23 ± 0.04
	77.4 ± 4.7	19.8 ± 3.1	0.23 ± 0.03
7 clusters	-1.74 ± 2.8	9.8 ± 2.1	0.19 ± 0.05
	14.5 ± 7.0	20.0 ± 3.8	0.11 ± 0.04
	-22.2 ± 7.0	18.1 ± 3.8	0.16 ± 0.03
	59.7 ± 5.2	18.2 ± 3.7	0.15 ± 0.03
	-63.0 ± 4.1	10.0 ± 5.3	0.10 ± 0.04
	-85.8 ± 4.0	8.0 ± 4.0	0.13 ± 0.05
	80.5 ± 3.7	16.9 ± 4.6	0.15 ± 0.03

Figure 7.27. Distributions des valeurs de moyenne (gauche), écart type (milieu) et proportion (droite) obtenues par bootstrap sur les données d'origine et recherche de mélange gaussien. Les flèches indiquent la moyenne de ces valeurs (qui sont reportées dans le tableau : moyennes et écart types des moyennes, écart types et proportions obtenus par bootstrap). Graphique de haut en bas : modèle à 3 gaussiennes, 4 gaussiennes, 5 gaussiennes et 7 gaussiennes. Ce sont ces valeurs qui sont utilisées pour tracer le graphique de la figure 7.26.

Contradiction de 60°.

D'après les scores d'AIC (figure 7.28), c'est clairement le modèle à 5 gaussiennes qui ressort. Comme précédemment on s'attend à 3 gaussiennes principales au moins, les erreurs de réponse selon la référence VDU centrées sur 0° et les erreurs de réponses selon la référence TDU centrées sur plus ou moins 60°. La contradiction diminuant, les réponses correspondant à des fusions de perceptions se rassemblent autour de 60°. Il est donc peu probable de parvenir à distinguer les réponses de type VDU perturbées par le TDU de celles de type TDU perturbées par le VDU. On doit donc s'attendre à 1 gaussienne entre -60° et 0° et une autre entre 0° et 60°.

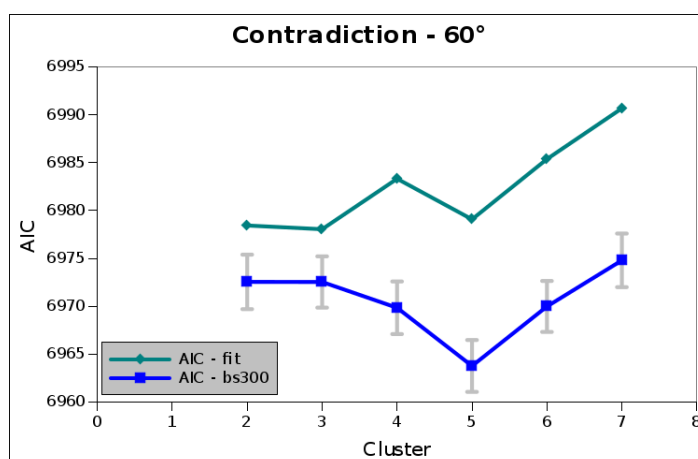


Figure 7.28. Scores d'AIC pour les mesures avec 60° de contradiction. Valeurs de l'AIC en fonction du nombre de clusters gaussiens utilisés. (voir Fig. 7.19)

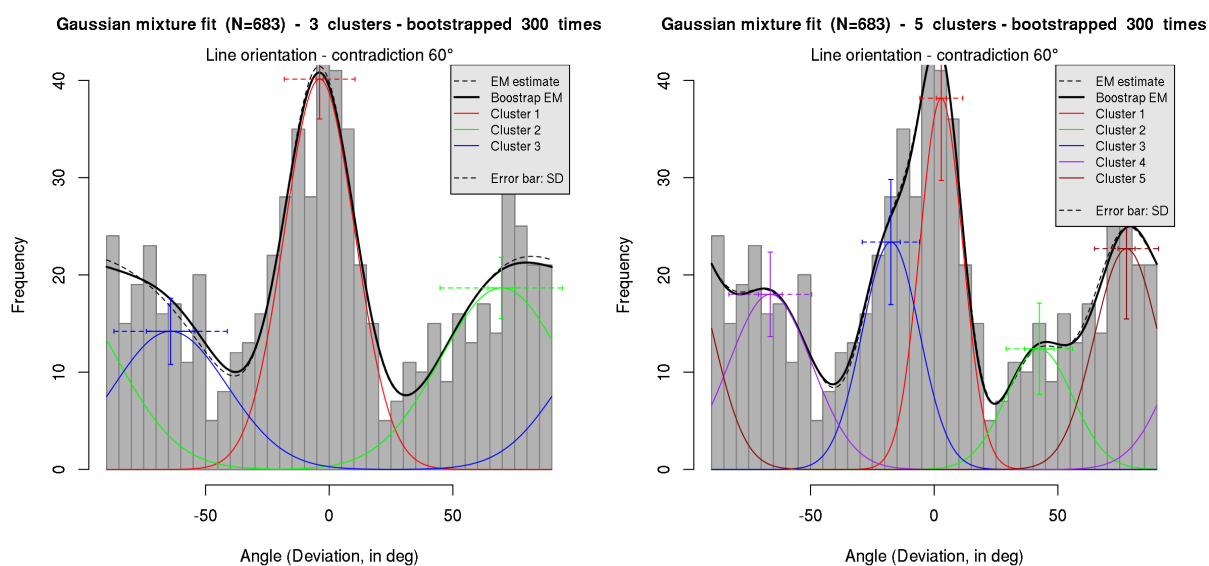
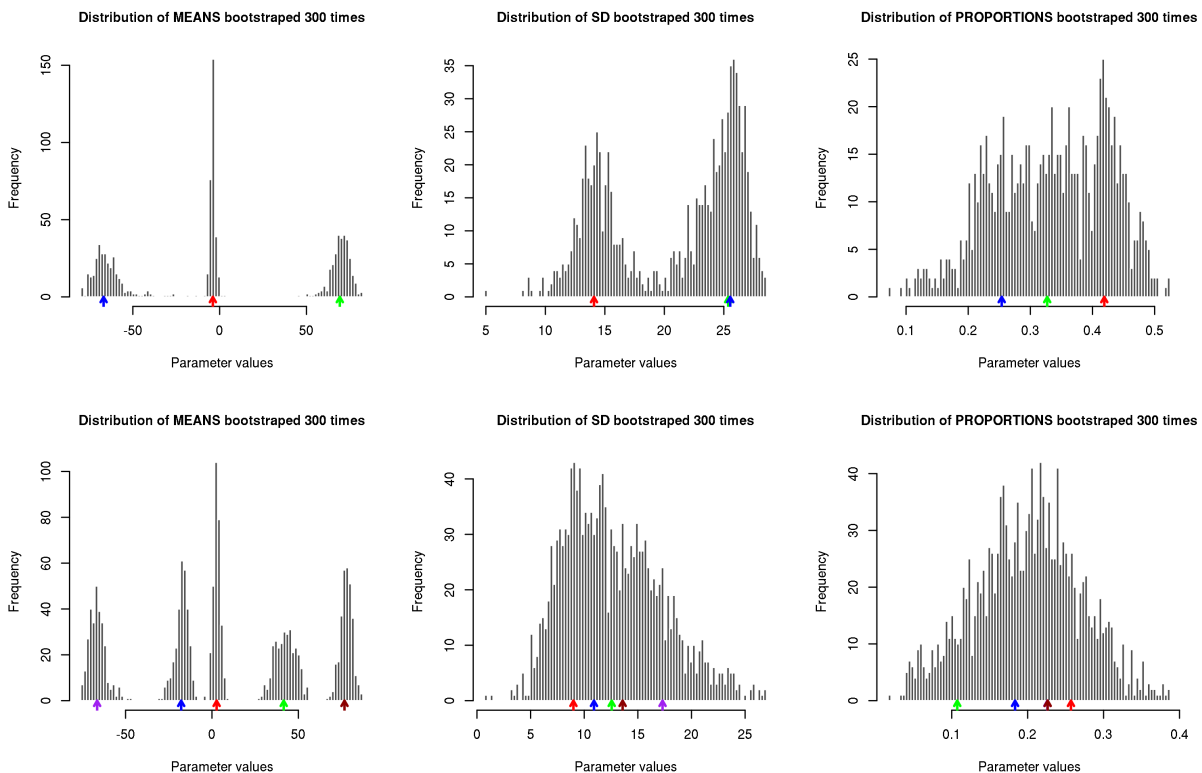


Figure 7.29. Histogrammes des angles de réponse mesurés et leur approximation par des sommes de gaussiennes (voir Fig. 7.20). Il s'agit des 2 meilleurs modèles pour cette distribution, d'après les scores d'AIC : un modèle à 3 gaussiennes et un modèle à 5 gaussiennes.

Dans les figures 7.29 et 7.30 nous ne montrons que les modèles à 3 et 5 clusters gaussiens qui correspondent à des minima dans les courbes d'AIC, le plus robuste étant le modèle à 5 gaussiennes. Comme attendu, le modèle à 3 gaussiennes donne une réponse centrée sur 0° (-3.9°) et deux réponses centrées sur une valeur proche de $\pm 60^\circ$ (-64.1° et 69.6°). Si l'on essaie de faire correspondre le modèle à 5 gaussiennes avec ce que l'on attend, on a comme gaussiennes principales : 2.9° pour le VDU et 66.3° (proche de 60°) et -77.8° (un

peu trop décalée par rapport à -60°) pour le TDU. Les réponses pour le VDU couvrent 24% de l'effectif contre plus de 40% pour le TDU. Les réponses intermédiaires ne sont pas positionnées à des valeurs similaires : -17.5° et 42.5° .

Il est difficile de conclure correctement sur ce modèle à 5 gaussiennes, mais une chose est claire : ce modèle qui est de loin le meilleur d'après le score AIC ne peut être expliqué (approximation gaussienne) correctement que par la présence de réponses intermédiaires.



Nb Gaussiennes	Moyennes	Écart Types	Proportions
3 clusters	-3.9 ± 1.5	14.3 ± 1.7	0.42 ± 0.04
	69.6 ± 5.5	24.8 ± 2.2	0.34 ± 0.06
	-64.1 ± 9.8	22.9 ± 4.7	0.24 ± 0.06
5 clusters	2.9 ± 2.0	8.7 ± 1.5	0.24 ± 0.05
	42.5 ± 6.0	13.4 ± 4.7	0.12 ± 0.05
	-17.5 ± 3.8	11.6 ± 3.0	0.20 ± 0.05
	66.3 ± 4.8	16.7 ± 4.3	0.22 ± 0.05
	-77.8 ± 3.4	12.9 ± 4.1	0.22 ± 0.07

Figure 7.30. Distributions des valeurs de moyenne (gauche), écart type (milieu) et proportion (droite) obtenues par bootstrap sur les données d'origine et recherche de mélange gaussien. Les flèches indiquent la moyenne de ces valeurs (qui sont reportées dans le tableau : moyennes et écart types des moyennes, écart types et proportions obtenus par bootstrap). (Graphique du haut) modèle à 3 gaussiennes. (Graphique du bas) modèle à 5 gaussiennes. Ce sont ces valeurs qui sont utilisées pour tracer le graphique de la figure 7.29.

Contradiction de 45°.

Avec cette contradiction faible, plus que dans les précédentes, il paraît encore plus difficile de mettre en évidence les réponses intermédiaires. On doit s'attendre cette fois à 5 réponses centrées respectivement sur 0°, -45° et +45°, entre 0° et +45° et entre -45° et 0°.

La courbe de scores AIC (Fig. 7.31) présente un minimum pour les modèles à 3, 4 et 5 gaussiennes. Notons que les modèles à 6 ou 7 gaussiennes ne sont pas bien éloignés. La distribution mesurée est en effet très difficile à approximer et l'ajout de gaussiennes, bien que coûteux en paramètres ne fait qu'améliorer cette approximation.

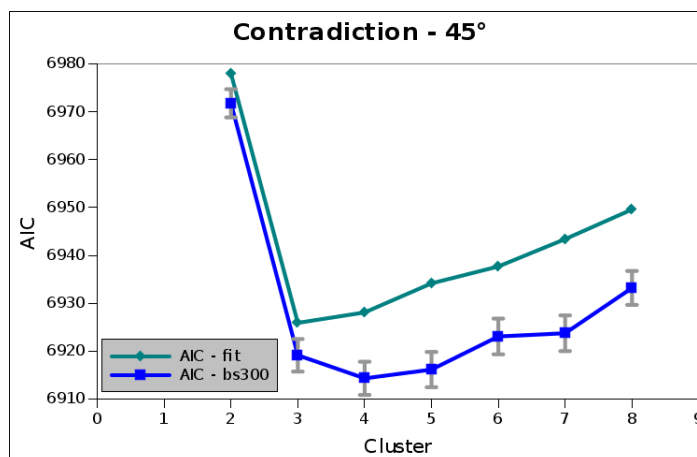


Figure 7.31. Scores d'AIC pour les mesures avec 45° de contradiction. Valeurs de l'AIC en fonction du nombre de clusters gaussiens utilisés. (voir Fig. 7.19)

Il semble que seuls les modèles à 3 et éventuellement 4 gaussiennes soient interprétables (Figs. 7.32 et 7.33). Au-delà, l'ajout de gaussiennes supplémentaires va se faire par répartition uniforme des moyennes de façon à approximer les données qui, elles-mêmes, semblent uniformes à l'exception des réponses centrées sur 0°.

Avec 3 gaussiennes, on obtient des erreurs de réponse centrées respectivement sur -3.4° (VDU), -54.2° et 42.5° (avec 6° d'écart type) pour le TDU, des valeurs comparables aux +/-45° attendus. Avec 4 gaussiennes, on obtient quelque chose d'assez semblable avec des erreurs de réponse centrées sur -2.7° (VDU) d'une part, et -34.4° (avec 10° d'écart type !) et 42.2° (6° d'écart type) pour le TDU, des valeurs encore une fois compatibles avec les +/-45° espérés. De plus, on obtient une gaussienne centrée sur -80° (avec 8° d'écart type), une valeur proche de 90° qui s'explique difficilement.

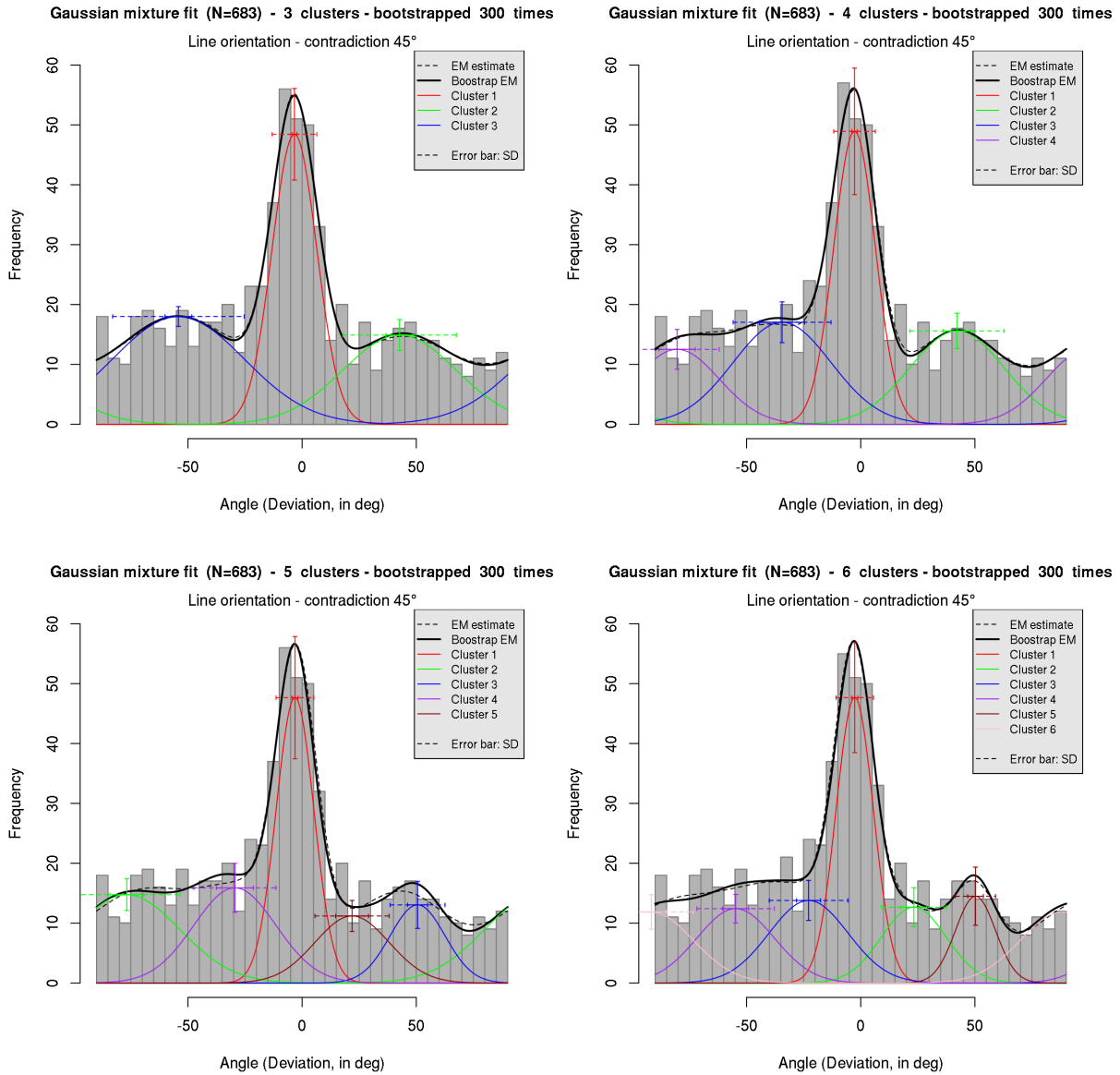
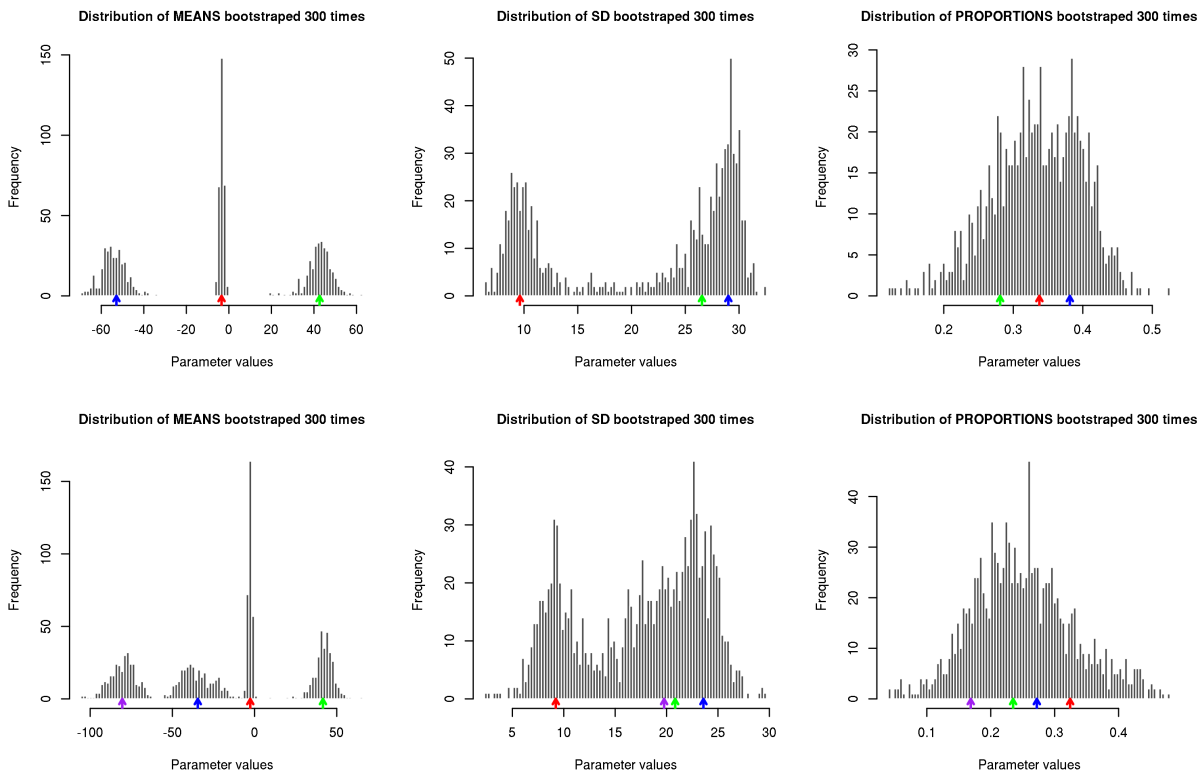


Figure 7.32. Histogrammes des angles de réponse mesurés et leur approximation par des sommes de gaussiennes (voir Fig. 7.20). Il s'agit des 2 meilleurs modèles pour cette distribution, d'après les scores d'AIC : un modèle à 3 gaussiennes et un modèle à 4 gaussiennes. Les modèles à 5 et 6 gaussiennes sont également montrés.



Nb Gaussiennes	Moyennes	Écart Types	Proportions
3 clusters	-3.4 ± 1.0	9.8 ± 1.5	0.35 ± 0.05
	42.5 ± 6.0	24.8 ± 4.7	0.27 ± 0.05
	-54.2 ± 5.9	28.8 ± 1.5	0.38 ± 0.03
4 clusters	-2.7 ± 1.2	9.1 ± 1.6	0.33 ± 0.07
	42.2 ± 6.0	20.7 ± 4.2	0.24 ± 0.04
	-34.4 ± 9.9	21.4 ± 3.5	0.27 ± 0.05
	-80.4 ± 8.0	18.4 ± 4.7	0.17 ± 0.04

Figure 7.33. Distributions des valeurs de moyenne (gauche), écart type (milieu) et proportion (droite) obtenues par bootstrap sur les données d'origine et recherche de mélange gaussien. Les flèches indiquent la moyenne de ces valeurs (qui sont reportées dans le tableau : moyennes et écart types des moyennes, écart types et proportions obtenus par bootstrap). (Graphique du haut) modèle à 3 gaussiennes. (Graphique du bas) modèle à 4 gaussiennes. Ce sont ces valeurs qui sont utilisées pour tracer le graphique de la figure 7.32.

IV - Discussion et Perspectives

Commençons par un rappel des différents points abordés dans ce manuscrit:

Nous avons introduit ce manuscrit avec une problématique de santé publique majeur qui préoccupe le monde entier et dans les conséquences sur la qualité de vie des personnes atteintes est dramatique : la vision et ses maladies. Ces maladies proviennent d'altérations génétiques ou apparaissent à partir d'un certain âge. Dans l'état de l'art, nous avons ainsi décrit l'anatomie de l'œil et les mécanismes physiologiques (micro-saccades oculaires, tremblements, dérive) qui assurent l'exécution du processus de perception visuelle. Certaines maladies, telles que la dégénérescence maculaire liée à l'âge (DMLA) et la rétine pigmentaire (RP), sont décrites avec les solutions médicales et technologiques existantes qui sont proposées pour y remédier. Parmi ces technologies, sont présentées celles reposant sur la vision prothétique ou sur l'exploitation du paradigme de substitution visuelle. Ce sont ces dernières qui nous concernent particulièrement. Aussi, entrant dans le détail, nous avons alors présenté les capacités perceptives tactiles de la peau permettant de traiter des informations spatiales et celles de la langue et le dispositif de substitution associé, le TDU. On conclut ainsi la partie « état de l'art » par une revue des technologies comme la vision prothétique et la substitution visuelle, par électro-stimulation notamment, qui fournissent des solutions palliatives aux personnes aveugles et malvoyantes afin qu'elles puissent percevoir une partie de l'information visuelle perdue. À l'heure actuelle, ces solutions manquent encore de maturité en termes d'efficacité et de qualité d'une part, d'ergonomie et d'acceptabilité d'autre part ; en tout cas elles sont difficilement commercialisables en l'état sous la forme de produits « grand public ».

La perception d'un signal électro-tactile persistant par un organe des sens (langue, œil) fait partie des objectifs à atteindre dans ces dispositifs. Les cellules sensibles de ces organes (neurones de la rétine, corpuscules mécano-récepteurs de la peau ou de la langue) soumis à un signal prolongé subissent rapidement un phénomène de désensibilisation nommé « fading ». Ce phénomène existant dans la vision normale, surtout quand il s'agit de perception d'informations statiques, et résolu par les mécanismes moteurs du système visuel, est à ce jour remarquablement ignoré à la fois dans le fonctionnement des dispositifs d'électro-stimulation (ex: TDU) ainsi que dans celui des implants de vision prothétique.

L'objectif principal de cette thèse était ainsi de déterminer si et dans quelle mesure il est possible d'améliorer la qualité (résolution) ou l'efficacité (perception d'un signal persistant ou d'un signal faible) de perception en vue d'applications de cette technologie d'électro-stimulation linguale, soit dans un cadre médical palliatif, soit dans d'autres applications de réalité augmentée. Nous avons donc cherché à déterminer dans un premier temps quels procédés pouvaient aller dans le sens d'une telle amélioration de la perception, ensuite ce qui pouvait au contraire l'affecter. Ces connaissances pourraient contribuer à améliorer non seulement les dispositifs de substitution visuelle par électro-stimulation, mais également les dispositifs de vision prothétique. Nous avons ainsi d'abord étudié l'effet sur la qualité et l'efficacité de perception par électro-stimulation linguale de procédés appliqués au signal image, procédés inspirés par ceux de la vision physiologique : des saccades et des tremblements du signal-image. Afin de qualifier les effets de ces procédés, nous avons mis en place des expériences de psychomotricité en utilisant une plate-forme expérimentale mise au point spécialement pour cette étude.

Les résultats de ces expériences ont été présentés et analysés. Nous allons maintenant les discuter avec un regard critique.

Dans la première partie de la thèse, nous avons évalué l'effet de procédés de modification du signal image, inspirés par les mécanismes de la vision, sur la perception électro-tactile. Notre hypothèse était que ces procédés, le scintillement ou des saccades de l'image, évitent au sujet de subir le phénomène de perte de sensibilité (fading) vis-à-vis du signal électrique, phénomène que nous pensions pouvoir corriger afin de permettre au sujet de « mieux percevoir » l'image électro-tactile. On peut décliner de plusieurs manières cette « amélioration de la perception ». En effet, l'amélioration peut être comprise comme une augmentation de la résolution effectivement perçue, mais elle peut venir aussi d'une réduction des temps de réponse ou encore du simple fait de continuer à percevoir le signal image, même après une longue stimulation. C'est ce qu'ont permis d'investiguer les expériences décrites au chapitre 6.

La première expérience consistait à tracer une ligne de taxels soit sur le VDU, soit sur le TDU, en présence ou en l'absence de saccades ou de scintillement de l'intensité de cette

ligne. Les sujets devaient donner comme réponse un tracé semblable avec un tracé sur une tablette graphique et ce avec le délai de réponse le plus court possible. Nous avons fait l'hypothèse d'une amélioration de la qualité de perception au sens de la résolution effectivement perçue ; l'expérience n'a cependant pas été concluante sur ce point : elle n'a pas permis de montrer une amélioration de la qualité de perception. Comme nous l'avons dit, la question était mal posée dès le début : ajouter du bruit (scintillement ou saccades) sur un signal de très basse résolution (12x12 taxels) peut difficilement améliorer la qualité de perception des orientations des lignes tracées sur le TDU. Nous avons montré que les saccades oculaires ne faisaient pas mieux !

Nous pouvons tout de même tirer avantage de cette expérience. Nous avons au moins pu vérifier que les sujets, sans apprentissage, se familiarisaient très vite avec le dispositif d'électro-stimulation linguale et parviennent avec une assez grande efficacité à reconnaître l'orientation d'une ligne tracée sur la matrice du TDU. Nous montrons toutefois que les erreurs sont d'autant plus faibles que les lignes tracées sont éloignées des diagonales, les lignes verticales étant les mieux perçues. Ces mesures correspondent à ce qui est ressenti par les sujets : le questionnaire que nous avons fait montre effectivement que les lignes horizontales et verticales sont mieux perçues que les diagonales.

Si le scintillement ou les saccades ne parviennent pas à augmenter la résolution perçue, nous nous attendions cependant à ce qu'elles permettent de ne pas perdre le signal image (phénomène de fading) lorsque celui-ci est maintenu constant. Une autre expérience a donc été faite pour voir si les saccades ou le scintillement d'un signal permettent de le percevoir même après une longue période de stimulation. Cette fois, 2 lignes orthogonales étaient tracées en même temps et la stimulation maintenue pendant un temps variable de 2s (court), 5s (intermédiaire) ou 10s (long). Au bout de ce temps, une des deux lignes était enlevée, cet événement facilement identifiable par les sujets grâce à un signal audible (un « bip ») ; les sujets devaient alors répondre comme dans l'expérience précédente le plus rapidement possible. Dans cette expérience aussi, le signal image était tracé en présence ou en l'absence de scintillement ou de saccades.

Cette fois, l'expérience a répondu de façon positive à notre hypothèse de travail : la présence de saccades ou de scintillement du signal image permet au sujets de continuer à percevoir correctement celui-ci et à répondre correctement. Cet effet était marqué après 10

secondes de stimulation : le signal-image statique n'était plus identifié correctement au contraire du signal-image dynamique (présence de saccades ou de scintillement).

Dans les expériences de guidage d'un geste chirurgical par des signaux affichés sur le TDU [Robineau 2009], les sujets doivent rejoindre une cible (virtuelle) à partir d'un point de départ en dirigeant une aiguille de ponction localisée dans l'espace (position et orientation). La trajectoire est rectiligne et d'une dizaine de centimètres. On espère que le guidage permettra de rester le plus proche possible de cette trajectoire. Lorsque l'instrument s'éloigne de la trajectoire idéale, le sujet reçoit sur le TDU un signal électrique (4 électrodes activées sur un bord de la matrice, en bas, en haut, à droite ou à gauche) d'une intensité proportionnelle à la distance et du côté de l'erreur (Fig. 8.1). Le sujet doit alors minimiser cette erreur, c'est-à-dire minimiser le signal électrique ressenti. Cependant, les sujets qui prennent le temps de faire un geste précis sont vite pénalisés par le fait que des déviations persistant du même côté et initialement faibles sont de moins en moins ressenties et induisent une dérive qui peut s'avérer conséquente. Le sujet n'a pourtant pas conscience de la présence d'un signal électrique signalant l'erreur. La présence de saccades ou de scintillement devrait suffire à corriger ce problème.

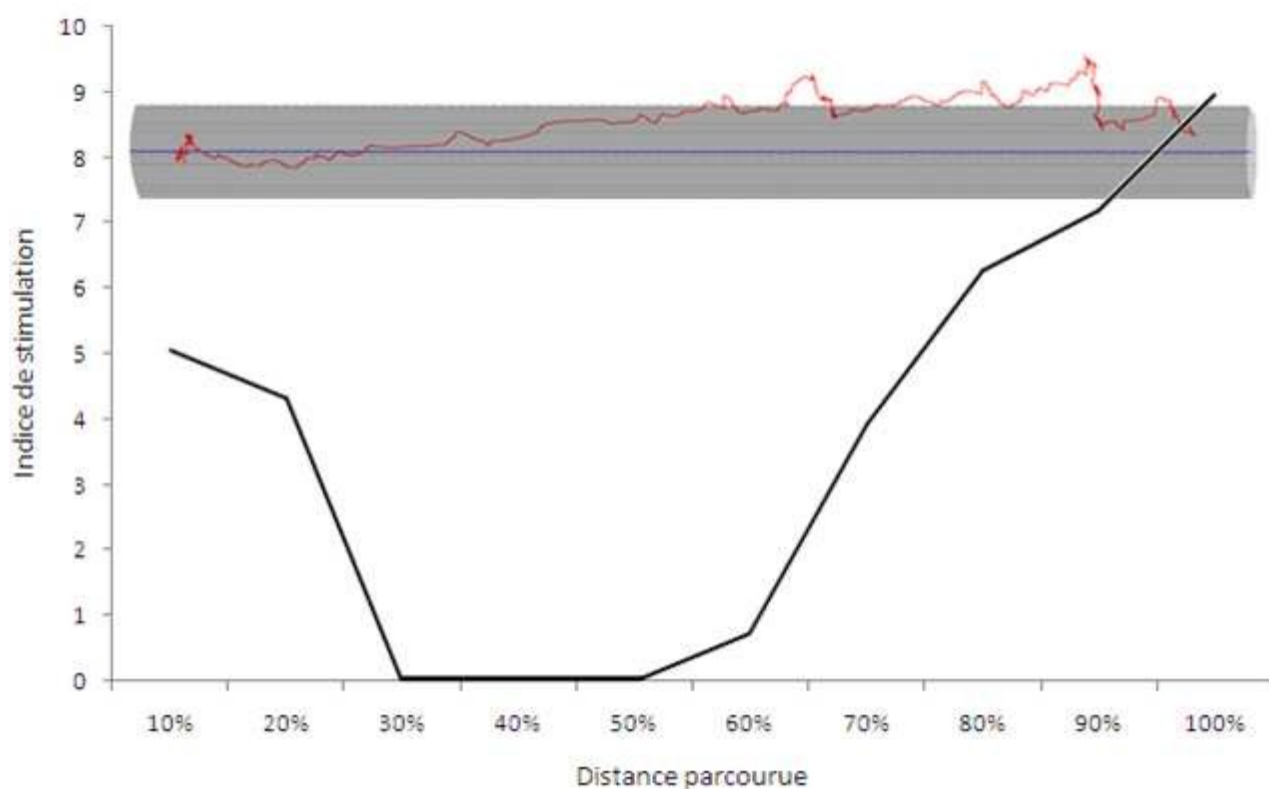


Figure 8.1. Evolution de l'indice de stimulation au cours de la distance parcourue relative (courbe noire) à la droite prédéfinie (ligne bleue). Un exemple de trajectoire (en rouge) et la zone de tolérance (cylindre gris) sont représentés.

A notre connaissance, ce procédé (saccades ou scintillement à faible fréquence) n'a pas encore été implémenté dans un TDU. En revanche, la littérature scientifique fait état des travaux de Loomis et al [Loomis 1986] qui ont utilisé des mouvements fins (tremblements) pour éliminer ce problème de fading dans des applications de stimulation tactile sur l'épiderme par des images. Ces résultats confirment que des saccades permettent le repos des capteurs sensibles aux stimulations vibro-électriques et le maintien de la perception. Un tel procédé pourrait être appliqué à l'ensemble des dispositifs qui se basent sur le paradigme de la substitution visuelle. Il peut aussi être appliqué à d'autres systèmes sensoriels palliatifs comme la vision prothétique. Les récepteurs de la rétine perdent en effet leur sensibilité aux stimuli après des longues stimulations lumineuses ou électriques. Zrenner et son équipe de recherche signalent ce problème et confirment récemment la présence d'un phénomène de saturation de capteurs (Peréz et al 2012, Zrenner et al 2011). L'ajout d'un tel procédé n'a

cependant pas encore été testé. Des saccades ou un scintillement du signal image transmis par une caméra à une rétine artificielle (ex: ARGUS II) ou à un HMD pourrait bien constituer une solution simple à mettre en œuvre.

Dans ces expériences, nous avons aussi observé que la présence de saccades ou de scintillement améliore sensiblement les temps de réponses. Ceux-ci chutent de l'ordre de 500 ms en moyenne. L'usage d'un des deux procédés signale la présence d'un signal aux sujets qui réagissent ainsi plus rapidement. De plus, leur application tout du long de l'expérience prévient la désensibilisation des récepteurs tactiles et permet aux sujets de rester alertes car toujours susceptibles de percevoir le signal.

Ce résultat est très intéressant du point de vue applicatif dans les cas où le temps de réponse est capital, par exemple lorsque la survie de la personne est engagée. C'est le cas dans les applications pour lesquelles un utilisateur doit éviter un obstacle (ex: en aéro et astronautique) ou éviter une chute (cas des sujets dont le sens de l'équilibre est affecté) ; le temps de réaction peut alors s'avérer critique.

Ces expériences nous ont amené à mettre en évidence un phénomène chez certains sujets, phénomène indésirable dans ces premières expériences notamment, « l'effet miroir » dont on montre qu'il est d'autant plus important que les lignes utilisées pour la stimulation sont proche des diagonales. Cet effet est vraisemblablement dû à une difficulté pour les sujets de mettre en correspondance le repère spatial de la matrice du TDU (disposée horizontalement sur la langue) et le repère « haut-bas ». Ce problème avait été évoqué discrètement par Robineau [Robineau, 2009] sans susciter de plus amples investigations car ce phénomène n'avait pas d'influence supposée sur la tâche de guidage du geste chirurgical. Brechmann et al [Brechmann 1989] soulignent aussi ce phénomène mais concluent qu'il peut être corrigé par un processus d'apprentissage. C'est certainement le cas et la meilleure solution (voir les résultats obtenus par Segond [Segond et al., 2012] concernant l'apprentissage dans les tâches de reconnaissance et d'orientation). Rappelons cependant que nous avons montré que ce phénomène persistait en situation de perception bimodale mêlant TDU et VDU. Ce résultat suggère que ce phénomène pourrait s'avérer plus robuste (et résistant à l'apprentissage) qu'attendu.

Dans nos expériences en tout cas, sa présence ne facilite pas l'interprétation des résultats puisque les réponses ainsi « inversées » peuvent être confondues avec des erreurs importantes de perception des angles.

Cela amène ainsi à revoir le design de l'expérience précédente : lorsque les deux lignes croisées sont diagonales, le phénomène d'inversion compromet considérablement l'interprétation des résultats puisqu'on ne sait pas si l'on observe la bonne réponse ou la mauvaise inversée (et réciproque). Une alternative à cette expérience serait de ne tester que des lignes verticales ou horizontales. Par exemple, il serait envisageable de tracer 2 lignes verticales, une à droite et l'autre à gauche de la matrice, d'attendre 2s, 5s ou 10s, et d'effacer une des deux. Cela présenterait le double avantage de ne pas avoir de phénomène d'inversion et de stimuler symétriquement la langue (à supposer que la sensibilité soit la même à gauche à droite), contrairement à un tracé avec 2 lignes horizontales, l'une sur le devant et l'autre à l'arrière de la langue.

Cela amène aussi à poser la question des mécanismes (naturels) et procédés (produits de l'ingénierie) susceptibles d'altérer la perception. Nous nous sommes notamment posé la question de l'effet de la bimodalité perceptive d'une part, et des contradictions d'information d'autre part.

La perception tactile par électro-stimulation linguale est un mode de perception qui pourrait être utilisé dans une diversité d'application par des malvoyants et des aveugles [Bach-y-Rita, 2004; Chebat et al., 2011; Segond et al., 2005], des chirurgiens [Robineau et al., 2007], des soldats [Van Erp and Self, 2008]. Quand un mode de perception (ex: le TDU, l'écoute d'une conversation à la radio ou au téléphone) est redondant ou pire, en contradiction partielle ou totale avec un autre mode de perception que la personne est habituée à utiliser (ex: la vision) pour accomplir ses tâches quotidiennes, des difficultés d'intégration de ces informations multimodales et des situations de conflits sensoriels se posent, la tâche de perception et l'action associée devenant alors plus compliquées.

Les deux expériences que nous avons menées à ce propos s'apparentent à la première expérience décrite dans laquelle on demande aux sujets d'identifier l'orientation (angle) d'une ligne de taxels tracée sur un périphérique de stimulation. Les sujets devaient également répondre le plus rapidement possible. Cette fois cependant, les deux périphériques étaient utilisés en même temps, l'utilisateur percevant le signal-image à la fois sur la langue grâce au

TDU et sur la rétine par l'intermédiaire du VDU. Nous avons réalisé cette expérience sans contradiction, le signal-image se présentant de la même manière sur le TDU et sur le VDU, étudiant ainsi l'effet de la perception bimodale. Nous l'avons aussi réalisée en introduisant des contradictions d'informations entre les deux périphériques, à savoir un décalage angulaire entre les lignes tracées sur le TDU et sur le VDU, décalages de 90°, 75°, 60° ou 45°.

Concernant la perception bimodale, on pourrait s'attendre à des résultats proches de ceux obtenus en vision normale (sans TDU). Nous savons en effet, grâce à la première expérience (identification de l'angle d'une ligne), que le degré de perception est bien moindre (plus d'erreurs plus importantes) avec le TDU qu'en vision normale (identification de l'orientation de la ligne sur le VDU). Cela est attendu, l'œil étant un capteur de grande qualité en comparaison de la langue et un organe sensoriel supposé dominant. Nous avons cependant constaté une augmentation très significative des erreurs, en nombre et en importance. En outre les temps de réponse augmentent dans ces conditions. Enfin, et c'est probablement le plus frappant, certains sujets continuent « d'inverser » certaines réponses sans toutefois ressentir (et signaler) de gêne comme un sentiment de contradiction. Cela signifie que l'usage du TDU en même temps que la vision a un effet significatif : les sujets ne parviennent pas à ignorer la perception linguale, leur perception visuelle étant grandement perturbée. De plus, leur charge cognitive semble grandement accrue comme le montre l'augmentation importante des temps de réponse. Ceci n'est pas sans rappeler des situations connues comme le téléphone au volant. Notre expérience est un excellent exemple des effets de la perception multimodale. Il est probable qu'un apprentissage réduirait cet effet.

L'expérience suivante concernant les contradictions d'information fut encore plus surprenante. Nous avons montré en premier lieu que les sujets répondaient invariablement selon l'une ou l'autre des deux modalités disponibles (TDU ou VDU) et ce de façon équilibrée. Encore une fois, l'œil ne domine pas ; les sujets font autant « confiance » à leur œil qu'à leur langue. En réalité, devant répondre rapidement, ils font probablement un choix frustrant (les sujets se rendent bien compte qu'il y a contradiction) mais demandant davantage de temps, les temps de réponse augmentant aussi à mesure que la contradiction augmente, donc que celle-ci est facilement identifiable : une contradiction de 90° est reconnue comme telle par tous les sujets sans équivoque tandis qu'une contradiction de 45° ne l'est pas forcément, les tracés étant moins éloignés (du point de vue de leur orientation).

Cependant, le résultat le plus frappant est l'obtention de réponses dans lesquelles les perceptions par l'intermédiaire du VDU et par l'intermédiaire du TDU ont fusionné. Nous observons en effet, en plus des réponses principales centrées sur les réponses attendues avec le TDU ou avec le VDU, des réponses intermédiaires en nombre important. Pour identifier ces réponses comme des modes de réponse à part entière, nous avons supposé leurs distributions individuelles (si tant est qu'elles puissent l'être) comme gaussiennes, c'est-à-dire que les réponses principales (TDU et VDU) comme les réponses intermédiaires (fusions) se voient associées chacune à une distribution gaussienne (avec 3 paramètres : moyenne, écart-type et proportion de l'effectif). Dès lors, nous avons été en mesure de rechercher, grâce à un algorithme adapté pour optimiser les paramètres d'une collection de distributions gaussiennes en se fondant sur la minimisation d'un score (l'AIC) prenant en compte le maximum de vraisemblance et la quantité de paramètres libres. Cette hypothèse gaussienne est discutable, mais nous pensons qu'elle est raisonnable étant donné le profil gaussien des distributions d'erreurs d'angles dans les expériences ne faisant intervenir qu'une seule modalité (TDU ou VDU).

Dans tous les cas de contradiction, le résultat est sans appel : tous les modèles de mélanges gaussiens pour obtenir un score d'AIC (meilleure qualité de modèle à moindre coût) font intervenir plus de réponses gaussiennes que les réponses principales associées au TDU et au VDU. Cela signifie que ces modes de réponse intermédiaires sont significatifs : les sujets formulent parfois des réponses sous forme de tracés dont la valeur angulaire est intermédiaire entre celle attendue pour les deux modes de perception. Nous avons en outre souvent pu identifier avec une certaine fiabilité ces réponses intermédiaires qui apparaissent comme des réponses proches du TDU mais influencées par la perception via le VDU (et réciproque). Il s'agit vraisemblablement d'une démonstration qu'un effet semblable à celui décrit par Mc Gurck et Mc Donald (contradictions visuo-auditives) peut être obtenu par une contradiction entre la vision et la vision substitutive.

Terminons ce retour sur les expériences par une remarque plus générale sur le paradigme de substitution visuelle. Ce paradigme, tel que formulé par P. Bach-y-Rita, dit que c'est le cerveau qui voit, pas l'organe.

Nos expériences confirment l'efficacité de la substitution visuelle par stimulation linguale. Le fait que dans l'expérience de contradiction les sujets choisissent le TDU ou le

VDU comme référence avec une même fréquence est marquant. De même en ce qui concerne le maintien du phénomène d'inversion dans l'expérience de perception bimodale. Il est maintenant sensé de dire que la cavité buccale voit, en l'occurrence qu'il s'agit d'un organe permettant un niveau d'acquisition d'informations spatiales proche de ceux de la vision et de l'audition ! Nous parvenons en effet à percevoir des textures avec notre cavité buccale (voûte palatine, joues et langue). Par exemple quand nous mangeons du poisson, nous parvenons sans peine à détecter les arêtes, à les localiser, voire même à en identifier la forme et la taille et les isoler pour s'en débarrasser, tout cela grâce à la perception fine de notre langue assistée du reste de la cavité buccale. [Warwick et Jeffs 2013] affirment que notre cavité buccale, langue incluse, est le prolongement d'un système ancestral de perception ayant une influence importante sur la survie et la santé des individus, et affirment également qu'il est possible que le cerveau soit ainsi équipé pour reconnaître et analyser des formes observées par ce système. De façon similaire, Hubel et al [Hubel 1994] ont mentionné que la rétine est un prolongement du cerveau, pré-traitant la complexité des tâches liées à la vision.

Cependant, tel quel, ce paradigme est incomplet : nous avons en effet montré que l'organe, l'œil en l'occurrence, possède des caractéristiques, motrices notamment (micro-saccades, dérives, etc.), nécessaires à une bonne perception. Ce sont des particularités que ne possèdent pas tous les organes (tels la peau ou la langue) pouvant servir de substituts à la perception visuelle.

Allons maintenant au-delà des résultats et concluons sur le rôle que peuvent jouer de telles recherches.

L'objectif initial de cette thèse était d'apporter une contribution scientifique et technologique qui aurait permis d'envisager l'utilisation du TDU (ou de toute autre technologie voisine) dans des applications médicales « grand public » comme fournir aux aveugles un moyen de recouvrer une partie de leur vision, ou dans des applications de réalité augmentée. Néanmoins on peut vraiment se poser la question de la pertinence d'une telle technologie pour de telles applications. En effet, à moins de faire évoluer, avec les appuis nécessaires cette technologie dans un cycle de production industriel, il est peu probable que le TDU (qui pose de nombreux problèmes d'acceptabilité liés au fait que son utilisation fait saliver, qu'on ne peut pas parler avec, qu'on soit identifié comme handicapé...) soit déployé

vers le marché médical pour des personnes malvoyantes ou aveugles. Pour cela, il faudrait, comme dans la majorité des « innovations » technologiques grand public actuelles, forcer son utilisation au détriment d'autres dispositifs comme les cannes blanches. Ce n'est bien entendu pas réaliste.

En revanche les résultats de cette étude ont remis le statut de la langue comme comparable sur de nombreux aspects à celui de l'organe auquel elle se substitue, l'oeil. De tels travaux sont donc un excellent moyen d'investiguer le rapport à nos sens. Cette thèse ouvre vers de nombreuses possibilités de continuation que nous avons commencé à explorer.

Nous avons en effet continué dans une direction technologiques d'une part : Nous avons entrepris en 2011-2012 la réalisation d'un TDU haute résolution (24 x 24 taxels) et dont le prototype est fonctionnel (Fig. 8.2). Nous avons aussi réalisé un système couplant la perception au geste. Une tablette graphique et un stylet permettent de dessiner librement, le tracé étant ressenti instantanément sur le TDU.

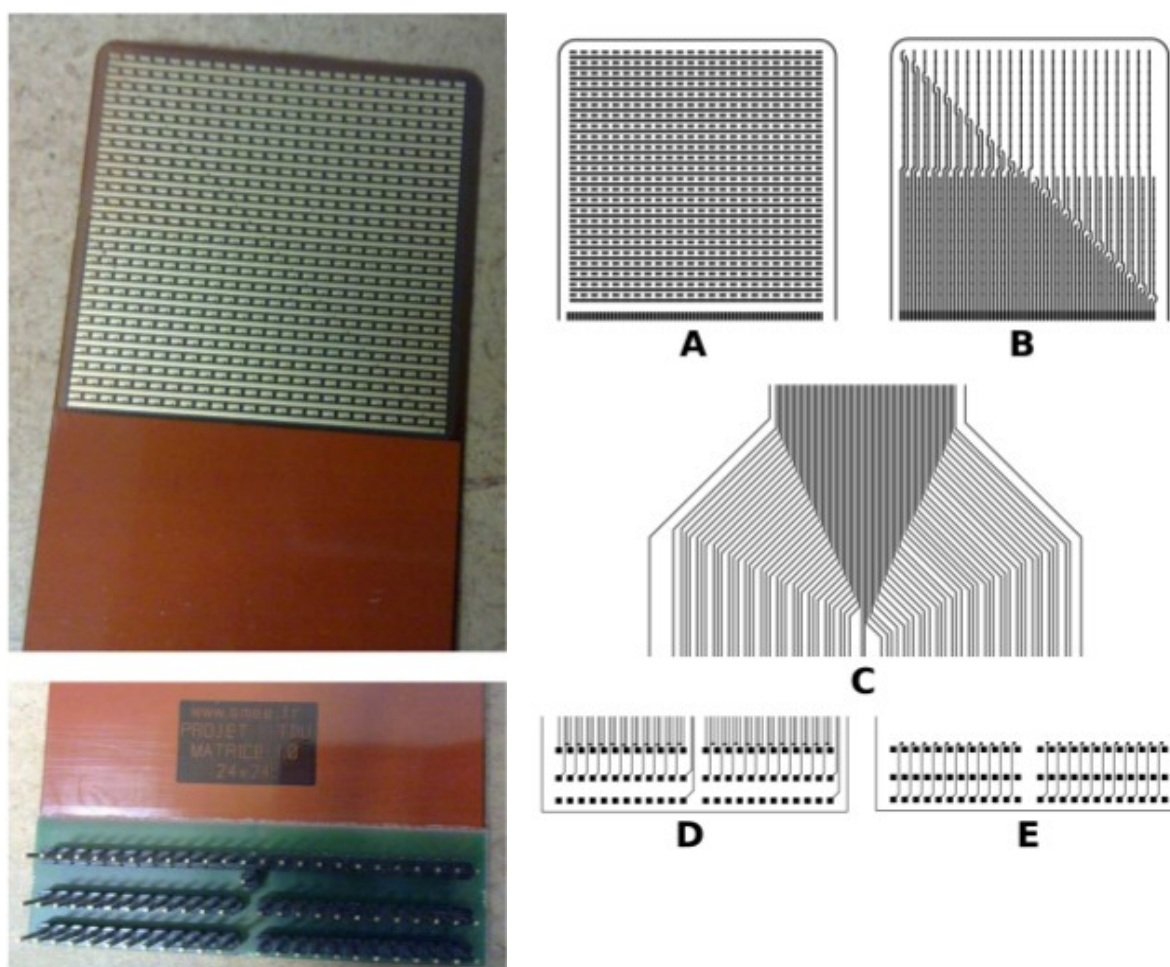


Figure 8.2. Matrice haute résolution (24x24 taxels) et son connecteur. Photographies à gauche ; schémas à droite.

Nous avons également entamé des recherches de psychomotricité d'une part en ce qui concerne le suivi de trajectoires bi et tri-dimensionnelles, d'autre part en nous intéressant au volet cérébral de la perception par l'intermédiaire de l'imagerie fonctionnelle.

Des expériences de suivi de trajectoires ont ainsi été réalisées (Fig. 8.3) et ont été publiées [Chekhchoukh 2011, Chekhchoukh 2013]. Dans ces expériences, nous affichons dans un espace en trois dimensions une série de points espacés que nous nommons « chemin lumineux » et qui s'apparente à un sentier parsemé de cailloux blancs. Le sujet, en utilisant un périphérique de saisie (soit le couple clavier+souris, soit le couple corps rigide+caméra stéréoscopique) déplace son point de vue et son angle de vue, et doit ainsi parcourir ce « chemin » d'un point de départ à une cible. Nous avons pu montrer l'efficacité de ce dispositif

qui pourrait servir dans des applications de navigation autonome, c'est-à-dire sans besoin d'être guidé et en utilisant seulement un ensemble de chemins prédéfinis.

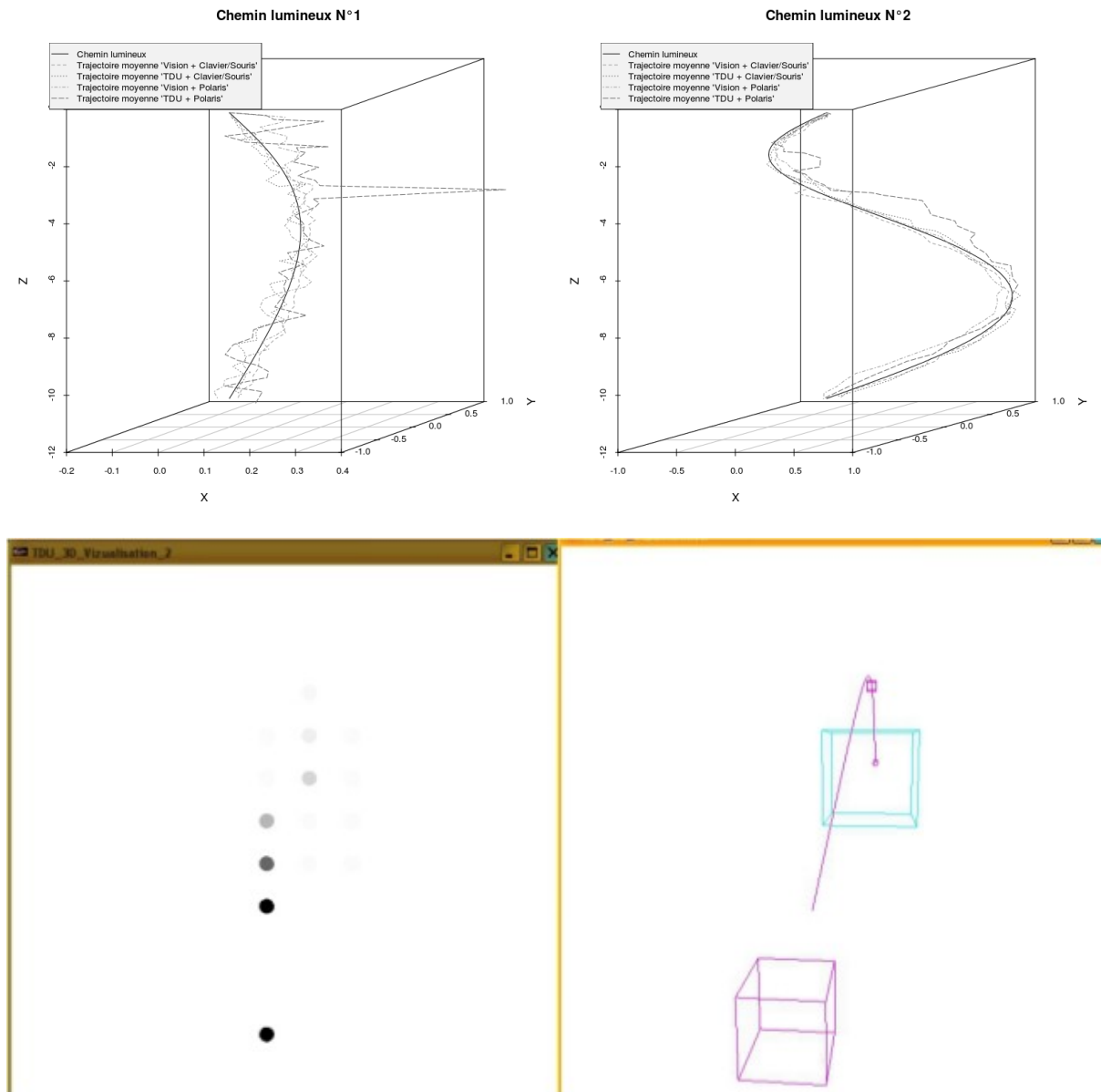


Figure 8.3. (Haut) Chemins lumineux (ligne pleine) et trajectoires moyennes suivies en fonction du mode de perception et du périphérique de saisie utilisé. Deux types de chemins sont montrés. Le parcours est d'environ 10 cm. (Bas à gauche) aspect du chemin lumineux tel que vu (ou perçu s'il s'agit du TDU) par le sujet. L'atténuation de l'intensité des points permet d'encoder la distance. (Bas à droite) position effective (petit carré) sur le chemin théorique. Les grands cubes matérialisent le départ et la cible.

Enfin, nous avons commencé à nous attaquer au versant anatomo-fonctionnel de ce travail. Parmi les méthodes utilisées pour mieux comprendre le fonctionnement cortical de la perception, on trouve l'Imagerie par Résonance Magnétique Fonctionnelle (IRMf) ainsi que

les techniques d'électrophysiologie. Afin d'investiguer l'activation des aires visuelles pendant une perception que nous supposons visuelle par électro-stimulation linguale, des expériences de reconnaissance de forme géométriques, ainsi que des tâches de détection d'orientation, ont été mise en œuvre dans le cadre d'une collaboration avec C. Scheiber au CHU de Bron. Ces expériences se font sous un IRMF 3 Tesla utilisant un TDU 12 x 12 compatible avec cet environnement magnétique.

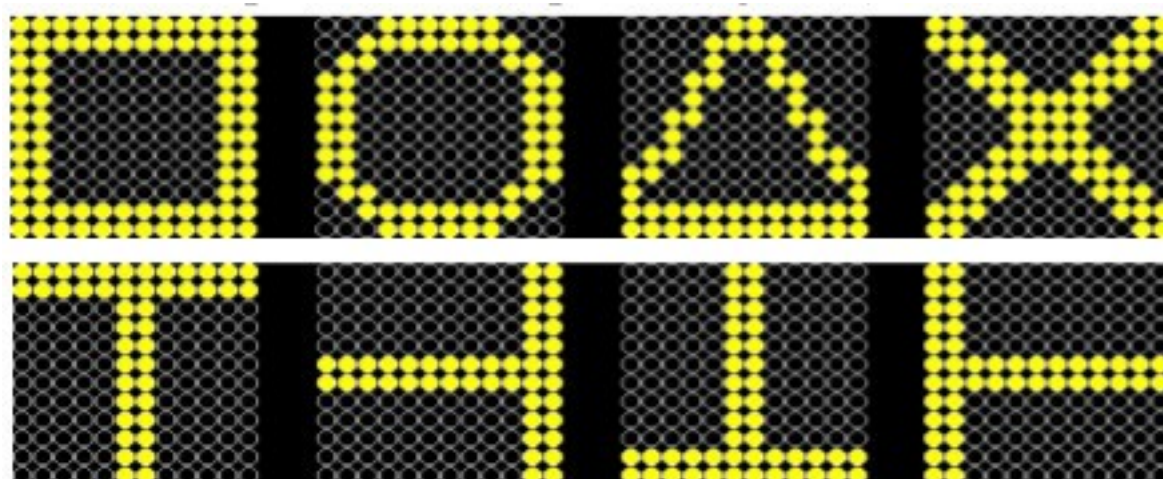


Figure 8.4. Deux listes de motifs pouvant être sélectionnés (choix de réponse). En haut : reconnaissance d'une forme géométrique simple (carré, cercle, triangle, croix). En bas : détection de l'orientation de la lettre T (0°, 90°, 180° et 270°)

Le volet anatomo-fonctionnel de la perception visuelle par électro-stimulation linguale a été initié pendant ma thèse mais malheureusement arrêté en cours de route, le TDU compatible avec l'IRMf n'étant pas au point pour accomplir l'expérience. Nous avons néanmoins désigné cette expérience conçue comme une tâche de détection de rotations et de reconnaissance d'objets. Des travaux antérieurs [Ptito M et al 2009] ont montré que chez les aveugles nés et même chez des sujets sains, le cortex visuel est bien activé pendant une expérience de détection à l'aide des stimulateurs vibro-tactiles sur l'abdomen des sujets. Nous souhaitons faire de même avec notre TDU. À ce stade, l'ensemble des scénarios des expériences est prêt, le code de contrôle du TDU et de l'IRMf est fonctionnel, et l'autorisation du CPP est accordée.

Ceux qui poursuivront ces travaux (trajectoires, IRMf ...) ont encore beaucoup de travail. Je leur souhaite de garder ce cap ... peut-être en se servant d'un TDU !

V - Bibliographie

- Bach-y-Rita, P., 2004. Tactile sensory substitution studies. *Ann.-N. Y. Acad. Sci.* 1013, 83–91.
- Bach-y-Rita, P., Collins, C.C., Saunders, F.A., White, B., Scadden, L., 1969. Vision substitution by tactile image projection. *Nature* Volume 221, pages 963, 964
- BeMent, S.L., Ranck Jr, J.B., 1969. A model for electrical stimulation of central myelinated fibers with monopolar electrodes. *Exp. Neurol.* 24, 171–186.
- Blake, D.T., Hsiao, S.S., Johnson, K.O., 1997. Neural coding mechanisms in tactile pattern recognition: the relative contributions of slowly and rapidly adapting mechanoreceptors to perceived roughness. *J. Neurosci.* 17, 7480–7489.
- Blamey, P., Clark, G., 1985. A wearable multiple-electrode electrotactile speech processor for the profoundly deaf. *J. Acoust. Soc. Am.* 77, 1619.
- Bliss, J.C., Katcher, M.H., Rogers, C.H., Shepard, R.P., 1970. Optical-to-tactile image conversion for the blind. *Man-Mach. Syst. IEEE Trans. On* 11, 58–65.
- Bluteau, J., 2010. Evaluation des effets de l'ajout d'interfaces haptiques sur le suivi manuel de trajectoires. Institut National Polytechnique de Grenoble-INPG.
- Bluteau, J., Dubois, M., Coquillart, S., Gentaz, E., Payan, Y., 2010. Vibrotactile guidance for trajectory following in computer aided surgery, in: *Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE. IEEE*, pp. 2085–2088.
- Butikofer, R., Lawrence, P.D., 1978. Electrocutaneous nerve stimulation-I: Model and experiment. *Biomed. Eng. IEEE Trans. On* 526–531.
- Button, J., Putnam, T., 1962. Visual responses to cortical stimulation in the blind. *J Iowa Med Soc* 52, 17–21.
- Cha, K., Horch, K., Normann, R.A., 1992a. Simulation of a phosphene-based visual field: visual acuity in a pixelized vision system. *Ann. Biomed. Eng.* 20, 439–449.
- Cha, K., Horch, K.W., Normann, R.A., Boman, D.K., 1992b. Reading speed with a pixelized vision system. *JOSA A* 9, 673–677.
- Challis, B.P., Edwards, A.D., 2001. Design principles for tactile interaction, in: *Haptic Human-Computer Interaction. Springer*, pp. 17–24.
- Chebat, D.-R., Schneider, F.C., Kupers, R., Ptito, M., 2011. Navigation with a sensory substitution device in congenitally blind individuals. *Neuroreport* 22, 342–347.
- Chekhchoukh, A., Glade, N., 2012. Influence of Sparkle and Saccades on Tongue Electro-Stimulation-Based Vision Substitution of 2D Vectors. *Acta biotheoretica* 60, 41-53.

- Chenu, O., Vuillerme, N., Demongeot, J., Payan, Y., 2009. A wireless lingual feedback device to reduce overpressures in seated posture: A feasibility study. *PloS One* 4, e7550.
- Chow, A.Y., Peachey, N.S., 1998. The subretinal microphotodiode array retinal prosthesis. *Ophthalmic Res.* 30, 195–196.
- Cohen, S.Y., Desmettre, T., 2005. *Dégénérescence maculaire liée à l'âge*. Elsevier.
- Collins, C.C., 1970. Tactile television-mechanical and electrical image projection. *Man-Mach. Syst. IEEE Trans. On* 11, 65–71.
- Cornsweet, T.N., 1956. Determination of the stimuli for involuntary drifts and saccadic eye movements. *JOSA* 46, 987–988.
- Diderot, D., 1901. *Oeuvres choisies de Diderot*. Garnier frères.
- DIES, J.S. 1965. Possibilités d'une Prothèse Sensorielle Acoustique par Stimulation Tactile. *Cybern. Nerv. Syst.* 17, 169.
- Dobelle, W.H., 2000. Artificial vision for the blind by connecting a television camera to the visual cortex. *ASAIO J.* 46, 3–9.
- Dodge, R., 1907. An experimental study of visual fixation. *Psychol. Monogr. Gen. Appl.* 8, i–95.
- Edwards, W.K., Mynatt, E.D., 1994. An architecture for transforming graphical interfaces, in: *Proceedings of the 7th Annual ACM Symposium on User Interface Software and Technology*. ACM, pp. 39–47.
- Ernst, M.O., Banks, M.S., 2002. Humans integrate visual and haptic information in a statistically optimal fashion. *Nature* 415, 429–433.
- Erp, V., Jan, B., Veltman, J., van Veen, H., Oving, A., 2003. Tactile torso display as countermeasure to reduce night vision goggles induced drift. DTIC Document.
- Flament, J., 2002. *Ophtalmologie: pathologie du système visuel*. Masson.
- Gibson, R.H., 1968. Electrical stimulation of pain and touch. *Skin Senses* 223–261.
- Gonzalez-Cordero, A., West, E.L., Pearson, R.A., Duran, Y., Carvalho, L.S., Chu, C.J., Naeem, A., Blackford, S.J., Georgiadis, A., Lakowski, J., others, 2013. Photoreceptor precursors derived from three-dimensional embryonic stem cell cultures integrate and mature within adult degenerate retina. *Nat. Biotechnol.* 31, 741–747.
- Howe, R.D., Matsuoka, Y., 1999. Robotics for surgery. *Annu. Rev. Biomed. Eng.* 1, 211–240.
- Hubel, D., This, P., 1994. *L'oeil, le cerveau et la vision: les étapes cérébrales du traitement visuel*. Pour la science.

- Hughes, R.G., Forrest, A.R., 1996. Perceptualisation using a tactile mouse, in: Proceedings of the 7th Conference on Visualization'96. IEEE Computer Society Press, p. 181–ff.
- Jansson, G., Brabyn, L., 1981. Tactually guided batting. Department [Psychologiska inst., Uppsala univ.].
- Javaheri, M., Hahn, D.S., Lakhanpal, R.R., Weiland, J.D., Humayun, M.S., 2006. Retinal prostheses for the blind. *Ann.-Acad. Med. Singap.* 35, 137.
- Kieninger, T., 1996. The “growing up” of HyperBraille—an office workspace for blind people, in: Proceedings of the 9th Annual ACM Symposium on User Interface Software and Technology. ACM, pp. 67–73.
- Kien, T.T., Maul, T., Bargiela, A., 2012. A review of retinal prosthesis approaches, in: *International Journal of Modern Physics: Conference Series*. World Scientific, pp. 209–231.
- Kobayashi, M., Watanabe, T., 2002. A tactile display system equipped with a pointing device—MIMIZU, in: *Computers Helping People with Special Needs*. Springer, pp. 527–534.
- Kowler, E., Steinman, R.M., 1980. Small saccades serve no useful purpose: reply to a letter by RW Ditchburn. *Vision Res.* 20, 273–276.
- Lenay, C., Gapenne, O., Hanneton, S., Marque, C., Genouëlle, C., Perceptive, G.S., 2000. La substitution sensorielle: limites et perspectives. *Touch. Pour Connaître Psychol. Cogn. Percept. Tactile Man.* Paris PUF 287–306.
- Loomis, J.M., Klatzky, R.L., Lederman, S.J., 1991. Similarity of tactual and visual picture recognition with limited field of view. *Perception* 20, 167–177.
- Margalit, E., Maia, M., Weiland, J.D., Greenberg, R.J., Fujii, G.Y., Torres, G., Piyathaisere, D.V., O’Hearn, T.M., Liu, W., Lazzi, G., others, 2002. Retinal prosthesis for the blind. *Surv. Ophthalmol.* 47, 335–356.
- Martinez-Conde, S., Macknik, S.L., Hubel, D.H., 2004. The role of fixational eye movements in visual perception. *Nat. Rev. Neurosci.* 5, 229–240.
- Martinez-Conde, S., Macknik, S.L., Hubel, D.H., 2000. Microsaccadic eye movements and firing of single cells in the striate cortex of macaque monkeys. *Nat. Neurosci.* 3, 251–258.
- Martinez-Conde, S., MACKNIK, T.N., 2007. Les mouvements secrets de l’œil. *Pour Sci.* 56–61.
- Mason, J.L., MacKay, N.A., 1976. Pain sensations associated with electrocutaneous stimulation. *Biomed. Eng. IEEE Trans. On* 405–409.
- McGurk, H., MacDonald, J., 1976. Hearing lips and seeing voices.
- Meffin, H., Tahayori, B., Grayden, D.B., Burkitt, A.N., 2013. Internal inconsistencies in models of electrical stimulation in neural tissue, in: *Engineering in Medicine and Biology Society (EMBC), 2013 35th Annual International Conference of the IEEE*. IEEE, pp. 5946–5949.

- Meffin, H., Tahayori, B., Grayden, D.B., Burkitt, A.N., 2012. Modeling extracellular electrical stimulation: I. Derivation and interpretation of neurite equations. *J. Neural Eng.* 9, 065005.
- Meijer, P.B., 1992. An experimental system for auditory image representations. *Biomed. Eng. IEEE Trans. On* 39, 112–121.
- Miletic, G., Hughes, B., Bach-y-Rita, P., 1988. Vibrotactile stimulation: An educational program for spatial concept development. *J. Vis. Impair. Blind.*
- Palacz, O., Kurcz, E., 1978. Przydatnosc zmodyfikowanego elektroftalmu EL-300 wg Starkiewicza dla niewidomych (The usefulness of modified Electrophthalm designed by Starkiewicz for the blind). *Klin. Oczna* 48, 61–63.
- Parier, V., Soubrane, G., 2008. La dégénérescence maculaire liée à l'âge. *Rev. Médecine Interne* 29, 215–223.
- Petrie, H., Morley, S., McNally, P., Graziani, P., Emiliani, P.L., 1996. Access to hypermedia systems for blind students. *Colloq.-Inst. Natl. SANTE Rech. MEDICALE Colloq. Semin.* 223–228.
- Petrie, H., Morley, S., Weber, G., 1995. Tactile-based direct manipulation in GUIs for blind users, in: *Conference Companion on Human Factors in Computing Systems. ACM*, pp. 428–429.
- Peyman, G., Chow, A., Liang, C., Chow, V., Perlman, J., Peachey, N., 1998. Subretinal semiconductor microphotodiode array. *Ophthalmic Surg. Lasers* 29, 234–241.
- Pfeiffer, E.A., 1968. Electrical stimulation of sensory nerves with skin electrodes for research, diagnosis, communication and behavioral conditioning: A survey. *Med. Biol. Eng.* 6, 637–651.
- Pritchard, R.M., 1961. *Stabilized images on the retina.* WH Freeman Company.
- Purves, D., Augustine, G., Fitzpatrick, D., Hall, W., LaMantia, A., McNamara, J., 2005. *Neurosciences (De Boeck Université, 3e édition ed.)*.
- Ramstein, C., 1996. Combining haptic and braille technologies: design issues and pilot study, in: *Proceedings of the Second Annual ACM Conference on Assistive Technologies. ACM*, pp. 37–44.
- Rattay, F., 1998. Analysis of the electrical excitation of CNS neurons. *Biomed. Eng. IEEE Trans. On* 45, 766–772.
- Ringel, R.L., Ewanowski, S.J., 1974. Oral perception: 1. Two-point discrimination. *Speech Hear. Sci. Sel. Read.* 8, 362.
- Roberts, J., Slattery, O., Kardos, D., 2000. 49.2: Rotating-Wheel Braille Display for Continuous Refreshable Braille, in: *SID Symposium Digest of Technical Papers. Wiley Online Library*, pp. 1130–1133.
- Robineau, F., 2009. Etude d'un dispositif de guidage de geste chirurgical de ponction par stimulation électrotactile linguale. *Université Joseph-Fourier-Grenoble I.*

- Robineau, F., Boy, F., Orliaguet, J.-P., Demongeot, J., Payan, Y., 2007. Guiding the surgical gesture using an electro-tactile stimulus array on the tongue: A feasibility study. *Biomed. Eng. IEEE Trans. On* 54, 711–717.
- Rock, I., Victor, J., 1964. Vision and touch: An experimentally created conflict between the two senses. *Science* 143, 594–596.
- Rollman, G.B., 1974. Electrocutaneous stimulation.
- Sampaio, E., Dufier, J., 1988. Suppléance sensorielle électronique pour les jeunes enfants aveugles. *J. Fr. Ophthalmol.* 11, 161–167.
- Schubert, M., Hierzenberger, A., Lehner, H., Werner, J., 1999a. Optimizing photodiode arrays for the use as retinal implants. *Sens. Actuators Phys.* 74, 193–197.
- Schubert, M., Stelzle, M., Graf, M., Stert, A., Nisch, W., Graf, H., Hammerle, H., Gabel, V., Hofflinger, B., Zrenner, E., 1999b. Subretinal implants for the recovery of vision, in: *Systems, Man, and Cybernetics, 1999. IEEE SMC'99 Conference Proceedings. 1999 IEEE International Conference on. IEEE*, pp. 376–381.
- Segond, H., Weiss, D., Kawalec, M., Sampaio, E., 2012. Perceiving space and optical cues via a visuo-tactile sensory substitution system: a methodological approach for training of blind subjects for navigation. *Perception* 42, 508—528.
- Segond, H., Weiss, D., Sampaio, E., 2005. Human spatial navigation via a visuo-tactile sensory substitution system. *Percept.-Lond.-* 34, 1231.
- Shimoga, K.B., 1992. Finger force and touch feedback issues in dexterous telemanipulation, in: *Intelligent Robotic Systems for Space Exploration, 1992. Proceedings. Fourth Annual Conference on. IEEE*, pp. 159–178.
- Skavenski, A., Hansen, R., Steinman, R.M., Winterson, B.J., 1979. Quality of retinal image stabilization during small natural and artificial body rotations in man. *Vision Res.* 19, 675–683.
- Solomonow, M., Lyman, J., Freedy, A., 1977. Electrotactile two-point discrimination as a function of frequency, body site, laterality, and stimulation codes. *Ann. Biomed. Eng.* 5, 47–60.
- Steinman, R.M., Collewijn, H., 1980. Binocular retinal image motion during active head rotation. *Vision Res.* 20, 415–429.
- Szlavik, R.B., de Bruin, H., 2000. The effect of anisotropy on the potential distribution in biological tissue and its impact on nerve excitation simulations. *Biomed. Eng. IEEE Trans. On* 47, 1202–1210.
- Tatur G, D.M., Marc I, Lafon D, Dupeyron G, Bardin F, 2011. Une approche fonctionnelle en vision prothétique : Etude préliminaire dans le contexte de la mobilité. pp. 6029–6032.
- Tomasi, C., 2004. Estimating Gaussian mixture densities with em—a tutorial. *Duke Univ.*
- Troxler, D., 1804. Über das Verschwinden gegebener Gegenstände innerhalb unseres Gesichtskreises. *Ophthalmol. Bibl.* 2, 1–53.

- Trulsson, M., Essick, G.K., 1997. Low-threshold mechanoreceptive afferents in the human lingual nerve. *J. Neurophysiol.* 77, 737–748.
- Van Erp, J., Self, B., 2008. Tactile displays for orientation, navigation and communication in air, sea and land environments.
- Velázquez, R., Pissaloux, E.E., Wiertelowski, M., 2006. A compact tactile display for the blind with shape memory alloys, in: *Robotics and Automation, 2006. ICRA 2006. Proceedings 2006 IEEE International Conference on. IEEE*, pp. 3905–3910.
- Veraart, C., Raftopoulos, C., Mortimer, J.T., Delbeke, J., Pins, D., Michaux, G., Vanlierde, A., Parrini, S., Wanet-Defalque, M.-C., 1998. Visual sensations produced by optic nerve stimulation using an implanted self-sizing spiral cuff electrode. *Brain Res.* 813, 181–186.
- Vuillerme, N., Chenu, O., Demongeot, J., Payan, Y., 2006. Improving human ankle joint position sense using an artificial tongue-placed tactile biofeedback. *Neurosci. Lett.* 405, 19–23.
- Vuillerme, N., Chenu, O., Pinsault, N., Moreau-Gaudry, A., Fleury, A., Demongeot, J., Payan, Y., 2007. Pressure sensor-based tongue-placed electrotactile biofeedback for balance improvement-Biomedical application to prevent pressure sores formation and falls, in: *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE. IEEE*, pp. 6113–6116.
- Waschkowski, F., Brockmann, C., Laube, T., Mokwa, W., Roessler, G., Walter, P., 2013. Development of a very large array for retinal stimulation, in: *Engineering in Medicine and Biology Society (EMBC), 2013 35th Annual International Conference of the IEEE. IEEE*, pp. 2748–2751.
- Yarbus, A.L., Haigh, B., Riggs, L.A., 1967. *Eye movements and vision*. Plenum press New York.
- Yobas, L., Durand, D.M., Skebe, G.G., Lisy, F.J., Huff, M.A., 2003. A novel integrable microvalve for refreshable braille display system. *Microelectromechanical Syst. J. Of* 12, 252–263.
- Zrenner, E., Stett, A., Weiss, S., Aramant, R., Guenther, E., Kohler, K., Miliczek, K.-D., Seiler, M., Haemmerle, H., 1999. Can subretinal microphotodiodes successfully replace degenerated photoreceptors? *Vision Res.* 39, 2555–2567.

VI - Annexes

Questionnaire

Merci de répondre honnêtement aux questions suivantes :

1/ Évaluez la facilité de définir la direction de la ligne en mode Visuel

très difficile

très facile

0	1	2	3	4
---	---	---	---	---

2/ Évaluez la facilité de définir la direction de la ligne en mode Hybride (Visuel+TDU)

très difficile

très facile

0	1	2	3	4
---	---	---	---	---

3/ Évaluez la facilité de définir la direction de la ligne en mode TDU

très difficile

très facile

0	1	2	3	4
---	---	---	---	---

4/ Dans le mode TDU, pensez vous être parvenus à « visualiser » une ligne ?

Si oui, avec ou sans saccades (et intensité) ?

Avec ou sans scintillement (intensité) ?

Quelles orientations sont plus faciles à visualiser (horizontales, verticales, diagonales) ?

5/ D'une séance à l'autre pensez vous vous être amélioré(e) ?

6/ Pensez vous que votre calibration était adaptée ?

pas adaptée

parfaite

0	1	2	3	4
---	---	---	---	---

7/ Évaluez l'ergonomie générale du dispositif pour le test que vous venez de faire.

pas adaptée

parfaite

0	1	2	3	4
---	---	---	---	---

8/ vos remarques et suggestions

Merci pour votre participation