



**HAL**  
open science

# Intégration de l'interaction au regard dans des systèmes optroniques : évaluation de l'influence du contexte

Romain Grosse

► **To cite this version:**

Romain Grosse. Intégration de l'interaction au regard dans des systèmes optroniques : évaluation de l'influence du contexte. Interface homme-machine [cs.HC]. Université de Technologie de Compiègne, 2018. Français. NNT : 2018COMP2418 . tel-02060996

**HAL Id: tel-02060996**

**<https://theses.hal.science/tel-02060996>**

Submitted on 7 Mar 2019

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

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

Par **Romain GROSSE**

*Intégration de l'interaction au regard dans des systèmes optroniques : évaluation de l'influence du contexte*

Thèse présentée  
pour l'obtention du grade  
de Docteur de l'UTC



Soutenue le 9 avril 2018

**Spécialité** : Informatique et Sciences et Technologies de l'Information et des Systèmes : Unité de recherche Heudyasic (UMR-7253)

D2418



UNIVERSITÉ DE TECHNOLOGIE DE COMPIÈGNE  
&  
SAFRAN ELECTRONICS & DEFENSE

THÈSE DE DOCTORAT EN INFORMATIQUE ET SCIENCES ET  
TECHNOLOGIES DE L'INFORMATION ET DES SYSTÈMES

Présentée par

Romain GROSSE

# **Intégration de l'interaction au regard dans des systèmes optroniques : évaluation de l'influence du contexte**

*Présentée et soutenue publiquement le 9 avril 2018 devant un jury composé de :*

*Président :*

*Directeurs de thèse :*

Dominique LENNE

Indira THOUVENIN

*Rapporteurs :*

Gaëlle CALVARY

Christophe KOLSKI

*Examineurs :*

Olivier GAPENNE

Patrick LECALLET

Stéphane AUBRY



# Remerciements

Parce qu'une thèse ne se fait pas seul, merci.

Merci tout d'abord à Safran Electronic & Defense d'avoir monté ce projet de thèse et de m'avoir sélectionné et fait confiance pour mener à bien ces travaux. Notamment François Coursaget, Laurent Huré et David Renaudet, merci de m'avoir tour à tour accueilli dans vos services.

Merci à l'UTC d'avoir permis l'existence de cette thèse, en particulier mes directeurs Indira et Dominique pour avoir su, malgré les difficultés, retrouver en moi l'énergie nécessaire pour mener la thèse jusqu'au bout.

Merci, évidemment, à Stéphane Aubry, pour m'avoir sélectionné d'abord, puis pour tous les échanges que nous avons eus. Tu as été un exemple à suivre pour moi durant cette thèse, et je suis heureux de te compter parmi mes amis maintenant. Je sais que tu m'as donné tout ce que tu as pu.

Sur le plan personnel, je n'aurai jamais pu réaliser ce travail sans le soutien et la bonne humeur infaillible de mes amis. Le CAT, mes frères, la team pas David, la colloc, Pessac, mes parents. Merci.

Merci enfin à Lucie pour sa compréhension, son soutien infaillible et son amour.



# Résumé

Les nouvelles versions de produits optroniques de Safran Electronic & Defense comme les jumelles multifonctions ou les lunettes de visée disposent de plus en plus de fonctionnalités, ce qui rend nécessaire une amélioration des dispositifs d'interaction mis à disposition des utilisateurs. L'intégration du regard comme modalité d'interaction semble notamment intéressante en raison de son caractère rapide, naturel et disponible.

Si pour des utilisateurs handicapés, l'interaction au regard est déjà bien développée, elle n'est pas encore une technologie mature pour des personnes valides. Lors d'interactions actives au regard, expressions explicites de l'intention d'agir d'un utilisateur, un problème appelé *Midas Touch* apparaît : il s'agit de l'incapacité pour un utilisateur à dissocier les phases d'analyse et les phases d'action, parce que l'œil est avant tout un organe senseur. Plusieurs modalités d'interaction au regard cherchent à outrepasser ce problème : on peut par exemple utiliser un temps de fixation minimum sur un item pour l'activer (*DwellTime*) ou un temps de fixation minimum sur un item spécifique disposé à côté de l'item d'intérêt (*DwellTime* délocalisé) ou encore associer une autre modalité d'interaction pour spécifier l'intention d'activation (multimodalité œil-bouton).

Chacune de ces modalités dispose d'avantages et d'inconvénients spécifiques, et déterminer la modalité d'interaction la plus adaptée n'est pas une question triviale. C'est d'autant plus difficile que les performances des modalités semblent dépendre de facteurs extérieurs variables, c'est-à-dire du contexte d'utilisation de la modalité. Afin de mieux intégrer le suivi du regard dans des systèmes et de choisir quelle modalité utiliser, il est nécessaire de bien comprendre quels sont les éléments du contexte de l'interaction et comment ils agissent sur les modalités.

Le but de cette thèse est de modéliser le contexte de la modalité d'interaction, c'est-à-dire de déterminer l'ensemble des éléments extérieurs à la modalité pouvant en influencer les performances. A partir d'un état de l'art et d'une étude de l'interaction au regard, nous proposons une séparation de ce contexte en quatre axes : l'utilisateur, la tâche, le système et l'environnement. Chacun de ces axes correspond à un ensemble de caractéristiques dont l'influence est justifiée par des travaux antérieurs ou par des raisonnements théoriques.

Nous étudions ensuite, au travers d'expérimentations, les caractéristiques qui nous ont semblé les plus pertinentes pour introduire le suivi du regard dans les systèmes optroniques, et nous comparons les modalités d'interaction au regard citées précédemment vis-à-vis de ces caractéristiques. La première caractéristique étudiée est le type de menu (linéaire ou circulaire). Contrairement à la souris, les modalités d'interaction au regard ne présentent pas de différences significatives en fonction du type de menu sur lequel elles sont utilisées. La seconde caractéristique étudiée est un des types de tâche visuelle de notre modèle : la réponse à des alertes visuelles attendues. Il s'agit d'étudier l'adéquation de l'interaction au regard avec la capacité à diviser son attention visuelle, c'est-à-dire à fixer un endroit tout en restant concentré sur un autre endroit du champ de vision. Cette capacité est nécessaire lors de tâches d'observation ou de suivi de cibles. Il semble que les interactions à base de temps de fixation permettent plus cette capacité. La troisième caractéristique est un autre type de tâche visuelle : la réponse à des événements visuels inattendus. La détection d'alertes visuelles en vision périphérique permet d'assurer que l'utilisateur puisse toujours être informé en temps voulu de données critiques. Il semble ici aussi que les modalités à base de temps de fixation soient plus adéquates et limitent moins le champ de vision.

Malgré une préférence des utilisateurs novices pour la multimodalité, en étudiant le contexte des modalités d'interaction, nous montrons qu'une interaction à base de temps de fixation est sûrement plus adaptée pour une intégration dans des systèmes optroniques. L'étude d'autres caractéristiques du contexte permettra d'affiner ces résultats et d'identifier la modalité d'interaction adéquate à une situation donnée.



# Abstract

## **Integration of eye-gaze based interaction in optronic systems: evaluation of the context's influence.**

New versions of Safran Electronics & Defense optronic products such as infrared binoculars or firearm sights are endowed with more and more functionalities. This leads to a need in the improvement of the user interface of those systems. The integration of eye-gaze based interaction modalities seems interesting because of the speed, the naturalness and the availability of the eye.

The eye-gaze based interaction is already well developed for impaired people but is not a mature technology for healthy people yet. During active eye-based interactions, which are explicit input from the user, a problem named Midas Touch arises. It consists in the incapacity of the system to differentiate scene analysis and voluntary user input. This is because the eye is a sensory organ over all. To overcome this problem, several interaction modalities have been designed: Dwell Time uses for example a minimum gaze dwell duration to trigger input, but the fixation area may also be located near the item to activate (relocated Dwell Time). It is also possible to associate the eye with another input modality such as a press button to indicate the intent from the user (multimodality eye-button).

Each of these modalities has pros and cons and cherry-picking the most suitable to a given situation is not trivial. Moreover, the performances of the interaction modalities seem dependent from external variables, which represents the context of an interaction modality. To integrate eye-based interaction in systems and to choose which modality to use, it is necessary to identify what are the context characteristics and how they affect the modalities.

Our goal is to propose an interaction modality context model, that is to define all the external characteristics affecting the modality performances. From a state of the art of the eye-based interaction, we propose a description of the context following four axes: the user, the task, the system and the environment. Each of these axes decomposed in characteristics whose influence is justified by previous works or theoretical reasonings.

Then we studied three characteristics which appeared to us as critical for the integration in optronic products, and we compared the performances of precited modalities against these characteristics. The first one is the type of menu (linear or circular). Contrary to the mouse, the tested interaction modalities present no significative differences depending on the type of menu they are used on. The second characteristic is linked to the user task. The aim is to evaluate the adequacy of interaction modalities with one's ability to split his/her visual attention, that is, to fixate an area while being visually focused elsewhere. This skill is necessary during target following tasks. The fixation-based modalities seemed more permissive concerning this ability. The third characteristic is about peripheral visual alert detection in order to ensure that the user can be warned at all time. Fixation based modalities seemed to less reduce the visual field than others.

Despite novice user preferences for the use of multimodality, we showed that fixation-based modalities may be more adequate for use in optronic systems. The further study of other characteristics of the context will allow to highlight the modality to use for each situation.

# Table des matières

<b>Chapitre 1 Introduction.....</b>	<b>11</b>
1.1 Contexte .....	11
1.2 Problématique.....	13
1.3 Approche .....	13
1.4 Organisation du manuscrit .....	14
<b>Chapitre 2 Le regard comme modalité d'interaction .....</b>	<b>15</b>
2.1 Anatomie et dynamique de l'œil.....	15
2.2 Eye-tracking.....	18
2.2.1 Classification des méthodes de suivi du regard .....	18
2.2.2 Méthode de la réflexion sur la cornée .....	19
2.2.3 Caractéristiques et limites des systèmes de suivi du regard. ....	21
2.3 Cognition liée à la perception visuelle .....	22
2.3.1 Méthodes de perception visuelle.....	22
2.3.2 Les différentes zones du champ visuel.....	23
2.3.3 L'attention visuelle .....	23
2.4 Interaction au regard .....	25
2.4.1 Restriction du domaine d'étude : le déclenchement de commande au regard. ....	25
2.4.2 Problématiques du déclenchement de commande au regard. ....	28
2.5 Modalités d'interaction au regard .....	29
2.5.1 Interaction gestuelle au regard.....	29
2.5.2 Interactions basées sur un temps de fixation ( <i>DwellTime</i> ).....	32
2.5.3 Interaction multimodale .....	34
2.6 Conclusion, approche et méthode.....	36
<b>Chapitre 3 Modélisation du contexte des modalités d'interaction au regard.....</b>	<b>37</b>
3.1 Modèle du contexte d'une modalité d'interaction au regard .....	39
3.2 L'utilisateur .....	43
3.2.1 Expertise de l'utilisateur.....	44
3.2.2 Etat de l'utilisateur .....	47
3.2.3 Caractéristiques individuelles .....	51
3.3 Tâche utilisateur.....	53

3.3.1 Tâches visuelles.....	54
3.3.2 Autres types de tâches.....	56
<b>3.4 La plateforme.....</b>	<b>58</b>
3.4.1 Système de tracking.....	58
3.4.2 Interface graphique.....	59
3.4.3 Autres modalités du système.....	61
<b>3.5 L'environnement.....</b>	<b>62</b>
3.5.1 Luminosité.....	62
3.5.2 Contexte social et auditif.....	63
<b>3.6 Conclusion.....</b>	<b>63</b>
<b>Chapitre 4 Evaluation de l'influence du contexte sur les modalités d'interaction.....</b>	<b>65</b>
4.1 Intégration de l'interaction au regard sur des menus linéaires et circulaires.....	65
4.1.1 Besoin opérationnel et état de l'art.....	66
4.1.2 Méthode et design expérimental.....	67
4.1.3 Résultats.....	75
4.1.4 Discussion.....	79
4.2 Effet de la réorientation de l'attention visuelle sur l'interaction au regard.....	81
4.2.1 Besoin opérationnel.....	81
4.2.2 Méthode et design expérimental.....	81
4.2.3 Résultats.....	88
4.2.4 Discussion.....	90
4.3 Effet de l'interaction au regard sur la capacité à détecter des signaux en vision périphérique.....	93
4.3.1 Besoin opérationnel et état de l'art.....	93
4.3.2 Méthode et design expérimental.....	94
4.3.3 Résultats.....	99
4.3.4 Discussion.....	101
<b>4.4 Conclusion.....</b>	<b>102</b>
<b>Chapitre 5 Conclusion.....</b>	<b>103</b>
5.5 Contributions.....	104
5.5.1 Proposition d'un modèle du contexte des modalités de l'interaction au regard.....	104
5.5.2 Expérimentations sur des critères pertinents pour les systèmes optroniques.....	105

5.5.3 Contribution industrielle .....	105
5.6 Perspectives .....	106
<b>Glossaire .....</b>	<b>109</b>
<b>Annexes.....</b>	<b>111</b>
Annexe A : Questionnaire d'utilisabilité de l'expérimentation 1 concernant l'impact du type de menu.....	113
Annexe B : Menus utilisés dans la première expérimentation.....	116
Annexe C : Adaptation du questionnaire UMUX à l'évaluation d'une modalité d'interaction au regard.....	119
<b>Bibliographie .....</b>	<b>125</b>

# Chapitre 1

## Introduction

### 1.1 Contexte

Nos travaux s'inscrivent dans le cadre d'un projet mené et financé par Safran en collaboration avec le laboratoire UMR CNRS Heudiasyc, de l'Université de Technologie de Compiègne, et plus précisément avec l'équipe ICI (Information, Connaissance, Interaction) qui s'intéresse à l'interaction homme-machine. Ce projet concerne la conception, la réalisation et la validation d'un nouvel équipement en optronique pour Safran, et devra à terme être utilisé dans cette industrie.

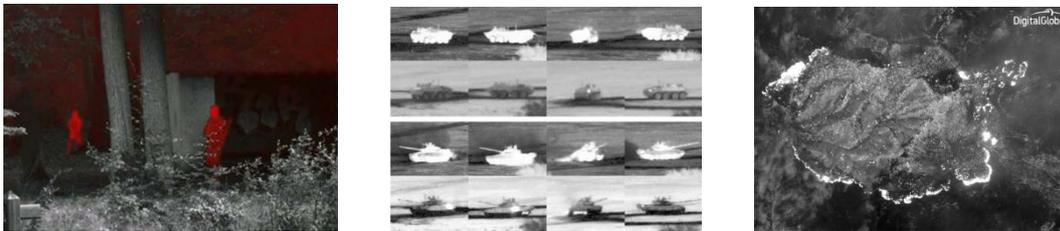
Ces produits sont destinés à être utilisés par divers types d'utilisateurs, militaires (chefs de section, soutien aérien tactique, forces spéciales) ou paramilitaires (pompiers, garde-frontières...), dans des situations difficiles où une erreur peut avoir des conséquences importantes. A ce titre, il est crucial d'avoir une interface homme-machine à la fois fiable et efficace.

L'optronique est une discipline combinant à la fois l'optique et l'électronique. Un système d'optronique associe un ou plusieurs capteurs optiques, une chaîne de traitement d'image et un écran pour présenter une image plus ou moins modifiée à l'utilisateur. L'intérêt de tels systèmes est de pouvoir proposer à l'utilisateur une réalité augmentée, c'est-à-dire de proposer un moyen de visualiser des informations qui ne sont pas directement visibles par l'utilisateur en temps normal. Utilisant des caméras thermiques ou des télémètres lasers, Safran produit des systèmes optroniques complexes comme des jumelles infrarouge multifonctions (comme JIM Compact, cf. Figure 1) ou des lunettes de visée.



*Figure 1. Jumelle infrarouge multifonctions JIM Compact.*

Ces systèmes permettent d'effectuer des actions spécifiques (cf. Figure 2) comme du décamouflage (mise en évidence d'unités cachées), de la reconnaissance automatique de véhicules en utilisant leur signature thermique voir même de suivre l'évolution d'un incendie à travers une épaisse fumée ou détecter un homme à la mer.



*Figure 2. Exemple de tâches réalisables à l'aide d'une jumelle infrarouge multifonction Safran. De gauche à droite : décamouflage d'un fantassin dans un sous-bois, reconnaissance de char par signature thermique, suivi de l'avancée d'un incendie.*

Si les premières générations de ces produits disposaient de fonctionnalités limitées au fait de zoomer/dézoomer et basculer de la vision naturelle (voie jour) à la vision infrarouge, de nos jours, les fonctionnalités sont bien plus nombreuses et vont continuer à croître. Prise de photo, tir télémétrique pour estimer la distance à un point, ajout d'une voie en « bas niveaux de lumières » plusieurs modes de fusion des voies, super-résolution, stabilisation d'image, cartographie, affichage de flux vidéo extérieurs, communication, mise en place de points d'intérêt partagés par plusieurs systèmes, suivi automatique de cibles, etc., ajoutent au potentiel d'utilisation de ces systèmes mais complexifient par la même occasion l'utilisabilité de ces systèmes. En effet, dans les premières versions, un bouton était dédié à chaque fonctionnalité, mais le nombre de doigts de la main n'a pas augmenté aussi vite que le nombre de fonctionnalités. Des menus plus ou moins complexes sont alors apparus afin de sélectionner les différentes fonctions des produits, mais des études des comportements des utilisateurs ont montré que certaines fonctionnalités ne sont pas utilisées, non pas parce qu'elles ne sont pas utiles, mais parce qu'elles sont difficiles d'accès. De plus, certaines tâches comme la **caractérisation de points**, consistant à donner des informations sur la nature et le nombre d'unités observées sont trop longues avec des boutons. Pour ces raisons, Safran s'intéresse à d'autres modalités d'interaction dans le but d'améliorer l'utilisabilité de leurs produits.

Plusieurs points laissent transparaître que le regard serait une modalité d'interaction intéressante. Premièrement, il existe un lien étroit entre les mouvements naturels du regard et la tâche cognitive

consciente de l'utilisateur (Liebling & Dumais, 2014). Qui plus est, les mouvements du regard précèdent l'action. Deuxièmement, le mouvement des yeux est le mouvement le plus rapide de tout le corps humain, ce qui laisse penser à une interaction rapide. Enfin, particulièrement dans ces produits, l'œil est disponible et déjà partie intégrante de l'utilisation des systèmes. Cependant, les premières tentatives d'utilisation du regard comme modalité d'interaction ont mis en avant le problème du *Midas<sup>1</sup> Touch* (Jacob, 1993) bien connu de la littérature. Le roi Midas se vit donner la faculté de transformer tout ce qu'il touchait en or. Une fois l'euphorie des premiers moments passé, il regretta ce don qui l'empêchait alors de boire, manger ou même toucher sa propre fille puisqu'il transformait tout en or, il supplia les dieux de rompre cette malédiction. Une interaction triviale au regard ou tout ce qui est fixé est activé induit le même phénomène, en activant tous les objets regardés, l'utilisateur a dans un premier temps le sentiment d'une communication rapide et naturelle avec le système. Mais rapidement, il devient impossible d'analyser l'interface sans en activer tous les éléments. L'œil étant par nature un organe senseur, il est en effet délicat de différencier les moments où il est utilisé de manière purement sensorielle pour analyser l'environnement des moments où il y a une intention d'interaction de la part de l'utilisateur. Pour résoudre ce problème, différentes modalités d'interaction utilisant le regard ont été développées durant les dernières décennies pour mieux permettre de discerner l'intention d'interaction de l'utilisateur. Ces modalités possèdent des qualités et des défauts qui font qu'elles disposent de situations d'utilisation différentes.

## 1.2 Problématique

Si les modalités d'interaction au regard ne sont pas toutes égales, il est nécessaire de pouvoir les évaluer. Certains travaux de la littérature actuelle s'attèlent à ce travail, mais il semble que le contexte dans lesquelles les modalités sont utilisées ait une influence importante sur leur utilisabilité. Nous cherchons alors à comprendre **quels sont les caractéristiques du contexte des modalités d'interaction au regard et comment elles influencent le comportement des utilisateurs vis-à-vis de ces modalités.**

Les situations dans lesquelles sont utilisés les produits Safran sont riches et variées. Elles donnent lieu à des cahiers des charges complexes. L'intérêt industriel d'un tel travail est alors clair : comprendre les modalités d'interaction au regard et choisir la/les plus pertinente(s) en fonction des cas d'utilisation afin de pouvoir les intégrer. L'intégration se faisant souvent sur des produits déjà conçus, il sera nécessaire de s'adapter aux interfaces existantes, ce qui limitera nos recherches à l'étude des modalités d'interaction au regard permettant d'interagir avec des menus.

## 1.3 Approche

L'intérêt scientifique de répondre à une telle question est de considérer le contexte comme un élément décisionnel important de la conception et du choix des modalités à utiliser. Adapter les interfaces aux besoins précis des utilisateurs est un axe de recherche complet, avec les interfaces plastiques par exemple (Calvary, Coutaz, & Thevenin, 2001), et la compréhension des impacts de facteurs extérieurs au système (du contexte) fait partie de ce but. Le contexte est à la fois un élément

---

<sup>1</sup> <https://fr.wikipedia.org/wiki/Midas>

perturbateur des modalités d'interaction, mais représente aussi un élément d'évaluation de ces modalités.

Nous avons opté pour une modélisation de ce contexte. Cette modélisation a été réalisée à la fois de manière top-down en s'inspirant des modèles de la littérature et de manière *bottom-up* en tirant parti de l'expertise métier des ergonomes de Safran Electronics & Defense et des observations réalisées lors des premières phases d'exploration de la technologie de suivi du regard. Nous avons cherché les éléments pouvant avoir un lien avec les mouvements de l'œil ou les aspects cognitifs de l'utilisateur.

Des expérimentations ont été réalisées sur les caractéristiques du modèle du contexte qui nous ont paru les plus pertinentes pour une intégration prochaine dans les produits. Ces expérimentations n'ont donc pas pour but de valider le modèle, mais de déterminer les modalités d'interaction au regard les plus adaptées à une situation opérationnelle précise. Nous nous sommes intéressés en particulier au cas de la caractérisation de points durant une phase d'observation. La caractérisation de points est une tâche demandant à l'utilisateur de spécifier un certain nombre de caractéristiques d'unités observées (taille, spécificités, allégeance, activité). Cette tâche s'effectue de manière méthodique, en spécifiant une caractéristique à la fois parmi des caractéristiques préétablies. L'utilisation de menus présentant les choix possibles pour les différentes caractéristiques est adaptée à cette tâche.

Les caractéristiques étudiées lors des expérimentations ont été déterminées en fonction de cette tâche. Nous avons choisi d'étudier la différence entre menus linéaires et circulaires ainsi que la capacité de l'utilisateur à continuer à utiliser les capacités perceptives de l'œil durant une interaction au regard.

## 1.4 Organisation du manuscrit

Le Chapitre 2 est consacré à un état de l'art sur le fonctionnement de l'œil et l'interaction au regard. Le fonctionnement physiologique et cognitif de l'œil ainsi que les systèmes de suivi du regard sont présentés. Les problématiques liées à l'utilisation motrice du regard sont abordées et nous présentons des différentes modalités d'interaction au regard de la littérature.

Le chapitre 3 est une modélisation du contexte des modalités d'interaction au regard. Il propose une décomposition du contexte selon 4 axes : Utilisateur, Tâche, Plateforme et Environnement. Chacun de ces axes est étudié afin d'identifier les facteurs d'impacts des modalités d'interaction.

Le chapitre 4 présente trois expérimentations qui ont été réalisées au cours de la thèse. La première consiste à comparer différentes modalités d'interaction au regard, dont une que nous avons conçue, sur des menus linéaires et circulaires disposant de caractéristiques proches. La seconde expérience s'intéresse à notre capacité à déporter notre attention visuelle tout en interagissant au regard et la troisième évalue notre capacité à détecter des stimuli en vision périphérique lors de l'interaction.

Le chapitre 5 conclut le mémoire et présente des perspectives à ces travaux de recherche.

## Chapitre 2

# Le regard comme modalité d'interaction

Afin de parler d'interaction utilisant le regard comme modalité d'interaction, il est nécessaire d'expliquer le fonctionnement basique de l'œil humain et de détailler ses caractéristiques propres, ses avantages et ses inconvénients. Nous décrivons ensuite le fonctionnement de la technologie actuellement la plus utilisée de tracking du regard qui est basé sur la capture vidéo de l'œil. Nous aborderons certains mécanismes cognitifs liés à l'utilisation du regard. Enfin, nous aborderons les problématiques de l'interaction au regard et nous détaillerons les principales méthodes d'interaction au regard.

### 2.1 Anatomie et dynamique de l'œil

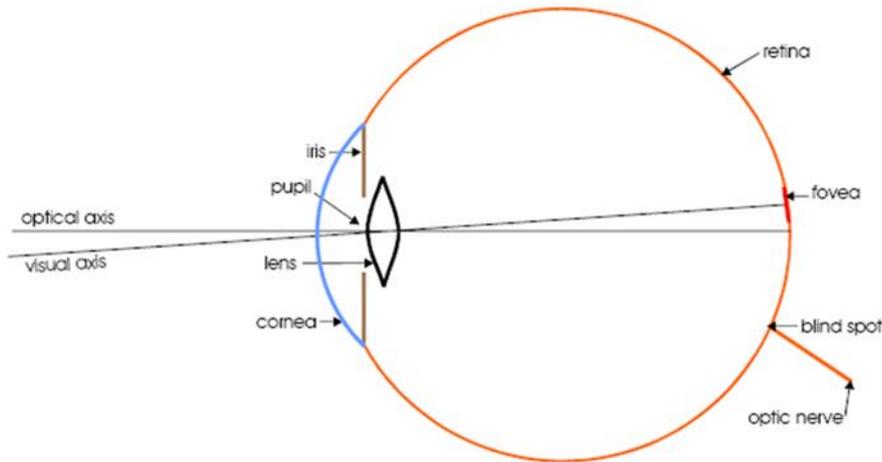
*“Evolution has over time turned the eye into an organ that carefully detects and filters a small specter of electromagnetic energy into meaningful information that can be interpreted by the brain.”*

(Skovsgaard, 2012)

L'œil est un système optique (cf. Figure 3) permettant de transformer des signaux électromagnétiques extérieurs en pulsation électriques dirigées vers le cerveau.

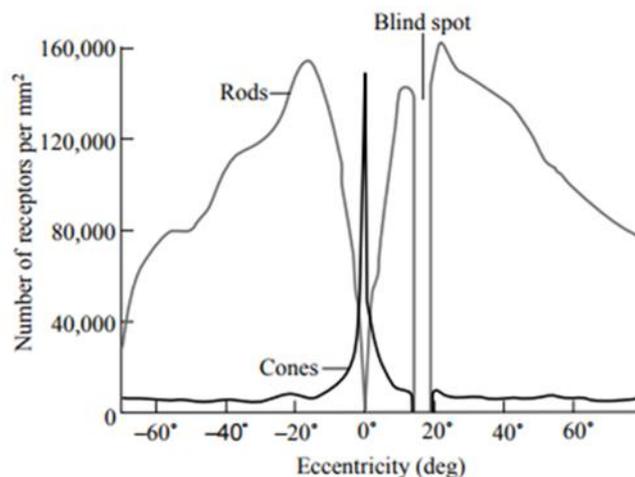
La première partie de ce système optique est la cornée, sur laquelle la lumière arrive et est réfractée. La lumière passe ensuite à travers le système pupille/iris qui agit comme un diaphragme permettant de réguler la quantité de lumière entrante. Derrière se situe la lentille, qui va focaliser les rayons lumineux sur la rétine qui agit comme un capteur sensible à la lumière. La rétine est composée de centaines de milliers de cellules lumino-réceptives : les cônes et les bâtonnets. Leur rôle est de transformer la puissance lumineuse en signaux électriques qui sont ensuite transmis au cerveau par le biais du nerf optique. Les bâtonnets sont sensibles à la luminosité et permettent entre autres de voir dans l'obscurité. Les cônes sont moins sensibles au contraste mais plus sensibles aux couleurs. Les cônes sont de trois types différents dont chaque type est sensible à différents intervalles de longueur d'ondes, l'association d'information permet alors de détecter des couleurs. Il existe

néanmoins une zone aveugle (blind spot) correspondant à la position de l'axe optique sur laquelle aucune cellule réceptrice n'est présente.



*Figure 3. Schéma représentatif du système optique visuel (Drewes, 2010).*

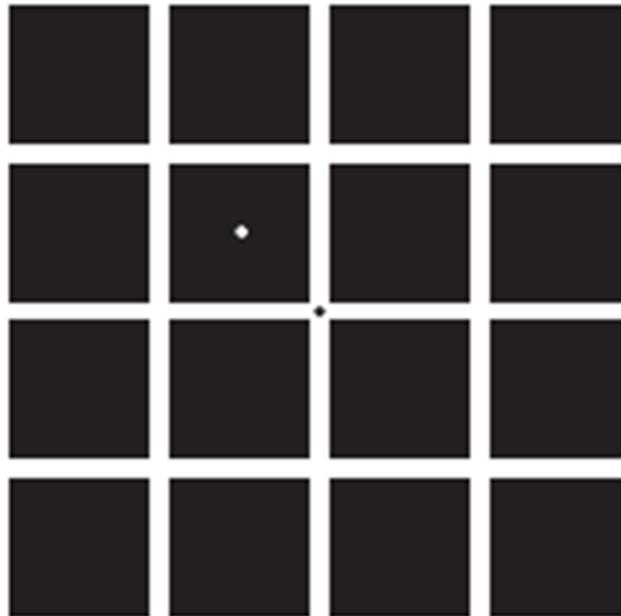
La taille du cône elliptique de perception des ondes lumineuses est de 180° horizontalement sur 130° verticalement (Drewes, 2010), mais la répartition de ces cellules sur la rétine n'est pas homogène (cf. Figure 4). Cette disparité crée des zones d'acuité différente dans l'œil. La densité de cônes est beaucoup plus grande dans une zone d'environ 5° appelée macula, zone de la vision dite centrale. Au centre de la macula existe une zone plus petite d'environ 1 à 2° qui correspond à la zone visuelle d'acuité maximum, c'est la **fovea**. En dehors de cette zone, l'homme dispose d'une vision à faible résolution dont l'utilité est principalement la détection de mouvements ambiants (Duchowski, 2007). Cette haute résolution localisée implique d'effectuer des mouvements rotatifs du globe oculaire pour aligner la fovea dans la direction souhaitée. Ainsi, lorsque l'on parle d'orientation de l'œil, on raisonne par rapport au centre de la fovea (axe visuel) et non pas par rapport au centre de l'axe optique de l'œil.



*Figure 4. Répartition des bâtonnets et des cônes sur la rétine en fonction de l'angle par rapport au centre de la fovea (Rodieck & Rodieck, 1998)*

Les mouvements oculaires sont des mouvements de rotation horizontaux, verticaux ou de torsion, synchrones pour les deux yeux, qui permettent une orientation en 3D de l'œil par rapport à la tête. On peut classer les mouvements de l'œil en plusieurs catégories :

- Le premier « mouvement » de l'œil n'est en réalité pas un mouvement à proprement parler, c'est l'état dans lequel l'œil est stable comme par exemple lorsque le regard s'arrête sur un mot durant la lecture. On parle alors de **fixation**. Une fixation dure classiquement entre 200 et 300ms mais peut être étendue sur une durée de plusieurs secondes. Le point de fixation fait ainsi référence à l'orientation désirée de l'axe visuel.
- Pour se déplacer d'un point de fixation à un autre, l'œil effectue un mouvement de **saccade**. C'est un mouvement balistique (qui ne peut pas être interrompu une fois déclenché) rapide (entre 30 et 80ms en fonction de l'ampleur de la saccade) durant lequel aucune information visuelle n'est envoyée au cerveau. Dans certaines conditions, la vitesse de rotation peut dépasser les 700°/s, faisant de l'œil l'organe le plus rapide du corps humain (Carpenter, 1988). Une saccade ne dispose pas d'une précision parfaite, ainsi, à la fin du mouvement de saccade, l'œil oscille avant de se fixer. Ce mouvement est appelé **glissade** (Holmqvist, et al., 2011).
- Lors d'une fixation, l'œil n'est en réalité pas tout à fait fixe et effectue des mouvements de faible amplitude non contrôlables et imperceptibles, ces mouvements sont les tremblements, la dérive et les microsaccades. Les tremblements, d'une fréquence d'environ 90 Hz, sont sûrement dus à la limite du contrôle des muscles impliqués dans le maintien de l'orientation de l'œil. La dérive est un mouvement lent (6-25' /s) éloignant progressivement l'axe visuel du point de fixation. Les microsaccades corrigent fréquemment cette dérive afin de réorienter l'œil vers le point de fixation. La Figure 5 permet de prendre conscience de ces micromouvements.



*Figure 5. Micromouvements de l'œil durant une fixation : si on fixe le point noir pendant environ 30 secondes et si on fixe le point blanc ensuite, l'image persistante des lignes blanches dans l'œil est perçue en mouvement constant par rapport à l'image réelle (Skovsgaard, 2012).*

- Le dernier mouvement synchrone est la poursuite lissée, qui est réalisable uniquement en suivant un objet en mouvement. Dans ce cas, le mouvement d'orientation ne se fait pas de manière saccadée mais de manière continue afin de conserver l'orientation du regard sur l'objet en mouvement.
- Il existe enfin le mouvement de vergence qui n'est pas un mouvement synchrone des deux yeux. Chaque œil effectue un mouvement différent afin d'orienter les deux axes visuels vers un seul point de l'espace. C'est ce qui permet une vision 3D nette.

D'un point de vue sensoriel, le phénomène d'adaptation neuronale fait que si le même signal est reçu en permanence, il finit par être ignoré. Ainsi, si tous les mouvements de l'œil étaient compensés, et que la même image était présentée exactement de la même manière sur l'œil, aucun signal visuel ne serait perçu par le cerveau. Ce sont donc les micromouvements qui, lors d'une fixation, permettent de continuer à voir.

## 2.2 Eye-tracking

### 2.2.1 Classification des méthodes de suivi du regard

Enregistrer les mouvements de l'œil n'est pas une idée nouvelle. Les mouvements de l'œil sont étudiés depuis plus d'un siècle (Drewes, 2010) comme le montre l'apparition du terme « saccade » en 1892 par Lamare. Le premier mécanisme d'enregistrement des mouvements de l'œil apparaît en 1891 avec un système proposé par Ahrens utilisant de petits leviers fixés sur le globe oculaire permettant de transférer les mouvements sur une surface couverte de suie. Par la suite, plusieurs méthodes d'enregistrement des chemins oculaires ont été inventées. Elles peuvent être classées en trois catégories : celles basées sur l'électro-oculographie, celles basées sur des lentilles de contact et celles basées sur la vidéo.

- La première méthode consiste à disposer des électrodes autour de l'œil qui vont mesurer des différences de potentiel lors de changement de la position de l'œil (Duchowski, 2007).
- La seconde méthode consiste à utiliser une lentille de contact munie d'une micro bobine alimentée à l'aide de fils électriques (cf. Figure 6). Les perturbations induites dans un champ électromagnétique contrôlé permettent ensuite de mesurer la position de la bobine et donc l'orientation de l'œil. Si cette méthode est la plus précise (Morimoto & Mimica, 2005), elle souffre d'une forte intrusivité et n'est donc pas utilisable pour de l'interaction au regard.



**Figure 6. Lentille de contact avec bobine (gauche<sup>2</sup>) et électrodes positionnées autour de l'œil (droite<sup>3</sup>) permettant de suivre l'orientation du regard.**

- Enfin, la troisième méthode consiste à filmer l'œil à l'aide de caméras et à extraire les informations de mouvement de l'œil de la vidéo. Cette méthode est souvent utilisée en addition avec un système d'émission infrarouge dirigé vers l'œil. Cette troisième méthode est la plus répandue en dehors des études très précises sur les mouvements de l'œil. De nombreux systèmes prêts à l'utilisation basés sur cette méthode sont accessibles dans le commerce. Elles sont commercialisées par des sociétés comme Tobii, EyeTribe, SMI, Smart Eye ou Mirametrix.

Les systèmes de suivi du regard peuvent aussi être classés en deux catégories, les systèmes montés et les systèmes déportés (Holmqvist, et al., 2011). La différence entre les deux types de système est le repère spatial de référence. Dans le cas d'un système monté, le repère est lié à la tête de l'utilisateur, les mouvements mesurés sont décrits par rapport à la tête considérée fixe. Il est ensuite possible d'ajouter un système de tracking de la tête pour déduire l'orientation de l'œil dans l'espace. Le second type de système est déporté par rapport à l'utilisateur et mesure alors la direction du regard sans avoir besoin de prendre en compte l'orientation de la tête. Ce dernier type de système a l'avantage de ne pas avoir à traiter les mouvements « parasites » de la tête afin de déterminer l'orientation de l'axe visuel. Pour l'interaction au regard, les systèmes déportés, c'est-à-dire liés directement au système d'affichage, sont donc plus adéquats.

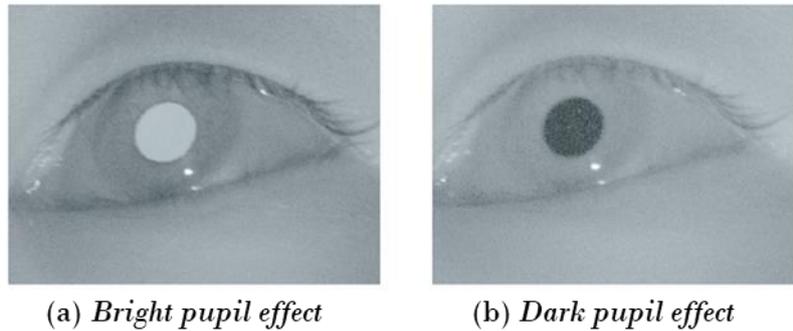
### 2.2.2 Méthode de la réflexion sur la cornée

La méthode de suivi du regard actuellement la plus utilisée est basée sur la détection de deux points d'intérêt, le centre de la pupille et la position d'un (ou plusieurs) reflet(s) sur la cornée de l'œil.

Deux méthodes existent pour détecter la pupille appelée méthode de la pupille noire et méthode de la pupille brillante. La première consiste simplement en l'acquisition vidéo de l'œil en s'assurant qu'aucune source lumineuse ne soit placée à proximité de l'axe optique de la caméra. La pupille apparaît alors noire et facilite sa détection (cf. Figure 7b). La seconde méthode consiste à éclairer l'œil avec une source lumineuse placée près de l'axe optique de la caméra. Dans ce cas, la rétine réfléchit la lumière infrarouge qui est alors captable par une caméra infrarouge. La pupille apparaît alors blanche et contraste avec le reste de l'œil (cf. Figure 7a). C'est le même effet qui cause les yeux rouges lors d'une photographie avec un flash.

<sup>2</sup> Tiré de chronos-vision.de

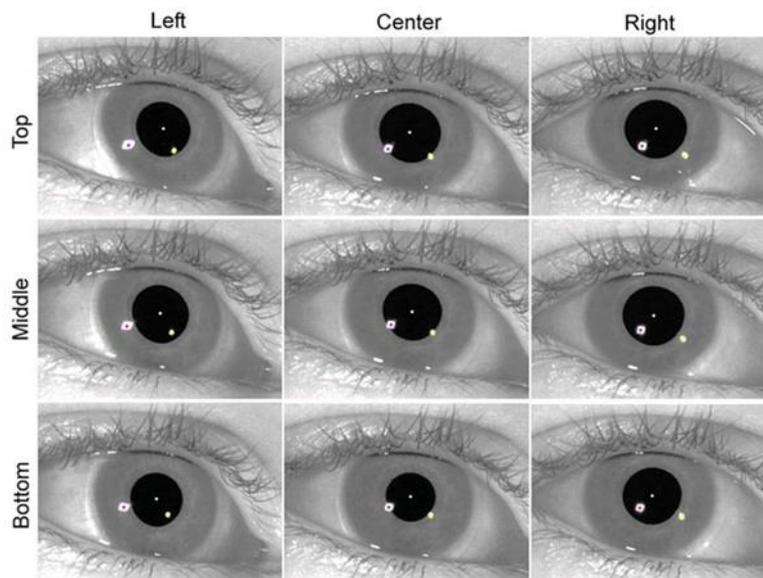
<sup>3</sup> Tiré de metrovision.fr



**Figure 7. Deux méthodes de détection de la pupille. Méthode de la pupille brillante (a) et méthode de la pupille noire (b). Image issue de (Morimoto, Koons, Amit, Flickner, & Zhai, 1999)**

Illuminer l'œil à l'aide d'un émetteur infrarouge entraîne quatre réflexions respectivement sur la surface de la cornée, sur le fond de la cornée, sur la surface de la lentille et sur le fond de la lentille, appelées images de Purkinje. Le plus souvent, une sélection appropriée de l'intensité de l'émetteur fait que seule la première réflexion apparait distinctement. Le centre de cette tâche réfléchie constitue le second point d'intérêt.

La surface de la cornée étant approximativement sphérique, les mouvements de rotation de l'œil n'influent pas sur la position de la tâche de Purkinje (Holmqvist, et al., 2011) alors que la position du centre de la pupille est dépendante de l'orientation de l'œil comme le montre la Figure 8. Il existe alors une bijection entre le vecteur séparant ces deux points et la direction du regard.



**Figure 8. Position du centre de la pupille et des images de Purkinje issues de deux sources lumineuses en fonction de l'orientation de l'œil.**

Afin de déterminer la position du regard sur un écran, aucune reproduction 3D de l'orientation du regard n'est effectuée. C'est en fait une phase de calibration qui permet de déterminer la valeur du vecteur <centre de la pupille, centre de l'image de Purkinje> pour des positions clés sur l'écran. Pour cela, l'utilisateur doit fixer ces points clés afin que le système mesure le vecteur pour ces positions. La Figure 8 présente 9 orientations sur 9 points différents de l'écran. La position réelle du regard est

ensuite déduite par interpolation inverse. Comme la physiologie des yeux de chacun est différente une calibration individuelle est nécessaire.

Cette méthode ne pose pas de problème pour les utilisateurs disposant de lunettes, l'image de Purkinje est décalée mais présente les mêmes propriétés. Cependant, elle peut poser problème pour les personnes ayant une déformation du globe oculaire (i.e. astigmatisme) ou utilisant des lentilles de contact.

### 2.2.3 Caractéristiques et limites des systèmes de suivi du regard.

Utiliser un système de suivi du regard pour l'interaction nécessite de prendre en compte différents facteurs influant sur la qualité du signal généré.

Tout d'abord, l'utilisation d'un tel système nécessite un environnement lumineux contrôlé afin de ne pas perturber la récupération des points d'intérêts. Une forte luminosité, certains spots lumineux, etc., peuvent empêcher ou dégrader la récupération de la position du regard.

Le système lui-même présente des limites. La résolution du capteur a un impact direct sur la précision estimée du centre de la pupille et de la tâche de Purkinje, une meilleure résolution donnant une estimation de la position du regard plus précise. La qualité des algorithmes de traitement d'image influe de la même manière sur les performances obtenues.

En anglais, deux termes font référence à la précision d'un tel système : *precision* et *accuracy* (Holmqvist, et al., 2011). La différence est résumée dans la Figure 9.

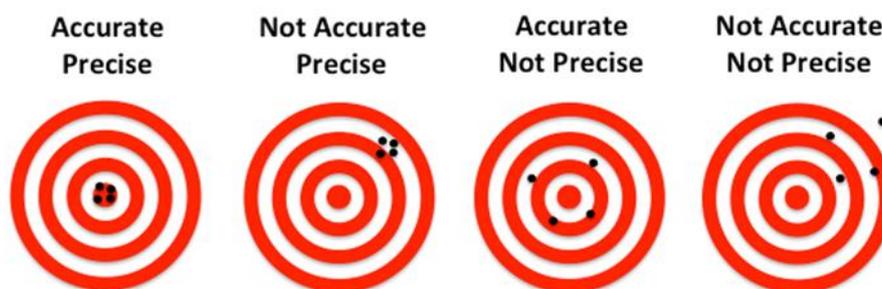


Figure 9. Illustration de la différence entre la précision et l'exactitude.

La *precision* correspond à la variabilité de la position des échantillons autour de la position réelle du regard. Cette variabilité a plusieurs sources : les micromouvements de l'œil détaillés dans la section 2.1, le bruit inhérent du système comme la résolution temporelle ou spatiale du capteur, les bruits environnementaux. Afin de corriger cette variable, la plupart des systèmes intègrent un filtrage passe-bas (moyenneur) des données brutes. Un filtrage induit cependant une latence et nécessite une plus haute fréquence d'acquisition de la part des systèmes. Le système Tobii EyeX utilisé pour nos expérimentations présente une fréquence d'acquisition de 60Hz. Une utilisation temps réel oblige à un filtrage causal, c'est-à-dire ne dépendant que des échantillons précédents, ce qui n'est pas le cas pour un traitement *a posteriori* des données d'un système de suivi du regard.

L'*accuracy*, aussi traduit par la précision en français, représente le décalage entre le point déterminé par le système de suivi du regard (après filtrage et traitements) et la position réelle du regard. Ce type d'erreur peut être dû à une mauvaise calibration, à une erreur d'approximation de la position due à l'interpolation à partir des positions clés. Lors de l'utilisation de systèmes déportés, la modification de la position de la tête par rapport au moment de la calibration impacte ce type de précision.

## 2.3 Cognition liée à la perception visuelle

Dans cette section, nous étudions plus en profondeur l'utilisation du regard et du champ visuel dans la tâche utilisateur. Nous commençons par définir les différentes manières naturelles d'utiliser le regard. Puis nous détaillons les différences entre la vision centrale et la vision périphérique. Nous verrons ensuite dans quelle mesure un utilisateur parvient à contrôler son regard avant de séparer orientation du regard et attention visuelle. Pour finir, nous évoquons les processus cognitifs liés à la vision.

### 2.3.1 Méthodes de perception visuelle

Nous avons décrit les différents mouvements de l'œil dans un chapitre précédent, mais nous n'avons pas encore parlé de l'utilisation de l'œil par un humain. En effet, l'œil est un système perceptif permettant de capter de l'information dans l'environnement, mais cette récupération d'information peut se faire de différente manière. Wickens (Wickens, Hollands, Banbury, & Parasuraman, 2015) présente six types de tâches utilisant l'attention visuelle. Nous reviendrons précisément sur le terme d'attention plus tard dans ce chapitre. Les six types de tâches sont :

- Le balayage visuel, c'est la forme d'observation primaire chez l'homme. Il permet d'établir une cartographie mentale de son environnement. C'est en particulier la manière dont nous observons une scène, un environnement, une image, pour la première fois. Le but est d'obtenir le plus d'information le plus rapidement.
- La supervision visuelle consiste en une série d'observations courtes afin de s'assurer que certaines variables sont dans des limites prédéfinies. Ce mode d'observation est par exemple utilisé en conduite où le conducteur doit vérifier que son véhicule est toujours orienté dans la direction de la route, que sa vitesse n'excède pas la limite autorisée, etc.
- La réponse à un événement, c'est le mouvement induit par l'apparition, ou tout du moins la détection d'un élément inattendu. Cela peut être un stimulus visuel par exemple mais aussi la réponse à un son.
- La recherche correspond au mouvement effectué lors de la recherche d'un objet ou d'un élément prédéterminé par le sujet. Le sujet cherche une correspondance avec une image mentale de sa cible.
- La lecture, mouvement systématique et linéaire du regard sur un texte qui permet le déchiffrement d'un texte.
- Le contrôle, qui consiste à vérifier que certaines actions attendues ont bien eu lieu, comme la prise en compte d'un feedback.

Il est évidemment possible qu'une tâche combine plusieurs de ces tâches visuelles. La lecture d'un site web fait par exemple appel à un balayage visuel en plus de la lecture (Russell, 2005).

Dans tous les cas, le système perceptif visuel, comme la plupart des systèmes perceptifs humains, est sensible aux changements (Bordwell, Thompson, & Smith, 2016), que ceux-ci soient d'ordres spatiaux ou temporels. C'est-à-dire qu'il est plus sensible aux mouvements, aux changements de couleur ou de luminosité qu'à un signal constant. Mais comme nous l'avons vu dans la section 2.1, le champ de vision humain n'est pas homogène. La répartition des différentes cellules sensibles à la lumière fait que le champ de vision peut être découpé en plusieurs secteurs dont l'utilité est différente : la vision centrale et la vision périphérique. La vision centrale représente une zone d'environ 5° contenant la fovea. Plusieurs études montrent que leur rôle est différent. La vision périphérique permet une appréhension rapide d'une situation visuelle alors que la vision centrale permet une analyse plus précise, mais aussi plus lente.

### **2.3.2 Les différentes zones du champ visuel**

La vision centrale et la vision périphérique influent différemment sur la perception du mouvement. Brandt (Brandt, Dichgans, & Koenig, 1973) montre par exemple une différence de perception si un mouvement observé en vision centrale implique une sensation de mouvement de l'environnement par rapport à l'observateur alors que le même mouvement observé en vision périphérique induit une sensation de déplacement de l'observateur.

De manière naturelle, la vision périphérique est plutôt utilisée de manière passive, sans considération consciente particulière, sauf exception (cf. section 2.3.3). Les formes, les couleurs, la position des objets ne peuvent pas être déterminées avec autant de précision que dans la vision centrale. Cela nous permet cependant de reconnaître des objets ou des formes connues et fréquentes sans avoir besoin de les regarder précisément comme des visages, des maisons, etc., et de détecter des anomalies dans une scène afin de déterminer par la suite, les zones les plus dignes d'intérêt pour la vision centrale. Il en est de même pour la réaction à des mouvements inattendus ou à des changements brusques de luminosité ou de couleur en vision périphérique.

La vision centrale est au contraire liée au processus cognitif conscient de l'utilisateur, c'est-à-dire à la tâche mentale qu'il est en train de réaliser. Des études (Epelboim & Suppes, 2001; Salvucci, 1999) utilisent même le chemin effectué par le regard pour déterminer la structure de la pensée de certains individus. Pelz (Pelz & Canosa, 2001) montre que, dans une tâche complexe, le regard précède l'action. C'est aussi le cas lors de l'utilisation de systèmes informatiques (Liebling & Dumais, 2014) ou le regard précède les mouvements du curseur de la souris par exemple.

### **2.3.3 L'attention visuelle**

En réalité, la zone du champ de vision utilisée par le processus cognitif conscient n'est pas forcément seulement la vision centrale. Pour décorréliser cette zone de la zone centrale de la vision, on parle d'attention visuelle. L'attention de manière générale se rapporte à la manière dont nous traitons activement et consciemment notre environnement. L'attention visuelle peut être imagée par un faisceau lumineux (Wickens, Hollands, Banbury, & Parasuraman, 2015). Dans cette métaphore, ce qui est éclairé est activement traité par notre processus cognitif principal.

La largeur du faisceau lumineux représente la zone dans laquelle notre attention est focalisée ; plus cette zone est large, moins intensément elle est éclairée (Treisman & Schmidt, 1982). De manière générale, si le regard est utilisé de manière active, la vision centrale est représentative de la zone d'attention.

Cependant, il est possible de délocaliser son attention (Awh, Pashler, & others, 2000; Tsotsos, et al., 1995), c'est-à-dire se concentrer visuellement sur une autre zone que celle dans laquelle se dirige le regard, comme si le faisceau lumineux était dévié de sa trajectoire originale. Dubois et ses collègues (Dubois, Hamker, & VanRullen, 2009) montrent cependant que cette délocalisation de l'attention n'est que temporaire.

Dans le cas où il est possible de se concentrer sur deux zones simultanément, on parle d'attention divisée. Notre capacité cognitive est limitée, ainsi nous sommes limités dans notre capacité d'attention (on peut voir ça comme la puissance lumineuse à notre disposition), ce qui explique que nous devons faire un choix sur l'orientation et la quantité d'attention que nous allouons à une zone. De plus, cette capacité d'attention est liée à notre charge mentale, c'est-à-dire à la « quantité de réflexion » à un instant donné. Plusieurs études montrent qu'une augmentation de la charge mentale réduit nos capacités d'attention (Kahneman, Beatty, & Pollack, 1967; Williams, 1982), on parle alors d'effet tunnel.

La plupart du temps, nous n'avons pas conscience de nos mouvements oculaires, les yeux fonctionnant de manière automatique afin de fournir l'information à notre cerveau. Ainsi, bien que notre attention (*a fortiori* notre attention visuelle) soit directement liée à un phénomène cognitif conscient, son orientation est inconsciente. Cela implique un contrôle seulement partiel de notre attention. Ce contrôle de l'attention peut être initié par les activités cognitives, on parle alors d'orientation endogène ou induit par des phénomènes extérieurs, on parle alors d'orientation exogène de l'attention.

Une orientation endogène de l'attention visuelle peut être explicite (i.e. se concentrer sur une zone et ignorer les autres (Posner, 1980)), ou implicite, c'est-à-dire automatiquement guidée par des régularités spatiotemporelles que nous avons assimilées au cours de notre évolution (Large & Jones, 1999).

L'attention visuelle exogène est captée par ce que l'on appelle des attracteurs visuels. D'après Oshberger, ces attracteurs visuels sont (Oshberger, Maeder, & Bergmann, 1998, September) :

- le contraste : les zones à fort contraste par rapport à l'environnement tendent à attirer l'attention visuelle
- la taille : les zones plus larges attirent généralement plus l'attention que les petites.
- la forme : les zones allongées et fines attirent d'avantage l'attention visuelle.
- la couleur : certaines couleurs comme le rouge sont plus aptes à capter notre attention. De plus, une couleur bien dissociée de la couleur de fond attire aussi l'attention.
- le mouvement : le mouvement est l'un des attracteurs visuels les plus importants. Comme cela a déjà été évoqué, notre vision périphérique joue un rôle important dans la détection de mouvements.
- la luminosité : les zones les plus lumineuses attirent plus l'attention visuelle.

- l'orientation : certaines formes comme les flèches disposent d'une orientation qui guide l'attention visuelle.
- les fins de lignes : de même que pour l'orientation, l'attention visuelle est dirigée vers la direction implicite formée par des lignes.

Les théories actuelles (Corbetta, Patel, & Shulman, 2008) tendent à montrer que les attentions endogène et exogène rivalisent sur de courtes périodes de temps afin de guider le comportement d'un utilisateur. D'un point de vue pratique, cela rend les mouvements du regard complexes et imprédictibles.

## 2.4 Interaction au regard

### 2.4.1 Restriction du domaine d'étude : le déclenchement de commande au regard.

Les systèmes de suivi du regard ont d'abord été utilisés comme outils d'analyse du chemin oculaire lors d'une tâche. Comme cette analyse est réalisée *a posteriori* et n'a pas d'impact sur l'utilisateur, on ne parle alors pas d'interaction au regard. Plusieurs domaines utilisent la récupération des données du regard de cette manière. Une cartographie des positions successives du regard permet de faire de l'aide à la conception de publicité (Wedel & Pieters, 2008) ou à la conception de sites web en optimisant le positionnement de contenus informatifs par exemple. L'analyse du chemin oculaire peut aussi permettre la détection de pathologies comme certaines formes d'autismes en détectant une position du regard plus éloignée du centre des visages ou encore de mieux comprendre différents processus cognitifs comme la comparaison de formes (Just & Carpenter, 1976), l'arithmétique (Salvucci, 1999) ou de tâches plus complexes comme la conduite (Robinski & Stein, 2013).

Nous utilisons aussi le regard pour communiquer avec notre entourage. Le seul fait de regarder une personne lorsque l'on est en train de parler lui indique qu'elle est la personne à qui on parle. Nous pouvons aussi effectuer des signes avec les yeux tels que « regarde par-là » ou « qu'ai-je fait au seigneur pour mériter ça ? ». Ainsi, il n'y a qu'un pas pour que les systèmes de suivi du regard soient aussi utilisés pour interagir. Motivés au départ par des besoins pour les personnes en situations de handicap, les travaux sur l'interaction au regard prennent aujourd'hui plusieurs formes. L'interaction est aussi séparée en deux catégories : l'interaction active et l'interaction passive.

*Interaction passive* : La principale caractéristique d'une interaction passive est qu'elle n'est pas instanciée par l'utilisateur. Il n'est pas décisionnaire quant au déclenchement de la boucle d'interaction, c'est le système qui l'est. Le système collecte des informations sur l'utilisateur sans intention de la part de ce dernier et utilise ces informations pour faire évoluer son état. L'interaction peut alors être décomposée en 3 étapes : la collecte des données, l'interprétation (optionnelle) puis la modification du système.

Hillaire et ses collègues (Hillaire, Lécuyer, Cozot, & Casiez, 2008, March) proposent une génération automatique du flou, normalement issu de mouvements ou issu de la distance de focalisation, en fonction de la position du regard de l'utilisateur. Ici, la position du regard de l'utilisateur est utilisée

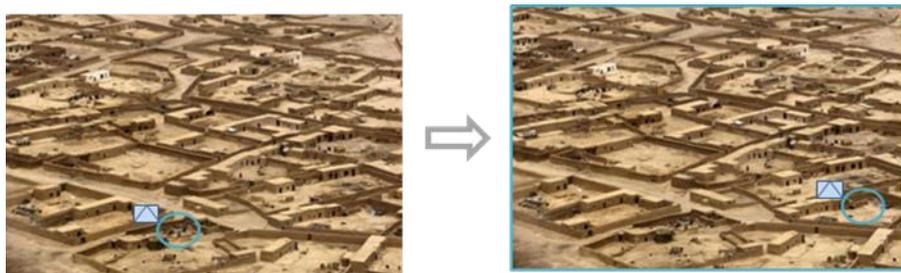
brute afin d'appliquer un traitement de rendu. Il est aussi possible de transformer les mouvements du regard pour générer une information de plus haut niveau. Dans iDict (Hyrskykari, 2006), Hyrskykari propose une aide automatique à la lecture en langue étrangère. Les données du regard sont interprétées afin de signifier une difficulté de compréhension. L'interface propose alors automatiquement une traduction ou des détails sur une formulation. Dans ce cas aussi, l'instigateur n'est pas l'instituteur de l'interaction. On parle aussi d'interface attentive.

*Interaction active* : Contrairement à l'interaction passive, une interaction active est initiée intentionnellement par un utilisateur dont le but est la modification du système. Norman (Norman D. A., 1986) décompose l'activité d'un utilisateur (c.a.d. l'interaction active) en 7 étapes :

- Détermination d'un but
- Génération de l'intention d'action
- Détermination des actions à réaliser (traduction de l'intention abstraite en séquence d'actions concrètes)
- Exécution
- Perception de la modification du système
- Interprétation
- Evaluation de l'état du système par rapport aux buts initiaux

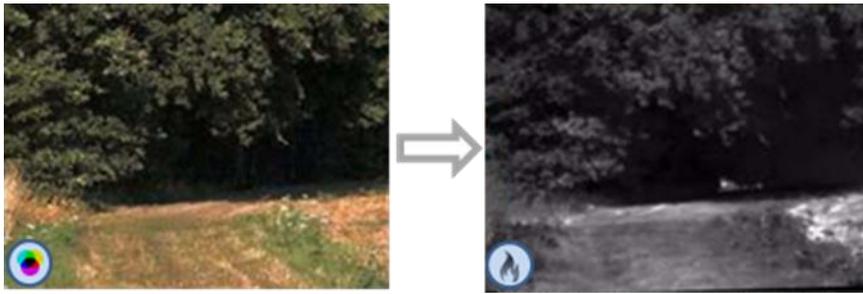
Cette décomposition met en avant l'intention et l'action de l'utilisateur. L'interaction active est séparée en deux catégories : la manipulation directe et le déclenchement de commandes.

La manipulation directe consiste à agir directement avec l'objet d'intérêt afin de modifier dynamiquement ses propriétés physiques comme sa position, sa taille, sa forme. Certaines actions plus abstraites peuvent découler de cette modification physique. Par exemple déplacer un fichier sur la corbeille du bureau d'un ordinateur déclenche la suppression de ce fichier. Lors d'une manipulation, l'évolution du système est faite de manière continue au fur et à mesure de l'action (e.g. Figure 10).



**Figure 10. Exemple de manipulation directe : déplacement d'un point d'intérêt sur une photographie.**

Le déclenchement de commande consiste à activer en une fois une séquence d'actions, pouvant être plus abstraites que lors d'une manipulation. Contrairement à la manipulation, la commande est instantanée. Le déclenchement de commande peut se faire via des interfaces en ligne de commandes mais aussi via des éléments d'interface dédiés, de menus ou par des gestes (e.g. Figure 11).



**Figure 11. Exemple de déclenchement de commande : passage d'une vue « jour » à une vue infrarouge.**

Une même tâche peut parfois être réalisée de plusieurs manières comme c'est le cas pour la suppression d'un fichier. Glisser le fichier dans la corbeille est une manipulation, alors qu'effectuer un clic-droit sur un élément puis sélectionner « supprimer » est un déclenchement de commande. La manipulation exige la présence d'un objet d'intérêt, une action comme remplacer toutes les occurrences d'un mot dans un document ne peut être faite par manipulation.

Actuellement, interagir avec une jumelle infrarouge multifonction de Safran se fait uniquement à l'aide de boutons et l'ensemble des fonctionnalités disponibles (demandant une action de la part de l'utilisateur) sont des commandes (déclenchement d'un traitement d'image, lancement de l'affichage infrarouge, prise de photo, etc.). L'évolution des besoins va nécessiter des interactions complexes à réaliser telles que la définition d'une zone d'intérêt, le déplacement de *waypoint*<sup>4</sup> ou la gestion de l'espace d'affichage. De telles actions sont plus adaptées à la manipulation directe puisqu'elles agissent sur des objets graphiques.

Pour supporter la manipulation directe, il est nécessaire de doter le produit d'un medium d'interaction permettant ce type d'interaction. Un système de suivi du regard est un medium permettant de la manipulation directe. Dès 1982, Bolt (Bolt, 1982, March) propose par exemple le « put that there », une interaction multimodale de manipulation directe associant un pointage de l'objet d'intérêt au regard associée à une verbalisation de la transformation à apporter à l'objet. Le regard pourrait donc autoriser ce type d'interaction.

D'un autre côté, de par l'augmentation du nombre de fonctionnalités associées à des commandes, il est aussi nécessaire d'étudier d'autres mediums d'interaction que les boutons pour permettre le déclenchement de ces commandes. Le regard pourrait aussi être une modalité d'interaction permettant de faciliter et d'accélérer l'accès aux différentes commandes.

Dans la suite, nous allons nous concentrer sur le déclenchement de commande au regard. Nous allons montrer quelles sont les problématiques liées à l'utilisation de l'œil pour sélectionner ou activer des éléments graphiques avant de présenter des solutions à ces problèmes.

<sup>4</sup> Point de cheminement, point notable pour la définition d'une trajectoire (i.e. virage, point de passage obligatoire, etc.)

## 2.4.2 Problématiques du déclenchement de commande au regard.

*“We believe that many fundamental limitations exist with traditional gaze pointing. In particular, it is unnatural to overload a perceptual channel such as vision with a motor control task.”*

(Zhai, Morimoto, & Ihde, 1999)

L'interaction active au regard possède néanmoins certaines limites. Certaines sont d'ordre physiologiques, comme la précision maximale atteignable avec le regard. D'autres sont dus au fait que l'œil est normalement un organe purement senseur, et à ce titre, le contrôle conscient du regard n'est pas trivial et il est nécessaire de dissocier analyse et intention d'interaction.

### Précision :

Ce problème, qui a été évoqué dans la section 2.1, est la limite en termes de précision (environ 1°) obtenable au regard. Cette précision est due à la précision du système de suivi du regard et aux micromouvements de l'œil, même lorsque celui-ci est considéré immobile. Cela induit un problème comparable au phénomène du « fat finger » lors de l'utilisation d'un écran tactile, c'est-à-dire que les items avec lesquels il est possible d'interagir sont plus petits que la tolérance du dispositif de pointage (le doigt dans le cas du tactile), engendrant une interaction intempestive avec d'autres items proches.

### Contrôle de l'œil :

Une question se pose sur notre capacité à contrôler les mouvements de notre œil. Comme cela a été évoqué dans la section 2.3, notre œil est conditionné à réagir de manière inconsciente à certains attracteurs visuels, ce qui risque de compromettre l'utilisation du regard comme medium d'interaction. Le contrôle des mouvements du regard consiste en deux aptitudes : supprimer les mouvements non consciemment désirés et avoir la capacité d'effectuer l'ensemble des mouvements désirés. Une utilisation motrice de l'œil en plus de l'utilisation sensorielle peut alors engendrer un conflit ayant pour résultat, soit de limiter les capacités perceptives de l'œil, soit d'empêcher une utilisation motrice efficace.

### Différence entre vision et action :

Un troisième problème qui a été nommé *Midas Touch* par Jacob (Jacob, 1993) représente le fait que l'œil est actif en permanence et de fait, il devient difficile de différencier une simple analyse de l'environnement (activité purement sensorielle) d'une intention d'action de la part de l'utilisateur (activité motrice de l'œil). Cela résulte en activations intempestives d'items qui sont seulement analysés par l'utilisateur. Contrairement avec la souris, qui peut être utilisée à la fois comme dispositif de pointage et comme déclencheur d'action le regard comme modalité d'interaction pêche dans ces deux domaines. Comme dispositif de pointage, il devient impossible de laisser le regard à une position sur l'écran alors que l'attention visuelle est localisée à un autre endroit. En tant que dispositif actionneur, l'œil ne dispose pas d'élément comme le bouton de la souris permettant de signifier la demande d'action et de la situer dans le temps.

Les problèmes de précision se corrigent par des mécanismes identiques à ceux de l'interaction tactile : création d'éléments plus imposant, répartition équitable de l'espace, etc. Afin de contrebalancer le problème du *Midas Touch*, plusieurs méthodes ont été proposées afin de

permettre à l'utilisateur de signifier plus efficacement son intention d'interaction. Les méthodes existantes ainsi qu'une nouvelle proposition sont présentées dans la section suivante.

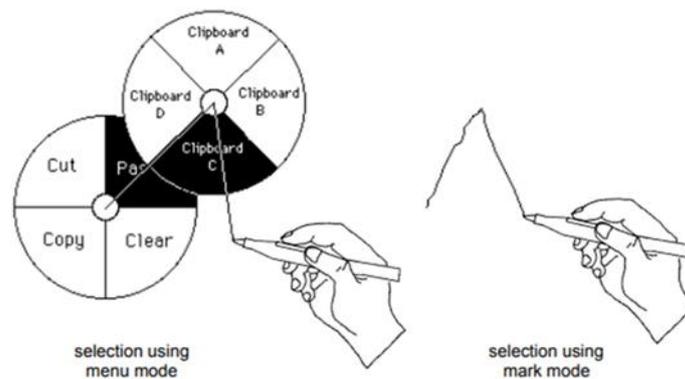
## 2.5 Modalités d'interaction au regard

Les différentes méthodes d'interaction sont appelées modalités d'interaction. Selon Vernier & Nigay (Vernier & Nigay, 2000), une modalité d'interaction est définie comme un couple < système physique, langage d'interaction >. Le système physique (medium d'interaction) représente le dispositif utilisé pour l'interaction (souris, clavier, regard) et le langage d'interaction décrit la façon dont le système est utilisé pour générer de l'information (ex : double-clic, appui sur une touche, clin d'œil). Dans ce document, lorsque ce n'est pas précisé, le terme modalité peut à la fois déterminer une modalité simple ou plusieurs modalités multiple (on parle alors de multimodalité).

Les techniques de déclenchement de commande au regard sont séparées en trois catégories : celles basées sur des gestes oculaires, celles basées sur un temps d'attente (*DwellTime*) et les modalités dites multimodales, nécessitant un second medium d'interaction complétant le regard.

### 2.5.1 Interaction gestuelle au regard

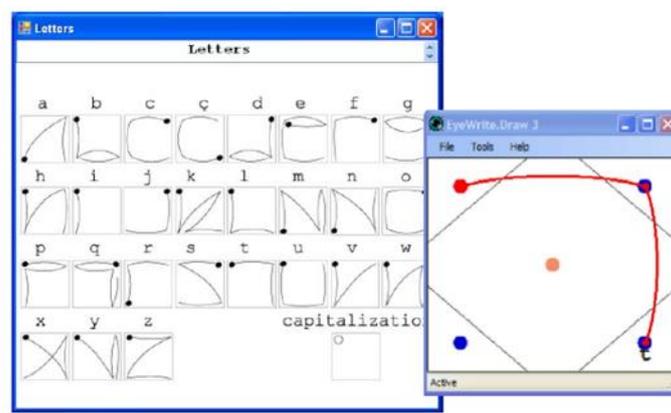
L'interaction gestuelle est déjà utilisée avec d'autres mediums d'interaction que le regard. Des gestes effectués avec un stylo, une souris, la main ou encore le corps peuvent être utilisés pour communiquer avec des systèmes. Avec la main ou le corps, il est possible de concevoir des gestes statiques, c'est-à-dire ne nécessitant pas de trajectoire. C'est le cas du « ok » en plongée ou le pouce et l'index forment un cercle, le mouvement permettant d'arriver à cette configuration de la main n'importe pas, seule la représentation finale importe. On parle de signe. Avec un dispositif de pointage comme un stylo, une souris ou le regard, un geste est défini par la trajectoire empruntée par le dispositif de pointage. Dans le cas de l'écriture, c'est la trace de la trajectoire qui dispose d'une sémantique, indépendamment de la manière d'obtenir cette trace (i.e., il est possible de commencer le tracé d'une lettre par la fin). Il existe aussi des interactions gestuelles où la sémantique est incluse dans la dynamique. C'est le cas des *marking-menus* par exemple, utilisables à la souris ou de manière tactile (Kurtenbach & Buxton, 1994) (cf. Figure 12).



**Figure 12.** Utilisation novice (gauche) et experte (droite) d'un marking-menu (Kurtenbach & Buxton, 1994). La trajectoire du geste dispose de la sémantique.

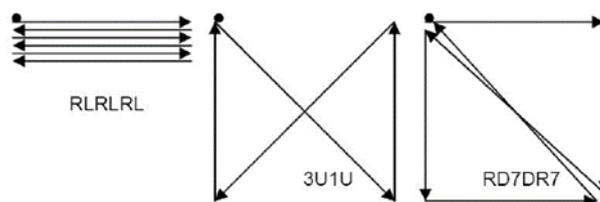
Drewes et Schmidt (Drewes & Schmidt, 2007) introduisent le geste au regard, qui se caractérise comme un geste dynamique dont la trajectoire dispose d'une sémantique. Et bien que dans les communications humain-humain, un geste de l'œil puisse être composé de mouvements du sourcil, de la paupière, etc., nous restreignons ici l'étude aux gestes liés aux mouvements de l'orientation du regard obtenus par des systèmes tels que ceux qui ont été présentés dans la section 2.2.

A l'instar d'une utilisation novice des *marking-menus*, le geste oculaire peut être guidé par l'interface graphique (Bee & André, 2008, June). Le geste peut alors être mémorisé comme une navigation à travers un menu, la trajectoire du geste est définie par cette navigation dans les menus et sous-menus. EyeWrite (Wobbrock, Rubinstein, Sawyer, & Duchowski, 2008, March) considère la création de gestes pour la saisie de caractères en concevant des gestes par similitude à la trajectoire effectuée avec un stylo (cf. Figure 13).



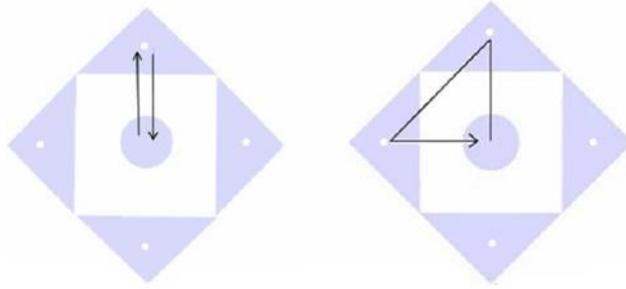
**Figure 13. Gestes oculaires de chaque lettre de l'alphabet dans le logiciel EyeWrite. A droite, un guide visuel pour faciliter la gestuelle des yeux.**

Il est aussi possible que le geste n'ait pas de sémantique particulière prédéfinie. Dans ce cas, l'alphabet gestuel est arbitraire. C'est le cas des gestes proposés par Drewes & Schmidt (Drewes & Schmidt, 2007) permettant le contrôle du *Windows media player* (cf. Figure 14).



**Figure 14. Gestes au regard étudiés par Drewes (Drewes & Schmidt, 2007).**

Istance et ses collègues (Istance, Hyrskykari, Immonen, Mansikkamaa, & Vickers, 2010) proposent un autre alphabet de gestes au regard contenant moins de saccades que les gestes proposés par Drewes. Les gestes sont basés sur une suite de régions à survoler avec le regard. C'est l'ordre de survol de ces différentes régions qui représente la sémantique du geste oculaire. Les deux gestes simples représentés sur la Figure 15 permettent un alphabet de 12 actions.



**Figure 15. Gestes au regard étudiés par Instance composés de 2 ou 3 saccades (Istance, Hyrskykari, Immonen, Mansikkamaa, & Vickers, 2010).**

Bien que la plupart des expérimentations concluent sur une expérience utilisateur acceptable, notamment pour des personnes en situations de handicap, l'interaction gestuelle au regard dispose de plusieurs limites :

- Déclenchements intempestifs

Dans les différentes études, sont mises en évidence des cas de détection d'un geste sans intention de la part de l'utilisateur de l'avoir effectué. En effet, dans l'ensemble des mouvements effectués par l'œil au naturel (comme organe senseur), certains se confondent avec les gestes des alphabets proposés. Cela produit des déclenchements de commande non voulus de la part de l'utilisateur et surtout difficiles à interpréter. Comprendre qu'il a effectué un geste et comprendre quelle commande a été effectuée demande un effort de la part de l'utilisateur. Pallier ce problème est possible en sélectionnant des alphabets plus complexes, coûtant à l'utilisateur un plus grand nombre de saccades à effectuer et un effort plus important de mémorisation.

- Manque d'Intuitivité

L'interaction gestuelle au regard n'est pas intuitive. Elle demande de la part de l'utilisateur l'acquisition d'une certaine habileté à manier et contrôler son regard. Cela demande l'apprentissage d'un alphabet gestuel et des actions qui lui sont associées. Les mouvements à effectuer sont souvent éloignés des mouvements naturels du regard afin de limiter les déclenchements intempestifs. Afin d'effectuer avec précision les différents gestes, l'utilisateur peut avoir besoin d'un guide visuel (cf. Figure 13), les gestes ne peuvent pas être effectués dans le « vide ».

- Sensibilité aux perturbations

Le geste au regard n'est pas décomposable, ainsi il doit être réalisé d'une traite et peut être facilement perturbé par un manque de concentration ou par un attracteur visuel.

- Fatigue

Les études citées ci-dessus concluent sur une faible fatigue liée au déplacement du regard. Cependant, une grande quantité d'énergie est nécessaire pour corriger d'éventuels déclenchements intempestifs, et pour apprendre puis se rappeler de l'alphabet gestuel.

L'interaction gestuelle peut tout de même être appliquée dans le cas où un faible nombre de commandes est disponible et d'une fréquence d'activation faible, impliquant un alphabet limité, la

possibilité de choisir des gestes bien discriminés du chemin naturel du regard et une perturbation limitée des fonctions perceptives de l'œil.

## 2.5.2 Interactions basées sur un temps de fixation (*DwellTime*)

Lors de la mise en évidence du *Midas Touch*, Jacob (Jacob, 1993) reprend les travaux de Ware et Mikaelian (Ware & Mikaelian, 1987) qui proposent de ne sélectionner un élément que si celui-ci est regardé pendant un certain temps. Il nomme cette méthode « *DwellTime* ». Il propose à cette époque un temps d'attente de 200 ms pour une tâche dont la sélection intempestive n'a pas d'impact négatif. La sélection non intentionnelle d'un item est corrigée par la sélection d'un autre item. Cette méthode a été régulièrement réétudiée et des tentatives ont été faites pour l'améliorer.

Le premier critère qui doit être fixé est le temps d'attente sur un élément avant que celui-ci ne soit activé. La détermination de ce temps se fait par un compromis entre la réactivité de l'interface et la minimisation du nombre d'activations involontaires. Lorsque l'activation intempestive d'un élément induit une correction non-triviale, les 200ms proposées par Jacob ne sont pas suffisantes. Dans la plupart des applications (Hansen, Johansen, Hansen, Itoh, & Mashino, 2003; Helmert, Pannasch, & Velichkovsky, 2008; Ware & Mikaelian, 1987) un temps de fixation d'environ 500ms (jusqu'à parfois 1000ms) est utilisé et semble être un bon compromis entre vitesse d'activation et faible taux d'erreur comme le montre Figure 16 (Spakov & Miniotas, 2004). Majaranta et ses collègues (Majaranta, Ahola, & Spakov, 2009) proposent de laisser l'utilisateur adapter la durée de fixation lors d'une saisie de caractères. Pour cette tâche répétitive, la durée est initialement fixée à 850ms et diminue rapidement aux environs de 250ms au fur et à mesure que l'utilisateur gagne en expertise. Špakov (Spakov 2004), lui, propose un ajustement automatique de cette durée. Le temps optimal semble converger aux alentours de 500ms  $\pm$  100ms. Deux approches existent sur la manière de calculer le temps de fixation, l'approche continue et l'approche cumulée. Dans l'approche continue, le temps est remis à zéro si l'utilisateur regarde ailleurs ne serait-ce qu'un instant. Dans l'approche cumulée, il ne l'est pas.

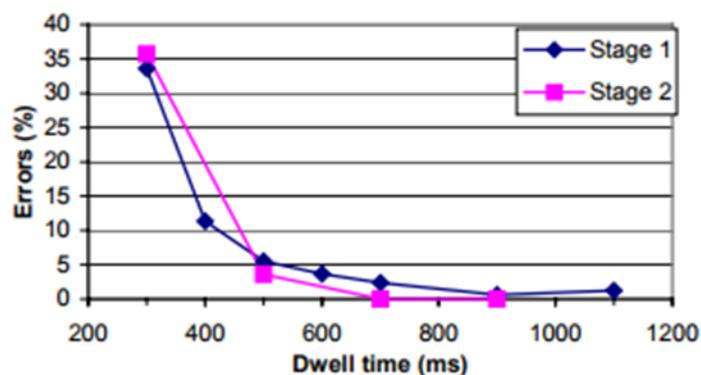


Figure 16. Pourcentage d'activations involontaires en fonction de la durée de fixation.  
(Špakov2004)

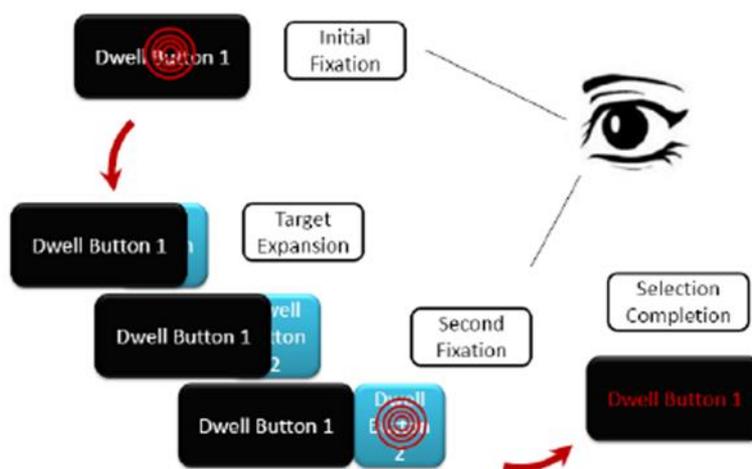
Un deuxième critère qui a été largement étudié est le type de feedback visuel lié à ce temps d'attente. Ce feedback permet à l'utilisateur d'avoir conscience que son regard a un effet sur l'élément fixé. Bednarik (Bednarik, Gowases, & Tukiainen, 2009) utilise la saturation d'une couleur

pour montrer l'avancement de la durée. Cependant, cette variable ne permet pas à l'utilisateur d'évaluer le temps qui lui reste avant l'activation. GazeTalk (Hansen, Lund, Aoki, & Itoh, 2006) utilise une bordure éclairée qui apparaît sur l'élément fixé et se contracte vers le centre. Une fois le centre atteint, l'élément est sélectionné. Tobii propose un feedback sous forme de cercle sur l'élément (cf. Figure 17). Le rayon d'un cercle balaye l'élément, une fois le cercle entièrement dessiné, l'élément est activé. Dans ce cas, la nature opaque du cercle peut gêner l'interaction. Le retour visuel augmente la compréhension de l'interface mais ne doit pas être trop intrusif.



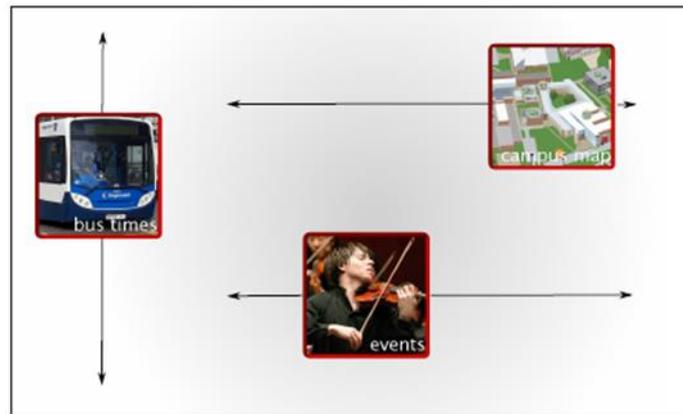
*Figure 17. Exemple de feedback visuel, l'item est sélectionné lorsque le cercle est entièrement complété.*

Des méthodes légèrement différentes basées sur un temps de fixation ont été proposées. NeoVisus (Tall, 2008) par exemple propose une activation en deux temps. La fixation d'un élément fait apparaître dynamiquement un second élément qui doit être regardé à son tour pour valider la sélection (cf. Figure 18).



*Figure 18. NeoVisus : sélection DwellTime en deux temps. Un temps court sur le premier élément fait apparaître le second élément de validation. (Image tirée de (Mollenbach, 2010)).*

Une méthode de fixation d'un élément basé sur l'utilisation de la poursuite lissée a été proposée par Vidal et ses collègues (Vidal, Pfeuffer, Bulling, & Gellersen, 2013). Les objets sont en mouvement permanent et la détection de la fixation d'un élément est faite par la détection d'une poursuite lissée suivant ce mouvement (cf. Figure 19). Cela permet notamment de s'abstraire des problèmes de précision (*accuracy*) du système de tracking. Cependant cela implique une interface en mouvement permanent.



**Figure 19. Pursuits :** *Les éléments de l'interface sont en mouvement permanent et la détection d'une fixation est liée à la détection d'une poursuite lissée.*

Les méthodes à base de *DwellTime* ont évidemment des limites :

- Adaptation du temps de fixation

Comme cela a été évoqué, la détermination du temps de fixation est un compromis entre une interaction rapide et naturelle et une réduction du nombre d'activation involontaires. La calibration de ce temps est complexe et sûrement dépendante de l'utilisateur. De plus, la découverte d'une interface est soumise aux effets du *Midas Touch* puisque l'utilisateur doit analyser l'interface pour en prendre connaissance.

- Attraction visuelles liée au feedback

L'utilisation d'un feedback dynamique représentant l'évolution du temps de fixation peut avoir un effet d'attracteur visuel, impliquant que l'utilisateur reste plus longtemps que prévu sur l'élément.

L'interaction au regard basée sur un temps de fixation permet d'utiliser l'œil comme dispositif de pointage. Le temps de fixation contrebalance les effets du *Midas Touch* tout en permettant une interaction assez naturelle en interagissant directement avec l'objet d'intérêt. Cependant, les erreurs sont courantes. Elles sont dépendantes du temps d'attente choisi.

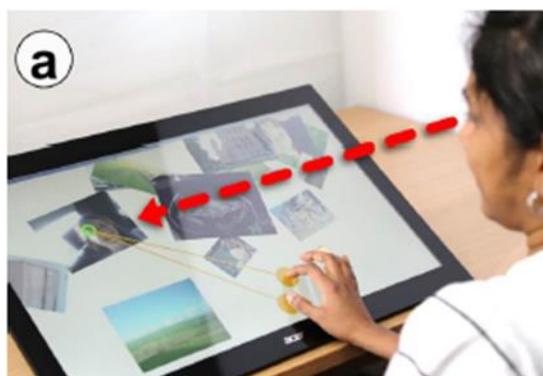
### 2.5.3 Interaction multimodale

La recherche autour de l'interaction au regard est souvent réalisée dans un contexte d'utilisateurs en situation de handicap. Les interactions monomodales (utilisant seulement l'œil comme moyen d'interaction) sont souvent privilégiées et l'interaction au regard est souvent vue comme un moyen de libérer les autres canaux de communication de l'homme. L'interaction multimodale est alors vue comme contradictoire à ce besoin. Il est possible d'utiliser le regard conjointement à d'autres modalités d'interaction.

Avec le « put that there », Bolt (Bolt, 1980) a étudié les prémisses de la multimodalité regard-voix. Le regard sert alors de dispositif de pointage alors que la voix permet de déterminer les commandes à effectuer. Le système OASIS (Glenn III, et al., 1986, September) est un exemple d'implémentation de cette multimodalité.

Il est possible de combiner le regard à des mouvements du visage tels que le clignement des yeux ou le froncement de sourcils (Majaranta & Rähä, 2002). Comme le clignement de l'œil est effectué à une fréquence d'environ 10 fois par minute, il est nécessaire d'effectuer un clignement plus long pour spécifier une interaction. Certaines personnes trouvent fatigant le fait de devoir effectuer un mouvement avec le sourcil alors que d'autres utilisateurs trouvent le clignement de l'œil trop disruptif (Grauman, Betke, Lombardi, Gips, & Bradski, 2003).

La simple présence d'un bouton physique comme une touche de clavier ou le bouton de la souris permet à l'utilisateur de communiquer son intention d'interaction. Dès 1987, Ware et Mikaelian (Ware & Mikaelian, 1987) mettent en avant un problème de synchronisation œil-main, l'utilisateur appuyant régulièrement sur le bouton avant même de regarder ce qu'il désire sélectionner. Jacob (Jacob, 1993) conclut aussi sur une plus grande efficacité des méthodes monomodales au regard par rapport à l'appui sur une touche. Cette méthode a cependant continué d'être étudiée car elle permet notamment de s'abstraire totalement du problème du *Midas Touch*. Pfeuffer et ses collègues (Pfeuffer, Alexander, Chong, & Gellersen, 2014) proposent plusieurs modalités d'interaction utilisant le regard en combinaison avec une modalité tactile multi-point pour effectuer de la manipulation d'objets avec des résultats convaincants (cf. Figure 20). Stellmach et Dachzelt (Stellmach & Dachzelt, 2012) proposent d'interagir avec des systèmes à distance en utilisant le regard comme dispositif de pointage et un smartphone comme interface tactile mobile distante. Salvucci et Anderson (Salvucci & Anderson, 2000) montrent des temps similaires entre l'interaction à la souris et l'interaction multimodale bouton + regard. Malgré un taux d'erreurs plus élevé, ils constatent une préférence des utilisateurs pour sa méthode, notamment pour des tâches de sélection.



**Figure 20. Interface multimodale regard + tactile pour de la manipulation d'objets virtuels (Pfeuffer, Alexander, Chong, & Gellersen, 2014).**

Le regard peut aussi être utilisé en coordination avec des interfaces cerveau-machine (BCI). Zander et ses collègues (Zander, Gaertner, Kothe, & Vilimek, 2010) décrivent ce type de multimodalité comme prometteuse pour contrebalancer le *Midas Touch*.

Notre cas d'application nous pousse à nous intéresser surtout à la multimodalité utilisant un bouton physique. Les interfaces vocales sont mises à l'écart par les situations opérationnelles bruyantes ou discrètes. Les BCI ne sont pas encore assez matures pour permettre une utilisation naturelle. Les gestes faciaux sont jugés trop contraignants et possiblement incontrôlables en situations de stress.

La principale limite de l'interaction multimodale regard + bouton est le besoin de synchronisation précise entre les mouvements de la main et les mouvements du regard. L'utilisateur n'ayant pas une

conscience parfaite de la position de son regard au cours du temps, il est fréquent qu'il appuie juste avant d'avoir fixé l'élément avec lequel il veut interagir ou juste après l'avoir quitté des yeux.

## 2.6 Conclusion, approche et méthode

Comme nous l'avons vu, les deux difficultés principales de l'interaction active au regard sont :

- la précision, somme de l'erreur due au système de suivi du regard et de l'erreur due à l'amplitude des micromouvements de l'œil,
- le discernement par le système des phases d'analyse des intentions d'interaction de l'utilisateur.

Ces difficultés ont conduit à la création de différentes modalités d'interaction au regard. Toutes ces modalités possèdent des avantages et des inconvénients distincts et sont censées s'adapter plus ou moins bien à certaines situations. A notre connaissance, la comparaison des différentes modalités d'interaction au regard dans des contextes variés n'a pas encore été étudiée en profondeur. Un des enjeux importants de la recherche sur l'interaction au regard est de comprendre quels sont les éléments du contexte des modalités d'interaction qui ont une influence sur les performances de ces modalités. Safran souhaite intégrer l'interaction au regard dans les systèmes optroniques. Cela suscite des enjeux similaires : il est nécessaire de comprendre avec précision les cas d'utilisation de ces systèmes afin de proposer des modalités d'interaction au regard optimales.

La volonté de Safran d'assurer la pérennité de ses interfaces nous a contraint à étudier l'intégration du suivi du regard sur des systèmes déjà conçus. Cette contrainte nous a conduit à nous intéresser plus précisément aux modalités d'interaction avec les menus présents dans les produits d'optronique Safran.

Dans un premier temps, notre but a été de structurer les éléments du contexte des modalités d'interaction au regard à travers une modélisation (cf. Chapitre 3). Cette modélisation est le résultat d'une approche top-down d'analyse des modèles de la littérature et d'une approche *bottom-up* qui a tiré profit de l'expertise métier en ergonomie de Safran et des observations lors des phases d'exploration de la technologie de suivi du regard. Elle vise à décrire des facteurs qui nous ont paru pouvoir impacter les performances des modalités d'interaction. Cette modélisation, qui sera présentée dans le chapitre suivant, vise à permettre de comprendre comment chaque caractéristique du contexte impacte les performances des modalités d'interaction au regard, et comment le fait d'interagir au regard modifie cette caractéristique.

## Chapitre 3

# Modélisation du contexte des modalités d'interaction au regard

Nous avons évoqué les difficultés pour créer une interaction générique au regard pour l'interaction dans des menus. En effet, les avantages et les inconvénients des différentes modalités d'interaction étudiées font qu'elles sont plus ou moins adaptées à des cas d'utilisation différents. Ces cas d'utilisation représentent différentes occurrences de ce que l'on appelle le contexte d'utilisation, c'est-à-dire de l'environnement au sens large dans lequel se déroule l'interaction. Depuis plus d'une vingtaine d'années, le contexte est considéré comme un élément important de la conception des systèmes interactifs. Un système est développé en tenant compte du cadre réel dans lequel il est utilisé. Cet environnement prévu impacte l'utilisation du produit et donc nécessairement aussi la conception.

Pour définir le contexte d'utilisation, nous partons de la définition du contexte d'utilisation d'un système de (Dey, Understanding and using context, 2001) :

*« Le contexte couvre toutes les informations pouvant être utilisées pour caractériser la situation d'une entité. Une entité est une personne, un lieu, ou un objet qui peut être pertinent pour l'interaction entre l'utilisateur et l'application, y compris l'utilisateur et l'application. »*

Nous préciserons que l'entité peut aussi être une tâche. La notion la plus importante de cette définition est la notion de pertinence afin de restreindre l'information relative au monde dans son ensemble à une information utile à la description du système étudié. Le contexte est extensible à volonté par transitivité. Si un objet A fait partie du contexte de l'objet d'étude O et si B fait partie du contexte de A, alors B fait partie du contexte de O. Dans la pratique, c'est le rôle de spécialistes de définir les limites du contexte, c'est-à-dire l'ensemble des entités pertinentes. Cela dépend du besoin de l'étude du contexte. Brézillon (Brézillon, 2007) avance « [qu']il n'y a pas de contexte sans contexte ». C'est-à-dire que l'étude du contexte d'un objet est définie par sa finalité. L'ensemble des informations d'un contexte doit servir un but précis. C'est ce but qui va aussi orienter la méthode d'analyse du contexte ainsi que la précision, la quantité et la nature des informations apportées.

Définir le contexte d'un objet d'étude peut avoir plusieurs fins. Lors de la conception, il permet d'évaluer quels éléments influencent le système. En les identifiant, les concepteurs peuvent anticiper des situations et adapter l'objet d'étude à son besoin en fonction du contexte. C'est une étape clé de la conception centrée utilisateur (ISO 13407 dans (Maguire, 2001), remplacée depuis par la norme ISO 9241-210). Cela peut englober de simples trivialisés comme de ne pas utiliser de restitution graphique pour un système destiné à des personnes aveugles. Mais cela permet de répondre à des problèmes moins évidents. Equiper les téléphones portables de vibreurs est par exemple dû au fait que ceux-ci sont régulièrement dans la poche et que l'utilisation publique du système implique un besoin de discrétion. Le contexte de la localisation et le contexte social ont bien eu un impact sur la conception de cette fonctionnalité des téléphones.

Comme on a pu le voir, le contexte est utilisé au moment de la conception pour anticiper le comportement de l'objet d'étude. Mais le contexte est aussi utilisé de manière dynamique. Chez l'homme, l'environnement influence grandement la prise de décision dans une fenêtre temporelle autour du moment de la décision, pas au moment de sa conception. Par exemple, s'il fait chaud, on a plutôt tendance à choisir une boisson fraîche. Afin de créer une expérience agréable pour l'utilisateur, il est nécessaire que certains systèmes adaptent leurs comportements de manière automatique. Une application capable de détecter si un utilisateur est connecté à internet pour adapter les fonctionnalités qu'elle propose ou un site web qui détecte l'appareil sur lequel il est visualisé pour adapter automatiquement son affichage sont des bonnes pratiques bien intégrées dans les développements informatiques actuels.

Le contexte peut aussi jouer un rôle évaluatif. En intégrant un système dans un environnement, il peut être nécessaire d'utiliser cet environnement pour évaluer le degré d'intégration de l'objet d'étude.

Lors de l'utilisation d'un système optronique durant une mission militaire, le contexte joue possiblement un rôle important sur la **tâche d'interaction**. Ce cas d'utilisation induit un besoin de performance du système en temps et en fiabilité. L'importance de la mission entraîne du stress et une grande concentration de la part de l'utilisateur. La **tâche utilisateur** elle-même fait partie du contexte d'utilisation puisqu'elle est la raison d'être de la tâche d'interaction. Elle détermine notamment des facteurs comme la fréquence de réalisation de la tâche d'interaction ou comme le type de système à utiliser en parallèle.

Notre méthode pour modéliser le contexte des modalités d'interaction au regard a consisté en deux approches : une approche top-down et une approche *bottom-up*.

L'approche top-down a consisté à nous inspirer des différents modèles du contexte issus de la littérature et à les confronter à nos besoins de décrire les phénomènes pouvant impacter les performances des modalités d'interaction au regard.

L'approche *bottom-up* a consisté à remplir et compléter notre modèle à partir de l'expertise métier des ergonomes de Safran et de nos observations lors des différentes phases d'utilisation de la technologie de suivi du regard. La mise en évidence de caractéristiques pouvant influencer les modalités d'interaction au regard est le résultat de séances de créativité avec les ergonomes Safran, d'interviews avec des représentants des utilisateurs finaux et de l'analyse de la littérature sur des domaines transverses tels que la perception.

## 3.1 Modèle du contexte d'une modalité d'interaction au regard

La première étape de la création d'un modèle du contexte d'utilisation est de définir son but et sa finalité. Ce que nous cherchons à faire est **d'étudier des modalités d'interaction au regard dans des menus afin de les évaluer** vis-à-vis du contexte dans lequel elles peuvent être utilisées. L'objet d'étude principal (O) de notre contexte est donc une modalité d'interaction au regard. Notre but est de déterminer statiquement et *a priori* quelles sont les entités extérieures qui peuvent impacter les performances des différentes modalités d'interaction au regard.

Les premières approches de définition du contexte ont suivi cette ligne de conduite et ont été spécifiques à des cas d'utilisation. Bucur (2005) ou Alarcón (2004) ont modélisé le contexte à l'aide d'ontologies pour le travail coopératif. Des travaux ont été réalisés pour essayer d'abstraire la description du contexte, motivé la plupart du temps par la nécessité que le contexte soit implanté dans des systèmes informatiques, pour des interfaces plastiques par exemple (Calvary, Coutaz, & Thevenin, 2001). Dey (Dey, 2001) formalise l'information des éléments du contexte comme des couples clé/valeur et donne ainsi une méthode de gestion de l'information. Cette représentation est facilement transposable informatiquement et facilite l'intégration du contexte dans les systèmes et la communication d'informations sur le contexte. Crowley (Crowley, Coutaz, Rey, & Reignier, 2002) suivi par Rey (Rey, 2006) considèrent le contexte de réalisation d'une tâche comme un ensemble de situations, c'est-à-dire un ensemble de réalisation du contexte. Il décrit ces situations (aussi appelées « états » dans (Rousseau, Bellik, Vernier, & Bazalgette, 2006) par exemple) comme un ensemble de rôles et de relations où les rôles sont les entités à disposition pour réaliser la tâche et les relations représentent les liens entre ces différentes entités. Cela dans le but de fédérer différents processus, qui faisant appel à ces entités, partagent alors leur vision du contexte. Cette approche permet aussi de modéliser l'évolution du contexte au cours du temps.

Afin de faciliter la compréhension du contexte et à des fins de tentative de normalisation, plusieurs modèles ont été proposés, mais qui ne satisfont pas tout à fait à nos besoins :

- (1) Jang et Woo (Jang & Woo, 2003) classent les éléments du contexte dans des catégories associées aux questions : *Who (utilisateur)*, *What (système)*, *Where (localisation)*, *When (temps)*, *Why (intention/émotion) and How (gestes)*. L'objet d'étude du modèle est ici l'application (considérée comme le système), qui a vocation à être dépendante du contexte. Ces questions ne sont pas explorées plus en profondeur. Par exemple, à la catégorie *Who* est seulement associé un nom d'utilisateur.

Ce modèle est complet en ce qui concerne les questions à se poser mais n'a pas la prétention d'être descriptif et d'approfondir les questions.

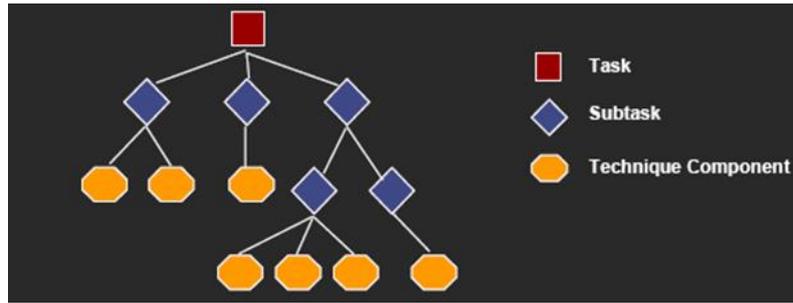
- (2) Bazire et Brézillon (Bazire & Brézillon, 2005, July) proposent un modèle générique du contexte. Il le décompose en la somme des contextes liés à l'utilisateur, au système, à l'environnement et à l'observateur. Le modèle inclut aussi les relations entre chacune de ces catégories. Il a pour vocation à être restreint en fonction des cas d'utilisation.

Le modèle porte ici sur l'activité et dispose de l'ensemble des catégories qui nous intéressent mais ces catégories ne sont pas explicitées. Les éléments qui composent ces catégories ne sont pas décrits.

- (3) Dey et Abowd (Dey & Abowd, 2000) considèrent que les quatre informations primordiales du contexte sont la position, l'identité, le temps et l'activité. Toutes les autres informations sont des sous-informations de ces quatre méta-informations. Son but est de fournir une méthode générique d'implémentation d'applications ayant « conscience » du contexte. De fait, son modèle n'est pas descriptif et ne permet pas de décrire directement une situation. Cependant, en appliquant ce modèle à l'interaction, le système ou l'utilisateur ne sont pas pris en compte comme faisant partie du contexte.
- (4) Dans sa modélisation du contexte, Crowley et ses collègues (Crowley, Coutaz, Rey, & Reignier, 2002) modélisent le contexte de la réalisation d'une tâche par un utilisateur. La tâche et l'utilisateur sont alors des paramètres du contexte et ne sont plus considérés dans le contexte. Ils ne sont pas décrits.
- (5) Pour la création d'interface plastique, Calvary et ses collègues (Calvary G. , et al., 2003) proposent une décomposition du contexte du système comme l'association de l'utilisateur, de la plateforme (moyens physiques et logiciels d'interaction avec le système) et de l'environnement. Dans ce modèle, la tâche à réaliser n'entre pas en compte dans le contexte d'utilisation du système.
- (6) Zimmermann et ses collègues (Zimmermann, Lorenz, & Oppermann, 2007) étendent la définition du contexte de Dey en y ajoutant les relations, représentant le lien pouvant être établi avec d'autres entités. De fait, cette catégorie peut intégrer le système et l'utilisateur, considérés comme des entités distinctes. Cette modélisation est à un niveau méta par rapport à nos besoins et n'explique pas les critères du modèle applicable à une interaction.
- (7) Kubicki et ses collègues (Kubicki, Lepreux, & Kolski, 2013) se basent sur le modèle (5) et définissent le contenu de chaque élément du triplet <Plateforme, Utilisateur, Environnement> pour une modélisation du contexte de tables tactiles connectées. Définir les éléments du contexte permet de définir d'un point de vue logiciel les caractéristiques pertinentes à l'adaptation dynamique du système au contexte.
- (8) Pour leur besoin de la définition du contexte de systèmes critiques de santé, Baas et ses collègues (Baas, et al., 2014) appliquent la méthode proposée par Coutaz et Calvary (Coutaz & Calvary, 2012) à une vision étendue du triplet <Plateforme, Utilisateur, Environnement> en y ajoutant la tâche et l'organisation pour correspondre à la pyramide AUTOS proposée par Boy (Boy, 2011).

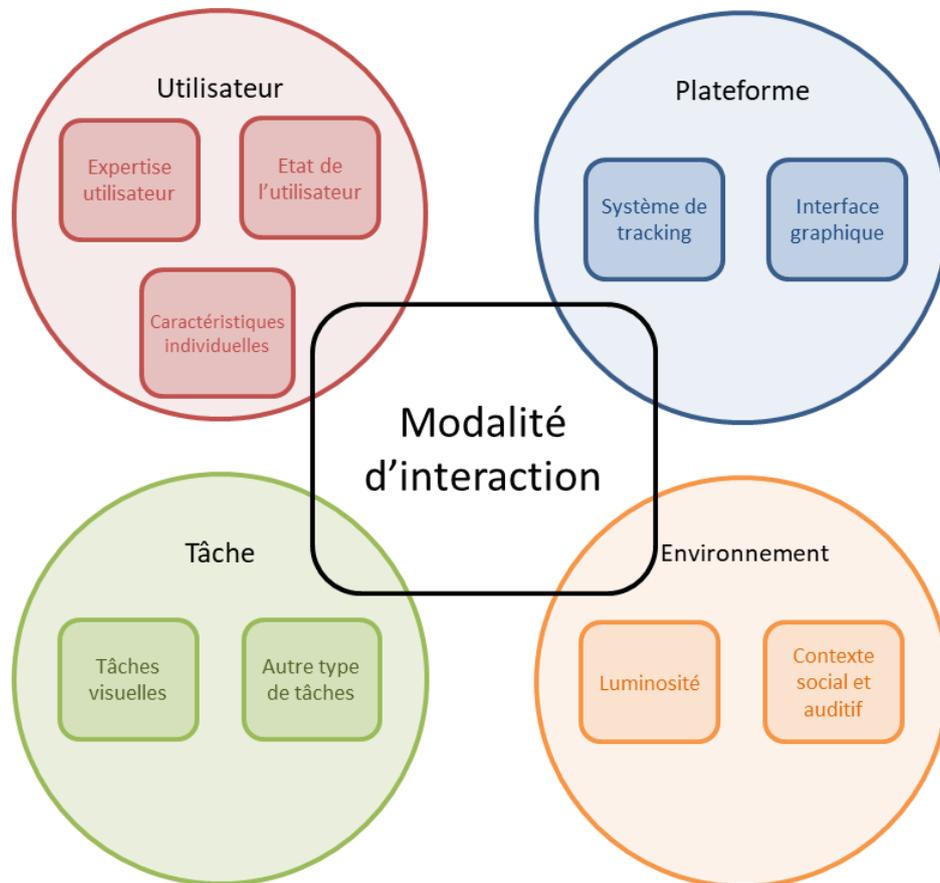
En nous inspirant de ces modèles, nous proposons un modèle du contexte des modalités d'interaction au regard à partir de 4 catégories : l'utilisateur, l'environnement, la tâche et la plateforme.

Pour plus de précision sur le positionnement de la définition d'une modalité d'interaction par rapport à la tâche utilisateur, on peut utiliser une décomposition de la tâche comme celle proposée par Bowman (Bowman, 2000) (cf. Figure 21). La tâche est décomposée en une somme de sous-tâches élémentaires auxquelles sont associées des méthodes techniques de réalisation. Ce sont ces *technique components* qui sont les modalités d'interaction, aussi appelées techniques d'interaction, qui sont l'objet de notre étude. On voit que les branches différentes de celle reliée à une feuille représentant la modalité d'interaction étudiée font possiblement partie du contexte, tout comme les caractéristiques générales de la tâche.



**Figure 21. Décomposition de la tâche (Bowman 2000)**

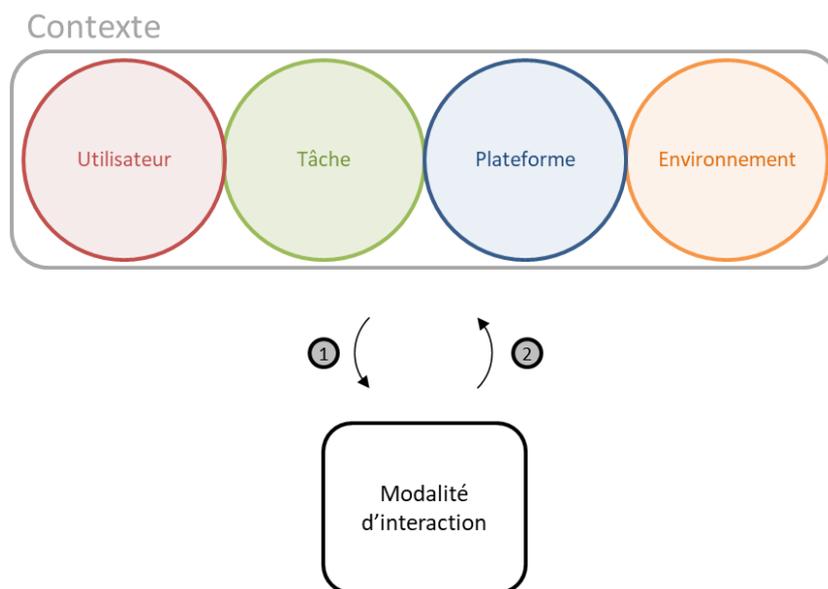
Dans le modèle du contexte de l'activité proposée par Calvary et ses collègues (Calvary G. , et al., 2003), la modalité d'interaction fait partie de la plateforme. Dans notre modèle, elle devient l'objet central du contexte. Ainsi, le modèle du contexte que nous étudions se présente comme la composition de contextes liés à l'utilisateur, à la plateforme, à la tâche et à l'environnement.



**Figure 22. Notre vision du contexte d'une modalité d'interaction au regard.**

Nous tenons à rappeler que le contexte n'englobe pas entièrement les catégories qui le composent, mais seulement leurs caractéristiques pertinentes. L'ensemble des catégories présentées sur la Figure 22 sont présentées dans les sections suivantes.

Ainsi, étudier l'influence du contexte revient à étudier comment les différentes caractéristiques du contexte impactent la modalité d'interaction. A des fins de simplification, nous considérons ces caractéristiques comme indépendantes. Cette simplification est discutée dans la conclusion.



**Figure 23. Influence réciproque du contexte sur une modalité d'interaction (1) et d'une modalité d'interaction sur les éléments du contexte (2).**

Cependant, l'inverse est vrai aussi, et l'utilisation d'une modalité influe sur les éléments du contexte (cf. Figure 23). Par exemple une modalité d'interaction difficile à utiliser peut stresser un utilisateur, modifiant par la même occasion son comportement. L'étude de l'intégration d'une modalité d'interaction dans un contexte doit alors aussi étudier comment cette dernière modifie les éléments du contexte. L'efficacité d'une modalité n'est pas seulement jugée par ses performances propres, mais aussi par son influence sur les éléments de son contexte. Il est possible d'utiliser des éléments du contexte comme variables d'évaluation des modalités d'interaction. Par exemple, le niveau de stress engendré par une modalité d'interaction peut être une variable d'évaluation pertinente dans certains cas.

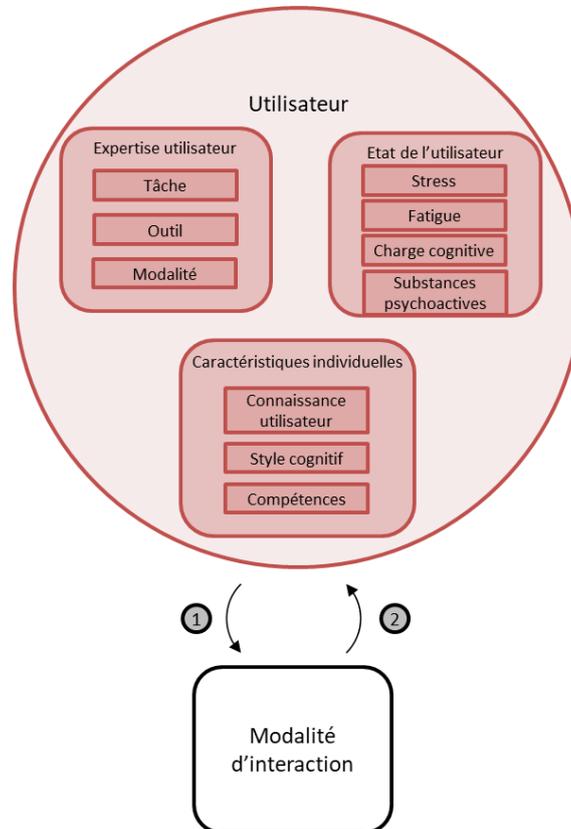
Evidemment, il est très probable que des effets de bouclage existent. C'est-à-dire que la modification de la modalité fasse évoluer les éléments du contexte, influant eux-mêmes sur l'évolution des performances de la modalité. Reprenons l'exemple d'un utilisateur stressé, imaginons qu'une modalité lui engendre du stress et que cette modalité est moins performante lorsque l'utilisateur est stressé. Une modalité moins performante va rendre encore plus stressante l'interaction, etc. C'est un cas de système divergeant, mais il est aussi possible d'obtenir un équilibre. Si au contraire, la modalité d'interaction est plus performante lorsque l'utilisateur est stressé (le stress peut avoir des effets positifs comme une amélioration des capacités mémorielles, et sensorielles)<sup>5</sup>, alors dans ce cas le système en boucle fermée est convergeant (ou oscillant) et c'est l'équilibre atteint qu'il faut étudier. Il est aussi possible qu'il existe des termes croisés, c'est-à-dire des effets d'interaction entre plusieurs éléments du contexte. Notre première approche est de considérer les critères indépendamment afin d'évaluer leur impact propre sur les modalités d'interaction.

Notre but est de décrire ces quatre catégories à l'aide de critères plus précis appartenant au contexte. Les chapitres suivants représentent le travail de construction de ces catégories en gardant l'objectif de générer un contexte pour l'interaction au regard dans des menus. Lorsque cela est

<sup>5</sup> <http://www.healthguidance.org/entry/15537/1/The-Positive-Effects-of-Stress.html>

pertinent, nous avons inclus les observations sur l'utilisation du regard en tant que modalité d'interaction obtenues lors de nos expérimentations mais qui n'ont pas été étudiée plus en profondeur.

## 3.2 L'utilisateur



**Figure 24. Etude de l'impact réciproque de l'utilisateur sur les modalités d'interaction au regard.**

L'utilisateur du système, aussi appelé utilisateur final ou opérateur, est la personne interagissant avec le système, et donc celle qui utilise la modalité permettant la communication avec le système. Il est bien évidemment une des composantes majeures du contexte d'une modalité d'interaction. Lors de la conception d'un système interactif, l'utilisateur est souvent le centre d'intérêt. L'utilisateur est un terme générique employé pour décrire l'entité qui utilise un système, mais il existe autant d'utilisateurs différents que de personnes. C'est cette singularité des individus, de leurs besoins, de leur expérience et de leurs attentes qui fait que le système doit être adapté à ceux qui l'utilisent vraiment. Ainsi, par exemple, l'utilisabilité d'un système n'est pas forcément la caractéristique la plus désirable. Fischer (Fischer, 2001) avance « qu'avoir été concentré sur la création de systèmes utilisables, a souvent réduit la puissance d'expression de ces systèmes ou de leurs interfaces pour satisfaire à la fois des novices et des utilisateurs occasionnels » et dénonce une cible mal déterminée pour certains systèmes. Il donne en exemple la différence de l'utilisateur cible entre un automate pour retirer de l'argent, qui est un système simple et qui doit être accessible à tous, en opposition à un système complexe pour professionnel qui doit être utile, et ce même au détriment d'une utilisabilité parfaite à la première utilisation.

Deux types de modélisation de l'utilisateur peuvent être distinguées : une modélisation basée sur des stéréotypes ou une modélisation basée sur les caractéristiques.

La première a été développée par Rich (Rich, 1979). Elle consiste à considérer l'ensemble des utilisateurs comme une somme de sous-ensemble complémentaires, appelés stéréotypes. Ces sous-groupes ne sont pas arbitraires et sont définis en fonction des besoins, souvent parce qu'ils sont représentatifs d'une classe de personnes ayant un comportement identique vis-à-vis du système. En pratique, la création de ces stéréotypes est réalisée à partir des caractéristiques, et la modification d'une des caractéristiques de l'utilisateur oblige seulement à réassigner celui-ci à une nouvelle catégorie.

La seconde modélisation de l'utilisateur s'appuie sur des caractéristiques pertinentes par rapport au système étudié (les modalités d'interaction au regard ici) (Brusilovsky & Millan, 2007). L'utilisateur est alors représenté par un ensemble d'aspects supposés indépendants. Cela augmente considérablement le nombre d'utilisateurs caractérisables en fonction de ces caractéristiques mais permet une description plus fine. D'un point de vue dynamique, si une des caractéristiques de l'utilisateur évolue, le système doit s'adapter au changement de cette caractéristique seulement.

L'intérêt d'un système basé sur les stéréotypes est la gestion de la granularité de la caractérisation de l'utilisateur. Pour un système adaptatif par exemple, cela permet de prévoir plus facilement les modifications dynamiques à apporter au système en fonction du passage de l'utilisateur d'un stéréotype à un autre. Au contraire, pour de l'aide à la conception de systèmes, et encore plus pour l'observation de l'évolution d'un système, qui est une phase d'apprentissage, une modélisation basée sur des caractéristiques nous paraît plus adéquate.

L'utilisateur sera aussi moteur de l'évaluation des différentes modalités d'interaction. Les préférences des utilisateurs, leur expérience lors de l'utilisation des modalités sont des critères à prendre en compte.

Les caractéristiques de l'utilisateur qui nous ont semblé pertinentes pour l'interaction au regard sont : l'expertise de l'utilisateur, son état, et ses caractéristiques individuelles. Nous les présentons dans les sections suivantes.

### **3.2.1 Expertise de l'utilisateur**

Tout d'abord, et surtout dans le cas de l'étude d'une modalité d'interaction, il est nécessaire de séparer l'expertise dans la réalisation de la tâche, aussi appelée niveau de connaissance de l'utilisateur, de l'expertise dans l'utilisation de l'outil d'interaction et enfin de l'expertise dans l'utilisation de la modalité d'interaction. Ces trois caractéristiques sont orthogonales, c'est-à-dire que l'expertise d'un utilisateur dans la tâche « haut-niveau » à effectuer n'a par exemple pas de relation avec sa capacité à maîtriser telle ou telle modalité d'interaction. L'inverse est tout aussi vrai.

#### ***Expertise de la tâche***

L'expertise dans la tâche utilisateur correspond au niveau de connaissance des mécanismes permettant l'atteinte d'un objectif haut niveau de l'utilisateur. Ces mécanismes sont indépendants du système utilisé (voir expertise dans l'outil) ou des modalités d'interaction utilisées pour communiquer avec ce dernier (voir expertise dans la modalité d'interaction). Un utilisateur expert

dans sa tâche réalise plus aisément et plus naturellement la planification des actions à réaliser et optimise ses actions pour améliorer sa performance. Bransford et ses collègues (Bransford, Brown, & Cocking, 2004) retiennent les cinq caractéristiques suivantes :

- Un expert repère des traits et des patterns d'information que des novices ne perçoivent pas.
- Un expert a acquis un grand lot de connaissances organisé d'une manière qui reflète une compréhension profonde du sujet.
- La connaissance d'un expert ne peut pas être réduite à une somme de connaissances isolées mais reflète plutôt des contextes d'applicabilité.
- Un expert est capable de retrouver un élément de connaissance presque sans effort.
- Un expert possède plusieurs manières d'approcher une situation nouvelle.

D'un point de vue pratique, un expert est souvent plus rapide et plus précis dans la prise de décision et dans la réalisation de ses actions. Ces actions s'enchaînent plus rapidement puisque moins de ressources sont allouées à la réflexion pour la planification des actions.

Lors de tâches d'observation, comme cela peut être le cas dans l'utilisation de produits d'optronique portable, la stratégie de recherche visuelle est différente entre utilisateurs novices et experts comme elle peut l'être chez des radiologues analysant des radiographies ou des professionnels du jeu d'échec observant un échiquier (Reingold & Sheridan, 2011).

### ***Expertise de l'outil***

Dans l'utilisation d'un système, l'expertise de l'outil est différente de l'expertise de la tâche car elle ne concerne plus les connaissances générales sur la réalisation d'un but de haut niveau mais les connaissances sur le système ou le logiciel lui-même. Dans un logiciel, c'est par exemple connaître la fonctionnalité associée à chaque bouton et savoir localiser ces fonctionnalités sans effort en fonction des besoins de la tâche. Un dessinateur professionnel utilisant pour la première fois un nouveau logiciel de dessin sera expert dans la tâche (il connaît les actions à réaliser) mais novice dans l'utilisation de l'outil (il ne sait pas où se situent les fonctionnalités recherchées, ni même si elles existent). Souvent, les logiciels actuels intègrent des raccourcis vers les fonctionnalités offertes pour permettre une amélioration des performances des experts avec l'outil.

Lors de l'utilisation d'une interface graphique, l'expertise peut complètement changer la stratégie de l'utilisateur. Un utilisateur novice sera plutôt dans une phase d'exploration, découvrant les fonctionnalités disponibles et leur localisation. Un utilisateur moyen sera dans une phase de recherche, connaissant les fonctionnalités mais ne sachant pas précisément où elles se trouvent. Alors qu'un utilisateur expert effectuera un parcours quasi automatique pour accéder aux fonctionnalités qu'il connaît. Ce changement de stratégie peut grandement influencer sur l'adéquation d'une modalité d'interaction à des utilisateurs d'expertise différente dans la maîtrise de l'outil.

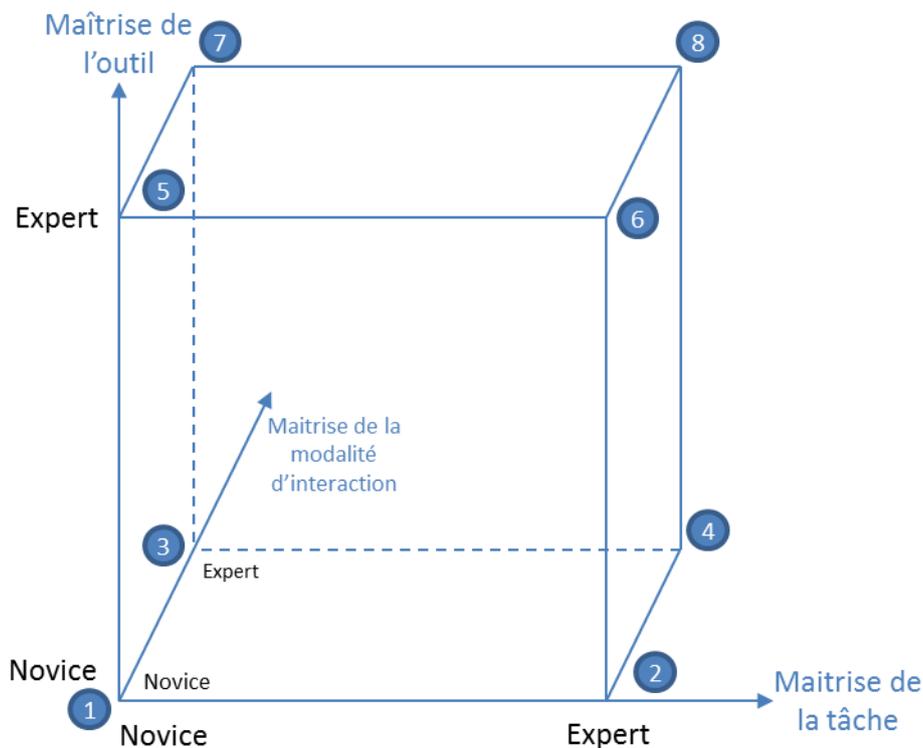
### ***Expertise de la modalité d'interaction***

L'expertise de la modalité d'interaction représente la connaissance des concepts liés à un médium d'interaction et l'habileté avec laquelle l'utilisateur est capable de manier ce dernier. S'habituer à utiliser une modalité d'interaction rend les actions à effectuer automatiques, on parle de mémoire musculaire (Kurtenbach G. P., 1993). Une modalité bien intégrée ne demande pas ou peu d'effort à

l'utilisateur. Au contraire, un utilisateur novice peut multiplier les erreurs mécaniques et avoir des difficultés à synchroniser ou séquentialiser des mouvements.

Pour l'interaction au regard, l'expertise se traduit par une capacité à détecter et compenser un décalage dû à la précision du système sans effort cognitif trop important. Cela engendre aussi une capacité de l'utilisateur à mieux « contrôler » son regard, c'est-à-dire à ne pas précipiter les mouvements du regard et à ne pas anticiper avec le regard les actions futures. A notre connaissance, la littérature sur l'interaction au regard n'a pas identifié de modalité d'interaction uniquement destinées à une utilisation par des experts.

Ces trois caractéristiques orthogonales donnent huit possibilités pour caractériser l'expertise de l'utilisateur (cf. Figure 25). Cela conduit à huit types d'expertise différents. Ces profils réagissent différemment en fonction des situations. On s'intéresse particulièrement à l'utilisabilité de la modalité d'interaction par rapport aux différentes expertises.



**Figure 25. Décomposition de l'expertise de l'utilisateur lors de l'utilisation d'un système.**

Nous analysons particulièrement les cas d'un utilisateur totalement novice (1), totalement expert (8) et les cas où l'utilisateur n'est novice que sur un des axes d'expertise (4,6,7) :

- 1 : un utilisateur novice dans les trois domaines. C'est une condition d'apprentissage difficile ne permettant pas d'appréhender chaque axe de manière efficace. Il est nécessaire que la modalité d'interaction soit propice à la phase d'exploration d'une interface dans laquelle l'utilisateur teste les différentes fonctionnalités qu'il a à sa disposition. Au regard notamment, la modalité doit permettre à l'utilisateur d'analyser les éléments sans pression de la part de la modalité.

- 4 : une non-expertise dans la maîtrise de l'outil entraîne un comportement différent de la part de l'utilisateur dont la stratégie est alors un mix entre exploration et recherche. La modalité d'interaction doit être adaptée à ces stratégies. Au regard, une interaction *DwellTime* ne permet par exemple pas d'analyser tranquillement un élément.
- 6 : l'utilisation d'une nouvelle modalité d'interaction peut poser certains problèmes à un utilisateur expert. Il essaie d'optimiser en temps la réalisation de sa tâche, et cela peut induire des erreurs sur l'utilisation de la modalité d'interaction. Le regard notamment est naturellement positionné à l'endroit du processus conscient de l'utilisateur (Pelz & Canosa, 2001), et cela rend difficile la réalisation de tâches simultanées ou l'anticipation efficace de tâches séquentielles. Un expert dans la tâche effectue des mouvements oculaires moins précis afin de générer une carte mentale de la zone observée. De ce fait, ses mouvements oculaires paraissent moins contrôlés (Cooper, Gale, Darker, Toms, & Saada, 2009). L'utilisateur a des chances de se retrouver frustré. Il n'arrive pas à accéder aux fonctionnalités dont il a besoin et l'accès à ces dernières lui demande un effort trop important par rapport à ses besoins.
- 7 : une non-expertise de la tâche peut influencer sur la manière qu'a un utilisateur d'utiliser un système qu'il connaît pourtant bien. D'un point de vue pratique, cela peut se traduire par des allers-retours entre l'interface et la zone de réalisation des actions. Lors d'une tâche d'observation par exemple, un novice effectue des mouvements plus systématiques qu'un expert qui aura des mouvements plus erratiques. Il est nécessaire que l'interaction ne soit pas gênée par ces mouvements experts « parasites ».
- 8 : c'est le cas d'utilisation visé dans les systèmes d'optroniques, les utilisateurs sont entraînés à la tâche, à l'outil et aux méthodes d'interaction avec le système. Ce cas permet des optimisations impossibles dans les 7 autres cas d'utilisation.

### 3.2.2 Etat de l'utilisateur

Ce que l'on nomme état de l'utilisateur correspond à des caractéristiques limitées dans le temps d'un utilisateur, physiologiques ou psychologiques, modifiant son comportement. Ainsi, bien que les effets de ces caractéristiques soient en pratique propre à chaque individu, l'influence de ces caractéristiques semble généralisable. Cette sous-section met en avant certaines de ces caractéristiques qui nous ont semblé avoir un impact potentiellement pertinent sur l'interaction au regard.

#### *Niveau de stress*

Nous ne parlerons pas ici des effets physiologiques à long terme dus à un stress prolongé, mais des effets immédiats, souvent positifs, d'une montée de stress. En effet, le stress dispose d'effets positifs sur le corps humain.

- Le stress a un impact physique. Il est lié à une libération d'adrénaline, qui augmente le rythme cardiaque et par là-même augmente la réactivité de la totalité de notre organisme. Cela se traduit par une amélioration des réflexes et de nos capacités perceptives. En diminuant aussi la douleur, le stress permet d'être plus endurant et de moins ressentir la fatigue.
- Le stress a un impact cognitif. Le stress permet à notre cerveau une meilleure concentration et nous aide à nous focaliser sur la tâche en cours. Les moments de *gaze aversion* (cf. sous-section sur la charge mentale) sont raccourcis voir supprimés.

- Il diminue aussi le temps de prise de décision et a un impact positif sur la mémoire.

Du point de vue de la modalité d'interaction au regard, le stress peut éventuellement avoir deux impacts.

- Le premier est lié à l'augmentation de la concentration et des réflexes. Lors d'une augmentation des capacités cognitives, un utilisateur d'un système critique, tel que des jumelles de combat, est en droit de s'attendre à ce que la réactivité et/ou la précision du système ne soit pas limitante de ses performances.
- Le second est le fait que le stress nous renvoie à notre état « sauvage », et le naturel reprend le dessus. Comme l'œil est par nature un organe senseur, il est possible qu'un certain état de stress empêche son utilisation efficace de manière motrice. Une meilleure conscience de notre environnement passe par une plus grande réactivité à des stimuli visuels et aux mouvements, et par une augmentation non maîtrisée de la fréquence des saccades.

Une augmentation du stress pourrait impliquer une diminution des performances générale de l'interaction au regard. Nous émettons l'hypothèse que les modalités à base de temps de fixations auraient tendance à être plus interrompues par des mouvements incontrôlés de l'œil. A l'inverse, le taux de faux positifs (déclenchements intempestifs) pourrait décroître à cause de la diminution du temps de fixation moyen. Cependant, les interactions multimodales seraient probablement affectées par un plus grand nombre de désynchronisations œil-main.

### ***Niveau de fatigue***

La fatigue a un impact négatif sur les performances humaines (Hancock and Desmond, 2001), en termes de perception, de réflexes et de concentration. Ainsi, nos mouvements sont plus lents mais peuvent aussi être plus réfléchis afin d'économiser de l'énergie. Lanthier (Lanthier, Risko, Smilek, & Kingstone, 2013) montre une baisse du temps de réaction à des stimuli visuels et une augmentation de la durée des fixations au fur et à mesure que la fatigue augmente.

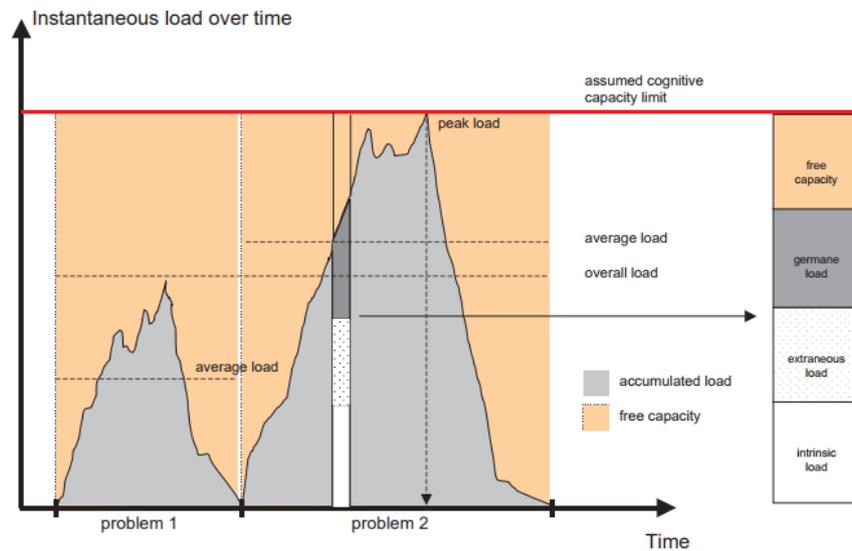
Cazzoli et ses collègues (Cazzoli, et al., 2014) ont étudié l'impact de la fatigue sur les mouvements oculaires et discerne deux types de fatigue différents. Le premier type de fatigue est dû à des facteurs chronotypiques, c'est-à-dire une fatigue due au moment de la journée et au chronotype de l'utilisateur (Est-il plutôt du matin, plutôt du soir ?). Les résultats montrent qu'un utilisateur présente des temps de fixation plus long en dehors de ses horaires de confort. Le second type de fatigue mis en avant est dû au temps passé à la tâche. Dans ce cas, plus le temps passe, plus les saccades sont lentes. Ces deux types de fatigue peuvent avoir des impacts sur les modalités d'interaction au regard.

Il est possible qu'une interaction basée sur un temps de fixation implique un plus grand nombre de faux positifs si l'utilisateur est dans un état de fatigue (ou nécessite un ajustement du temps de fixation). Au contraire, ce ralentissement de la cadence pourrait avoir des effets plus restreints (voir bénéfiques) sur une interaction multimodale, ou les problèmes de synchronisation seraient plus rare par exemple.

### ***Charge cognitive***

Dans une première approche, la charge cognitive représente la quantité de concentration mobilisée par le cerveau à un instant donné. C'est un état éphémère de l'utilisateur qui influe sur ses

performances. Cette concentration est due à l'ensemble des actions qui sollicitent le cerveau durant la tâche. Paas et ses collègues (Paas, Tuovinen, Tabbers, & Van Gerven, 2003) séparent cette charge cognitive en trois charges : la charge intrinsèque, la charge extrinsèque et la charge « *germane* » (liée aux processus d'apprentissage). Dans son modèle, les différentes actions réalisées par le cerveau contribuent au cumul d'une variable appelée charge mentale. Tout individu dispose alors d'un seuil de charge mentale qui ne peut pas être dépassé. Lors du dépassement de ce seuil, les ressources associées à chaque tâche sont réduites (cf. Figure 26).



**Figure 26. Représentation de la charge mentale selon Paas (Paas, Tuovinen, Tabbers, & Van Gerven, 2003).**

Dans des approches plus complexes, on considère que les différentes parties du cerveau traitent des informations différentes. Norman (Norman J. , 2002), sépare deux flux mentaux, le flux ventral et le flux dorsal (cf. Tableau 1). Le premier, aussi appelé flux « quoi » est plus lié à la reconnaissance, à la mémoire à long terme et aux actions conscientes alors que le second, flux « comment » ou « où » traite des actions motrices, de la mémoire à court terme. Ces deux flux interfèrent peu entre eux et peuvent fonctionner simultanément sans causer de surcharge mentale. Les baisses de performances dues à un excès d'actions à effectuer par le cerveau sont alors dues au fait de devoir effectuer plusieurs fonctions similaires au même instant.

Factor	Ventral system	Dorsal system
Function	Recognition/identification	Visually guided behaviour
Sensitivity	High spatial frequencies - details	High temporal frequencies - motion
Memory	Long term stored representations	Only very short-term storage
Speed	Relatively slow	Relatively fast
Consciousness	Typically high	Typically low
Frame of reference	Allocentric or object-centered	Egocentric or viewer-centered
Visual input	Mainly foveal or parafoveal	Across retina
Monocular vision	Generally reasonably small effects	Often large effects e.g. motion parallax

**Tableau 1. Séparation des actions mentales en deux flux.**<sup>6</sup>

Nous nous intéressons à deux aspects de la charge mentale. Le premier consiste à étudier l'influence d'une surcharge mentale sur les capacités à utiliser son regard comme modalité d'interaction. Le second consiste en l'étude de la réciproque : l'influence de l'interaction au regard sur la charge mentale.

Une première influence de la charge mentale est connue sous le nom d'effet tunnel (Kahneman, Beatty, & Pollack, 1967; Williams, 1982). Une forte charge mentale réduit les capacités perceptives d'un utilisateur. Du point de vue de la modalité d'interaction, cet effet est plutôt positif puisque le regard a moins de chance d'être attiré par des attracteurs visuels hors de la vision centrale.

Une seconde influence d'une charge mentale élevée est nommée « *gaze aversion* » (GA). C'est le fait qu'une grande demande cognitive pour se remémorer quelque chose ou pour réfléchir à une chose abstraite va pousser l'utilisateur à détourner son regard vers une zone de faible complexité visuelle dans le but de minimiser les interférences cognitives entre les deux processus (Schmidt & Zelinsky, 2006). Plus la complexité de la tâche augmente plus la GA s'intensifie. Il est donc possible qu'une charge cognitive élevée implique un besoin de l'utilisateur de diriger son regard vers des zones extérieures à l'interface.

Le second aspect consiste à étudier l'influence du fait d'interagir au regard sur la charge mentale, et plus généralement sur les différentes facultés cognitives de l'utilisateur. En effet, le fait d'utiliser l'œil de manière motrice n'est pas naturel et demande une certaine concentration de la part de l'utilisateur. Ainsi, la mesure de la charge mentale lors de l'utilisation des différentes modalités d'interaction peut servir de métrique d'évaluation. Plus la CM induite est basse, meilleure est la modalité sur ce critère.

### ***Substances psychoactives***

L'ingestion de substances psychoactives comme des stimulants, des déprimeurs, des hallucinogènes ou des antipsychotiques pourrait avoir des effets sur le comportement de l'utilisateur, et donc sur les capacités à interagir au regard. Aucune étude n'a été menée à notre connaissance sur le sujet.

<sup>6</sup> [https://en.wikipedia.org/wiki/Two-streams\\_hypothesis](https://en.wikipedia.org/wiki/Two-streams_hypothesis)

### 3.2.3 Caractéristiques individuelles

Les caractéristiques individuelles sont les caractéristiques les plus à même d'expliquer les différences inter-utilisateur. Elles ne dépendent pas ou peu de l'entraînement comme l'expertise utilisateur, et ne sont pas limitées dans le temps comme les caractéristiques de l'état de l'utilisateur. Toutes ces caractéristiques sont indépendantes du système et peuvent être transposées à d'autres modèles que celui de la modalité d'interaction au regard.

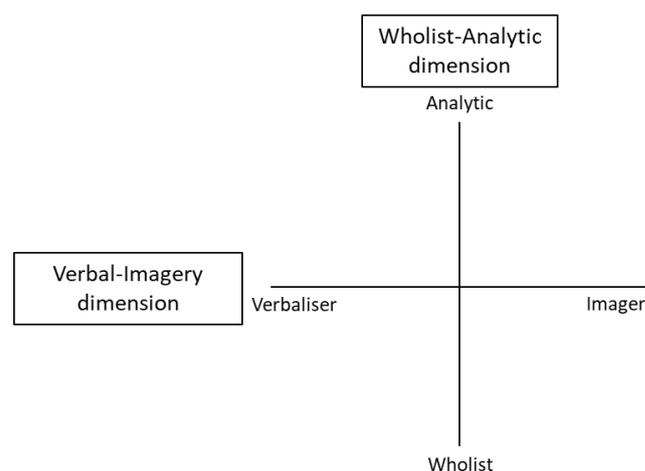
#### *Connaissances utilisateur*

La connaissance utilisateur est proche des éléments présentés dans la sous-section 3.2.1. Cependant, elle s'en distingue par le fait d'être indépendante d'un système donné. Elle correspond à la connaissance du type d'actions à réaliser, du type de système utilisé de manière plus générale. Fischer (Fischer, 2001) présente deux canaux de communication entre l'utilisateur et un système, le canal explicite et le canal implicite. La connaissance utilisateur permet une meilleure communication implicite, l'utilisateur est plus à même de comprendre des erreurs d'un système, et d'apprendre à s'en servir.

Cette connaissance peut être due à la culture de l'utilisateur (environnement de travail, lieu de vie, langue) ou aux autres systèmes qu'il est amené à utiliser.

#### *Style cognitif*

Les individus ne réfléchissent pas tous de la même manière. Le style cognitif fait référence à la manière dont un individu collecte, évalue et interprète des données (Harrison & Rainer Jr, 1992). Riding et Sadler-Smith (Riding & Sadler-Smith, 1997) proposent un modèle où le style cognitif de l'utilisateur suit deux dimensions : la dimension Verbal-Imagé et la dimension Analytique-Ensemble (cf. Figure 27).



**Figure 27. Deux dimensions du style cognitif (Riding & Sadler-Smith, 1997)**

Suivant ces auteurs, la dimension Analytique-Ensemble représente la manière dont l'individu organise et structure l'information. Les analytiques ont une vision *bottom-up* de l'information, ils considèrent en priorité les différents composants à traiter de manière indépendante. Les ensemblistes voient l'information de manière globale, avec une vue d'ensemble. Les ensemblistes ont plus tendance à traiter l'information en parallèle. Du point de vue de la modalité d'interaction,

cela signifie qu'ils ont tendance à récolter de l'information à plusieurs endroits pour les traiter simultanément, ce qui peut rendre plus difficile la concentration du regard à un endroit précis unique. Les analytiques, eux, disposent naturellement d'une concentration plus fine et d'un traitement de l'information et des tâches plus séquentiel, ce qui peut faciliter l'utilisation du regard comme modalité d'interaction.

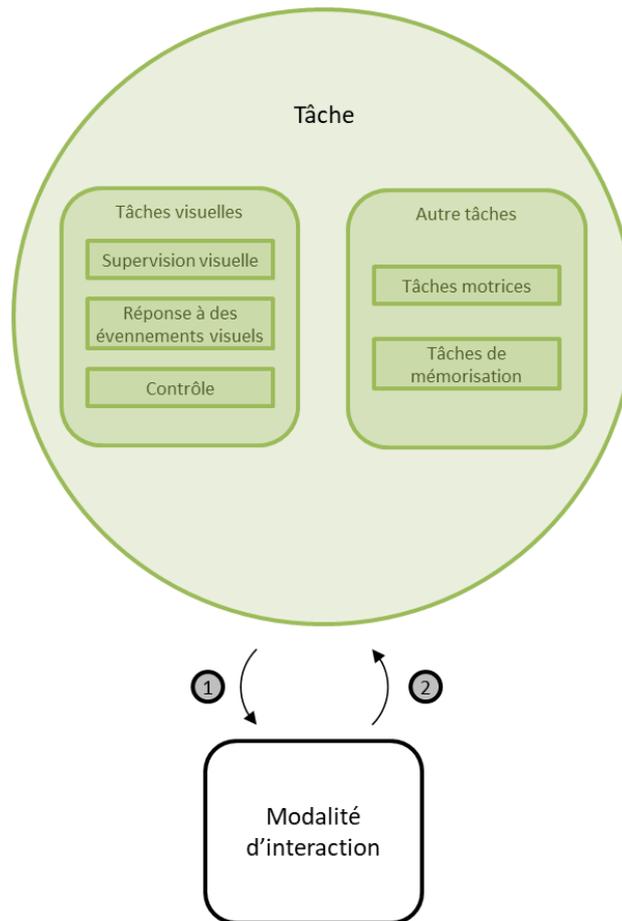
La dimension Verbal-Imagé décrit le mode de représentation de l'information. Les utilisateurs verbaux ont plus tendance à traiter et à retenir l'information sous forme de mots alors que les utilisateurs « imagés » perçoivent et traitent l'information à l'aide d'images mentales. Bien que cette dimension puisse impacter la représentation la plus adaptée à chaque individu pour une interface (texte ou icône par exemple), notre hypothèse est qu'il n'y a pas d'impact de cette dimension sur les modalités d'interaction.

### ***Compétences***

Les compétences d'un utilisateur correspondent à ses capacités physiques et psychologiques telles que la dextérité ou l'endurance. Ces compétences sont propres à chaque individu. Harrison et ses collègues (Harrison & Rainer Jr, 1992) ont montré une dépendance de ces caractéristiques à des caractéristiques génériques comme l'âge ou le genre. Nous avons listé les compétences pouvant avoir une influence sur l'interaction au regard :

- L'**acuité visuelle** est la capacité de l'œil à distinguer des détails d'une scène ou un motif visuel observé. Plus ces détails sont fins, plus l'acuité visuelle est élevée. Les défauts de l'œil non corrigés comme un faible astigmatisme ou une faible myopie, un astigmatisme ou la presbytie modifient l'acuité visuelle et peuvent avoir un impact sur l'interaction au regard. Toute dégénérescence de l'œil a sûrement pour effet de limiter la vision et donc l'interaction au regard.
- Le **contrôle du regard** représente la capacité d'un individu à restreindre les besoins d'utiliser son œil inconsciemment de manière sensorielle. Lors de nos expérimentations, nous avons pu constater que certains utilisateurs ne réussissaient pas à garder un contrôle suffisant pour pouvoir interagir au regard. Ils ne pouvaient pas réfréner les mouvements naturels de leur regard. Les personnes disposant d'un bon contrôle de leur regard sont plus à même de délocaliser leur attention visuelle (cf. sous-section 2.3.3)
- L'**agilité du regard** représente aussi une certaine forme de contrôle du regard. Elle correspond à la vitesse et la précision avec laquelle l'utilisateur manie son regard.
- La **dextérité et la synchronisation** sont des caractéristiques nécessaires, notamment pour des modalités d'interaction multimodales.
- La **vitesse d'acquisition d'informations visuelles** représente le temps nécessaire entre l'apparition d'une information dans le champ visuel et son traitement par le cerveau. Différents profils à ce niveau ont de grandes chances d'impliquer différents profils d'utilisation de l'interaction au regard. Thorpe et ses collègues (Thorpe, Fize, & Marlot, 1996) mettent en avant des temps de réaction allant de 220 ms à plus de 600ms. Cette compétence influe très probablement sur l'interaction au regard.

### 3.3 Tâche utilisateur



*Figure 28. Etude de l'impact réciproque de la tâche utilisateur sur les modalités d'interaction au regard.*

La **tâche utilisateur**, aussi appelée tâche de fond, se définit en opposition à la **tâche d'interaction**. La tâche d'interaction représente l'ensemble des actions qui sont effectuées pour communiquer avec le système. La tâche d'interaction sert la tâche utilisateur. La tâche utilisateur est donc le but réel de l'utilisateur, ce qu'il est en train de faire au-delà de l'utilisation du système, car l'utilisation du système est rarement une fin en soi. En choisissant cette définition en opposition à la tâche d'interaction, nous avons décidé d'inclure aussi l'ensemble des actions extérieures qui peuvent avoir un impact sur elle comme les différents mouvements de l'utilisateur ou d'autres tâches annexes comme parler avec une autre personne.

Dans les modélisations de la tâche comme celle de Bowman, le but est de mieux comprendre la tâche pour qu'elle puisse être exécutée plus simplement. Il existe aussi des modèles du contexte de la tâche, comme celui proposé par Bazire et Brézillon (Bazire & Brézillon, 2005, July), où la tâche est au centre de la modélisation. Mais, à notre connaissance, il n'existe pas de modélisation de la tâche utilisateur comme élément de contexte. Pour décrire la tâche utilisateur comme élément de contexte des interactions au regard, nous avons choisi de nous intéresser séparément aux tâches faisant intervenir des composantes de la vision des autres tâches.

La relation avec la tâche utilisateur peut être étudiée de deux manières : dans un premier temps, on peut étudier l'impact du fait d'interagir au regard sur les différentes tâches utilisateur. Cette étape est nécessaire pour intégrer l'interaction au regard dans des systèmes sans perturber la réalisation des tâches pour lesquelles les systèmes sont utilisés. Le centre de la conception d'un système interactif est la tâche utilisateur, pas la tâche d'interaction. Dans un second temps, on peut s'intéresser à l'inverse, c'est-à-dire à l'impact des différentes tâches utilisateurs sur les différentes modalités d'interaction au regard. Cette étape permet de mieux comprendre les modalités d'interaction au regard, et d'anticiper le comportement de l'utilisateur vis-à-vis de la tâche d'interaction lors de la réalisation des différentes tâches de l'utilisateur.

La suite de cette section présente différents types de tâche ainsi que leurs interactions avec une interaction au regard.

### 3.3.1 Tâches visuelles

Toute tâche d'interaction peut avoir un impact sur la tâche utilisateur, mais par la transformation de l'œil en organe moteur, les tâches d'interaction faisant intervenir le regard de manière motrice peuvent avoir des impacts différents sur les tâches utilisateur nécessitant l'utilisation de la vision. Reprenons les différentes tâches visuelles présentées dans la sous-section 2.3.1 : le **balayage visuel** (observation pure), la **supervision visuelle** (maintien de variables dans une fourchette), la **réponse à un évènement**, la **recherche**, la **lecture** et le **contrôle** (vérification de la réalisation d'actions). Pour ces 6 tâches, nous allons analyser l'impact éventuel de la tâche d'interaction sur la tâche de fond et inversement.

Tout d'abord, la lecture, la recherche et l'observation pure sont des tâches qui doivent être suspendues lors d'une tâche d'interaction. Il est tout de même possible que ces actions soient réalisées par intermittence, c'est-à-dire en alternant interaction et lecture ou interaction et observation. Dans ce cas, c'est le temps passé à regarder l'interface qui est limitant vis-à-vis de la tâche utilisateur. Plus l'utilisateur passe de temps à regarder l'interface, moins il en passe à lire ou à observer. L'impact est trivial et indépendant du fait que l'interaction soit effectuée au regard ou non.

#### *La supervision visuelle*

La supervision visuelle est la tâche qui consiste à vérifier qu'un ensemble de variables sont maintenues à l'intérieur de limites prédéfinies. Cette tâche visuelle demande d'alterner la supervision entre les différentes variables et donc de les visualiser à des fréquences plus ou moins élevées. Ce genre de tâche est classique lors de la conduite de véhicules notamment.

#### Impact de l'interaction au regard sur la supervision visuelle :

Le problème majeur ici est que lors d'une interaction au regard, le regard doit être présent dans la zone de l'interaction, il est monopolisé. Ainsi, s'il est impossible de faire une pause dans l'interaction pour continuer d'alterner la visualisation des différentes variables à analyser, la supervision visuelle se dégrade. Nous voyons deux solutions à ce cas : proposer une interaction qui n'est pas perturbée par des interruptions pour continuer la supervision, ou, pour des supervisions plus superficielles, proposer des modalités d'interaction qui permettent de continuer la supervision en vision périphérique.

Impact de la supervision visuelle sur l'interaction au regard :

La supervision visuelle implique des chemins visuels assez répétitifs et assez semblables. Imaginons qu'un utilisateur ait trois zones à visualiser pour superviser trois variables. Six chemins sont réalisés (1 vers 2, 2 vers 1, 1 vers 3, etc.). Ces schémas sont assez similaires aux mouvements à effectuer pour les interactions gestuelles au regard.

Si l'utilisateur priorise la supervision visuelle par rapport à l'interaction, ce sont les interruptions fréquentes de l'interaction qui font sortir le regard de la zone d'interaction qui vont le plus l'impacter. Comme nous venons de le dire, il est nécessaire d'avoir une modalité d'interaction robuste à ces interruptions.

***La réponse à des événements visuels***

L'apparition de signaux visuels peut être de deux types, attendue ou inattendue. Dans le premier cas, l'attention de l'utilisateur est sollicitée au préalable, et même sans dévier son regard de la zone actuellement fixée, l'utilisateur délocalise son attention afin d'être alerte vis-à-vis du signal qui doit apparaître. On parle alors d'attention visuelle endogène, parce qu'elle est guidée par le conscient de l'utilisateur. Le second cas représente l'apparition non prévue d'un signal visuel, le plus souvent en vision périphérique. Ce signal joue un rôle d'attracteur visuel et attire le regard, ou au moins l'attention (exogène dans ce cas) de l'utilisateur.

Impact de l'interaction au regard sur la détection de signaux attendus :

Habituellement, lorsque l'utilisateur est en attente de l'apparition d'un signal visuel, il va regarder la zone dans laquelle le signal peut apparaître, ou concentrer son attention visuelle dans cette zone s'il en a connaissance, ou il va défocaliser légèrement ses yeux, c'est-à-dire regarder légèrement dans le vague afin d'être plus alerte. Nous émettons l'hypothèse qu'en changeant la fonction de l'œil vers une fonction motrice, ces situations de division de l'attention visuelle et de délocalisation sont toujours possibles. D'un point de vue opérationnel, ce genre de tâche arrive régulièrement lors de missions d'observations où l'utilisateur est concentré sur une zone plus ou moins grande de son champ de vision pour détecter l'apparition éventuelle d'unités. Pour cette raison, nous avons choisi d'étudier ce type de tâche lors d'une expérimentation sur la capacité à délocaliser son attention visuelle tout en interagissant au regard dans la section 4.2.

Impact de l'interaction au regard sur la détection de signaux inattendus :

Dans le cas de signaux inattendus, l'attention visuelle est dirigée par des signaux extérieurs. Un trop grand focus sur la tâche d'interaction peut limiter les capacités à détecter des signaux en vision périphérique. Cela a été montré lors d'une augmentation de la charge cognitive (effet tunnel) mais il est possible qu'une composante supplémentaire existe, due à l'utilisation motrice de l'œil. Il est aussi possible que, pour privilégier l'interaction et afin de limiter les risques d'erreurs dus à la modalité, l'utilisateur limite lui-même sa perception visuelle périphérique en se concentrant fortement sur la zone d'interaction.

Dans les deux cas, qu'il s'agisse de signaux attendus ou inattendus, si l'interaction se fait sur un plan d'affichage plus proche que celui de la zone d'observation ou d'apparition des signaux visuels, cela va rendre plus difficile leur détection. En effet, Mazza et ses collègues (Mazza, Turatto, & Umiltà, 2005)

ont montré que l'attention visuelle est plus facilement dirigée sur les éléments de premier plan que sur les éléments d'arrière-plan.

#### Impact de la détection de signaux sur des interactions au regard :

Que le signal soit attendu ou inattendu, son apparition va solliciter l'attention visuelle de l'utilisateur et agir comme un attracteur visuel. Cela risque d'attirer le regard de l'utilisateur, et donc d'interrompre la tâche d'interaction. Ces interruptions risquent de perturber l'interaction au regard, mais il est possible que différentes modalités d'interaction au regard réagissent différemment à ce problème.

### ***Le contrôle***

Le contrôle est la vérification qu'une action attendue s'est produite. Soit le contrôle ne demande pas d'avoir le regard sur l'élément à contrôler lors du changement d'état de l'élément contrôlé, et dans ce cas, la réalisation est effectuée à la suite de la tâche d'interaction, soit, pour être détecté, le changement de l'élément doit être visualisé, dans ce cas, la réalisation de la tâche de contrôle est alors simultanée à la tâche d'interaction.

Dans le premier cas (e.g. vérifier qu'un changement de l'heure affichée a été pris en compte), il n'y a pas d'interférences entre la tâche d'interaction et la tâche utilisateur.

Un exemple du second cas est la mise en évidence d'un changement lors de l'interaction. Comme l'ont mis en évidence Rensink et ses collègues (Rensink, Kevin O'Regan, & Clark, 2000), la mise en évidence d'un changement (assez minime) nécessite de visualiser explicitement ce changement. En interagissant au regard, si l'action déclenchée est simultanée au déclenchement, il est alors impossible de discerner ce changement. La solution serait alors d'opter pour un délai après le déclenchement d'une action pour que celle-ci soit effectuée.

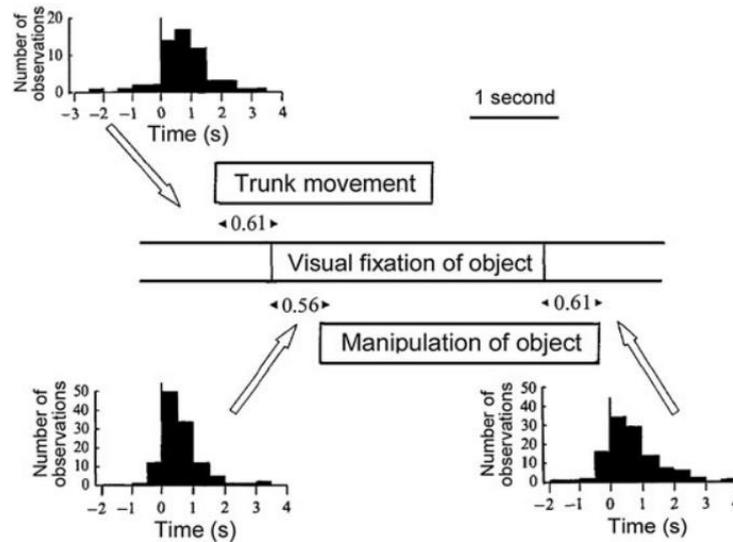
## **3.3.2 Autres types de tâches**

Parmi les tâches utilisateur ne faisant pas intervenir directement le regard, nous nous sommes intéressés brièvement à deux types de tâches, les tâches motrices et les tâches faisant intervenir la mémoire.

### ***Tâches motrices***

Dans une tâche motrice, l'état de la tâche est évalué en permanence et le mouvement est perpétuellement corrigé en fonction de cette évaluation. Cette évaluation peut être réalisée par plusieurs modalités sensorielles, de manière haptique par exemple, mais la perception visuelle est souvent mise en œuvre pour cette évaluation. C'est par exemple le cas lors de la conduite de véhicules.

Cependant, dans des travaux comme ceux de Land (Land, 2009), le regard est étroitement lié aux tâches motrices et précède légèrement le reste du corps (cf. Figure 29).



**Figure 29. Temps relatifs des mouvements du corps, de la fixation d'objets, et de la manipulation de ces objets.**

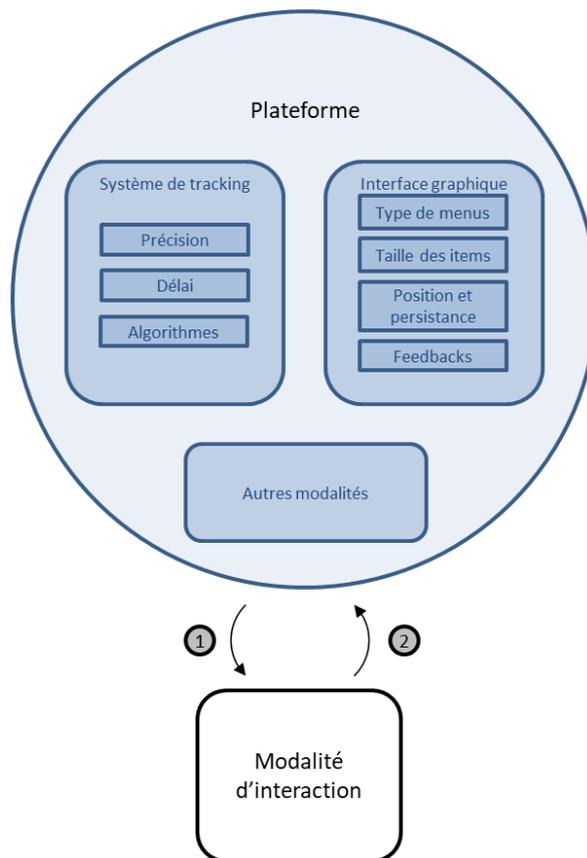
Il serait intéressant de comprendre comment le fait de solliciter le regard de manière active pour une interaction impacterait des tâches motrices.

### ***Tâches mémorielles***

Lors de tâches mettant en jeu la mémoire, le phénomène de *gaze aversion* (cf. sous-section 3.2.2) entre en jeu. Notre regard est alors attiré vers les zones de faible complexité visuelle. Nous avons dans ce cas l'impression de regarder dans le vide. Cette modification du comportement du regard peut avoir des impacts sur le fait d'interagir au regard.

L'inverse est aussi vrai, et la limitation du phénomène de *gaze aversion* implique une plus grande difficulté à utiliser ses capacités mémorielles durant une interaction au regard.

## 3.4 La plateforme



*Figure 30. Etude de l'impact réciproque de la plateforme sur les modalités d'interaction au regard.*

Selon Calvary et ses collègues (Calvary G. , et al., 2003), la plateforme représente l'ensemble des ressources logicielles et matérielles dont dispose le système pour l'interaction. La plateforme est le sous-ensemble du système faisant le lien entre l'utilisateur et le système. En considérant la plateforme comme faisant partie du contexte de l'interaction au regard, on suppose que le système dispose des prérequis physiques suffisants pour intégrer le suivi du regard. C'est-à-dire un système de suivi du regard, des performances nécessaires pour les algorithmes de traitement d'image, etc. Ainsi, selon nous, la plateforme comme contexte de l'interaction au regard se décompose selon trois axes : le système de tracking, l'interface graphique et les autres modalités d'interaction présentes.

### 3.4.1 Système de tracking

Les systèmes de suivi du regard ne sont pas tous équivalents ; les différents producteurs de systèmes de tracking proposent des produits aux caractéristiques différentes. Ils mettent en avant la précision, la fréquence d'acquisition ou encore des algorithmes de filtrage spécifiques. Des performances différentes impliquent des impacts éventuels sur les modalités d'interaction au regard. Les caractéristiques haut-niveau pertinentes sont la précision, le délai entre les mouvements de l'œil et leur utilisation par le système et les différents algorithmes mis en œuvre.

## ***Précision***

La précision dispose de deux aspects présentés dans la sous-section 2.2.3, l'offset de précision (*precision*) et la variance autour la position décalée (*accuracy*). Toutes les modalités d'interaction au regard ne répondent pas de la même manière aux problèmes de précision. L'interaction gestuelle par exemple, est robuste aux deux aspects de la précision puisque c'est la « forme » du chemin effectué par le regard qui importe. Il en est de même pour la fixation d'éléments mobiles (exemple donné à la sous-section 2.5.2) où c'est la direction du mouvement de l'œil qui importe et non le point de fixation. Les modalités comme le *DwellTime* ou une interaction multimodale ne sont pas robustes à un offset de précision, elles engendrent probablement des problèmes. Par exemple, une interaction nécessitant une validation explicite de la part de l'utilisateur lui permet de prendre le temps de corriger consciemment un offset de précision avant de lancer l'activation, là où une interaction à base de temps de fixation aurait déjà activé un autre élément proche.

Les systèmes visent la meilleure précision possible, mais la précision n'est pas limitée uniquement par le système de suivi du regard, mais aussi par l'amplitude des micromouvements de l'œil (cf. section 2.1).

## ***Délai***

Les concepteurs de systèmes de suivi du regard mettent souvent en place des algorithmes de filtrage pour moyenniser les micromouvements de l'œil lors des fixations. Mais tout filtrage induit un retard entre la position réelle du regard et la position détectée par le système. Augmenter la fréquence d'acquisition de la vidéo de l'œil permet de réduire le délai pour un filtre moyenniseur dont l'utilité est de diminuer l'erreur de précision. Le retard minimum, sans filtrage, est égal à l'inverse de la fréquence d'acquisition. Une fréquence de 50 Hz induit un retard minimum de 20 ms.

Dans certains cas, ce retard peut être perçu par l'utilisateur et nuire à son expérience d'interaction. Les modalités d'interaction disposant de feedbacks visuels sont plus aptes à laisser l'utilisateur percevoir ce délai.

## ***Algorithmes de prédiction***

Il existe des algorithmes de prédiction de la localisation précise du regard à partir d'estimations grossières de la position du regard comme le montre le travail de Hillaire et ses collègues (Hillaire, et al., 2010). En plus d'engendrer un délai (cf. ci-dessus), il est possible que ce genre d'algorithme ne soit pas compatible avec certaines modalités d'interaction (interaction gestuelle au regard par exemple), ou alors qu'il induise des erreurs de manipulation difficilement compréhensibles par l'utilisateur.

### **3.4.2 Interface graphique**

L'interface graphique joue un rôle majeur dans beaucoup de systèmes. C'est un des moyens les plus courants pour permettre de communiquer efficacement des informations sur l'état du système à l'utilisateur. Lors de l'interaction au regard, elle joue aussi un rôle central. Nous nous intéressons particulièrement à l'interaction au regard avec des menus, c'est-à-dire avec des éléments graphiques conçus pour l'interaction. Les caractéristiques décrites ci-après sont à la fois des contraintes à

respecter pour l'interaction au regard, mais peuvent aussi être des marges de manœuvre pour améliorer l'utilisabilité de certaines modalités d'interaction au regard.

### ***Type de menu***

Pour les interactions tactiles, un grand nombre de menus ont été développés (Bailly, 2009) qui s'adaptent mieux à la modalité que les menus linéaires classiques que l'on peut observer sur nos systèmes d'exploitation d'ordinateurs. Mais les différents types de menus impliquent des schémas d'analyse différents de la part des utilisateurs. Par exemple, les menus circulaires s'analysent de manière plus erratique que des menus linéaires qui s'analysent plutôt de manière systématique.

Des mouvements différents du regard sur l'interface peuvent avoir des impacts sur certaines modalités d'interaction au regard. Il est important de comprendre ces mécanismes car il est possible qu'une interaction au regard soit utilisée sur des menus conçus à la base pour d'autres modalités d'interaction. L'intérêt de Safran se porte actuellement vers l'intégration d'interactions au regard dans des menus composés de tuiles (cf. une expérimentation présentée au Chapitre 4).

### ***Taille des éléments des menus et distance inter items***

Comme il a été évoqué dans la section 2.4.2, l'interaction au regard sur des menus nécessite des éléments suffisamment larges (au moins 1° d'angle) pour bien fonctionner. Cependant, plus un item est large et plus les items sont éloignés, plus cela rend robuste l'interaction au regard. En effet, cela évite que l'utilisateur regarde un item et que le système considère qu'il regarde un item proche. Souvent, la zone fixable des items est conçue plus grande que les items eux-mêmes. Cela permet de compenser ce type d'offset.

Certains travaux proposent d'utiliser une taille d'éléments dynamique afin de pouvoir intégrer un plus grand nombre d'items (Spakov & Miniotas, 2005; Miniotas, Spakov, & MacKenzie, 2004). Lorsqu'un item est grossièrement regardé, il est mis en évidence en augmentant la taille et en décalant les items adjacents. Les mouvements du regard relatifs à l'item mis en évidence permettent de diminuer la précision nécessaire. Si l'utilisateur regarde vers le bas de l'élément actuellement mis en évidence, c'est l'item en-dessous qui est mis en évidence, et ce, que le point de fixation déterminé par le système de suivi du regard soit effectivement dirigé vers ce nouvel élément ou non.

### ***Position et persistance des menus***

Un des points qui nous paraît pouvoir impacter le fonctionnement des modalités d'interaction est le positionnement central ou non des menus. Les menus contextuels s'ouvrent à proximité d'un point d'intérêt et sont, du coup, souvent positionnés au centre de la zone d'affichage. Par leur positionnement central, ces menus sont survolés plus régulièrement par l'utilisateur si celui-ci continue d'observer ce qu'il y a à côté. Ces aller-retours sur le menu peuvent avoir un effet sur certaines modalités d'interaction. L'ouverture triviale au regard de sous-menus par exemple (ouverture instantanée d'un sous-menu lorsque l'item correspondant est survolé) nous paraît envisageable pour des menus excentrés, mais impensable pour des menus plus centraux.

Les menus contextuels sont souvent éphémères ; ils sont ouverts lorsque le besoin se présente puis se referment une fois utilisés. De fait, il n'est pas possible pour l'utilisateur d'analyser le menu en utilisant la vision périphérique pendant une autre action pour prévoir quel item sélectionner.

L'analyse et la sélection d'un item se fait une fois le menu ouvert. Sachant qu'un menu persistant et un menu éphémère proposent des interactions différentes, il est probable que les modalités les plus adaptées ne soient pas les mêmes.

### ***Types de feedback visuels***

Si un pointeur permanent à l'écran, comme pour une interaction à la souris peut être obstruant lors d'une interaction au regard, il est tout de même nécessaire de donner à l'utilisateur un retour sur l'interprétation de son regard par le système. Cela est nécessaire notamment pour comprendre et corriger un éventuel offset de position entre le regard de l'utilisateur et la position détectée par le système.

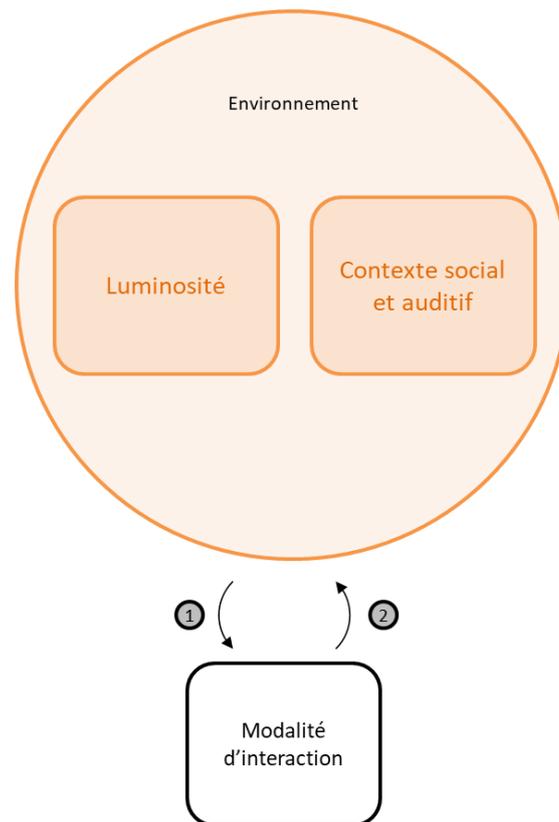
Plusieurs types de feedbacks visuels existent, le plus souvent, l'item actuellement regardé est mis en surbrillance, mais pour des interactions faisant intervenir un temps de fixation, il est utile que ce temps soit signifié à l'utilisateur. Cela peut être un cercle qui se dessine au-dessus de l'item (cf. Figure 17), un cercle autour de l'item, une jauge sur le côté de l'item ou une surbrillance qui s'étend depuis le centre de l'item vers les bords (cf. Figure 34). Tous les feedbacks visuels mettant en œuvre un mouvement agissent comme des attracteurs visuels et perturbent l'interaction en tant que tels (cf. sous-section 3.3.1). La conception de feedbacks adaptés est nécessaire à la création de modalités d'interaction au regard performantes.

### **3.4.3 Autres modalités du système**

Notons tout d'abord que la présence d'autres modalités d'interaction (i.e. boutons physiques, système de reconnaissance vocale, BCI, etc.) dans le système est un prérequis pour des interactions multimodales. Quoi qu'il en soit :

- L'utilisation d'autres modalités que le regard en fonction de la situation (du contexte) est une solution à certains problèmes évoqués dans ce chapitre. Comme il est peu probable qu'une modalité soit supérieure dans tous les domaines par rapport aux autres, cela peut être un choix de conception d'attribuer une autre modalité que le regard afin d'utiliser chaque modalité lorsqu'elle est la plus performante.
- L'utilisation simultanée d'autres modalités d'interaction (pédales et volant lors de la conduite par exemple) peut engendrer des interférences avec l'interaction au regard. Cela a été évoqué dans la sous-section 3.3.2.

## 3.5 L'environnement



*Figure 31. Etude de l'impact réciproque de l'environnement sur les modalités d'interaction au regard.*

Dans notre modèle du contexte, l'environnement représente les caractéristiques ne rentrant pas dans les autres catégories. Les critères de l'environnement sont extérieurs au couple utilisateur-système et à la tâche que l'utilisateur doit réaliser. Nous considérons les critères de l'environnement comme des contraintes, nous ne cherchons pas à modifier l'environnement pour satisfaire des besoins de la modalité d'interaction mais à adapter cette dernière à son l'environnement.

### 3.5.1 Luminosité

En étudiant l'interaction au regard, il était naturel d'étudier la luminosité comme facteur d'impact. La luminosité semble pouvoir impacter les modalités d'interaction au regard de deux manières différentes.

Dans un premier temps il faut considérer son impact système, c'est-à-dire comment elle influence le système de suivi du regard. La plupart des systèmes de suivi du regard actuels sont basés sur une méthode de réflexion cornéenne. Des réflexions lumineuses peuvent perturber voir totalement fausser les algorithmes de détection vidéo, remettant alors totalement en cause les performances de ces systèmes. Nombre des systèmes du commerce conseillent une utilisation dans un environnement lumineux contrôlé avec une lumière n'étant pas directement dirigé face à l'œil. Dans le cas où la luminosité ambiante a un effet sur la précision des systèmes de suivi du regard, les impacts sont les mêmes que ceux évoqués dans la section 3.4.1 sur les performances du système d'eye-tracking.

Le deuxième impact à considérer est celui de la luminosité sur les mouvements oculaires. En effet, dans le domaine de la photographie, il est bien connu que les zones lumineuses attirent le regard. La luminosité agit comme un attracteur visuel. L'inverse est aussi vrai, une trop forte luminosité va avoir tendance à écarter le regard, c'est le cas des phares pendant la conduite par exemple. La luminosité impacte aussi le temps traitement de l'information, ou augmente la fréquence de tâches de contrôle (Konstantopoulos, Chapman, & Crundall, 2010). Ainsi, des scènes à forts contrastes ou des zones trop lumineuses disposent de plus d'attracteurs visuels et sont plus à même de perturber l'interaction au regard.

Ces problématiques nous ramènent au choix de la luminosité d'une interface : trop lumineuse et elle gêne la tâche utilisateur, trop peu lumineuse, les performances de l'interaction risquent d'être dégradées.

Dans le cas de produits comme les jumelles optroniques, la luminosité extérieure peut être contrôlée pour ne pas perturber les algorithmes de détection, mais le problème de la luminosité des scènes d'observation reste pertinent.

### 3.5.2 Contexte social et auditif

Les éléments visuels ne sont pas les seuls à modifier le comportement du regard. C'est aussi le cas de l'audition et du contexte social.

Si le regard a une grande influence sur la représentation spatiale de notre environnement, notre audition sert aussi ce but, et des stimuli auditifs peuvent attirer le regard autant que des stimuli visuels. Lors de la chute d'un objet par exemple, le regard va être automatiquement attiré par le stimulus auditif lié à la chute.

L'environnement social impacte aussi les mouvements du regard, et ce plutôt de manière endogène. Il arrive de regarder une personne pour vérifier sa compréhension lorsqu'on lui explique quelque chose par exemple. Cela a le même effet que des stimuli auditifs et va engendrer des mouvements du regard. Cependant, l'environnement social perturbe un peu moins le contrôle du regard d'un utilisateur puisque les mouvements sont issus d'un processus cognitif conscient de la part de l'utilisateur.

Ces deux aspects ont en plus un impact sur la charge mentale (un fort bruit ambiant augmente la charge mentale par exemple), dont les implications sont évoquées dans la sous-section 3.2.2.

## 3.6 Conclusion

En nous inspirant des modèles de la littérature, des travaux antérieurs sur des domaines transverses à l'interaction au regard et grâce à l'expertise métier des ergonomes de Safran, nous avons proposé un modèle du contexte des modalités d'interaction au regard qui nous permet de mieux identifier les caractéristiques extérieures pouvant les influencer.

A partir de cette modélisation, les besoins industriels nous ont poussés à sélectionner trois caractéristiques de notre modèle afin d'évaluer cet impact (cf. Chapitre 4). Ces caractéristiques sont la forme des menus, notamment les menus linéaires et circulaires, et une tâche utilisateur visuelle de

réponse à des événements visuels attendus ou inattendus. La réponse à des événements visuels attendus correspond à la capacité de l'utilisateur à déporter son attention visuelle tout en interagissant au regard, c'est-à-dire à regarder à l'endroit de l'interaction tout en restant concentré à un autre endroit. La réponse à des événements visuels inattendus correspond, elle, à la capacité à détecter des stimuli en vision périphérique lors de l'interaction. Ces deux dernières caractéristiques sont importantes lors d'une tâche d'observation réalisée avec des jumelles où l'utilisateur doit être concentré sur la scène à tout instant, et attentif à l'apparition d'alertes visuelles.

## Chapitre 4

# Evaluation de l'influence du contexte sur les modalités d'interaction

Dans le Chapitre 3, nous avons présenté quatre axes du contexte d'une modalité d'interaction, chacun de ces axes étant associé à plusieurs caractéristiques. L'influence de tous ces éléments extérieurs doit évidemment être prise en compte lors de la création d'un système. L'étude de l'influence de ces éléments est un sujet trop large pour être traité intégralement ici. Nous nous sommes limités à trois caractéristiques qui sont en lien direct avec une éventuelle intégration dans un produit tel qu'une jumelle infrarouge multifonction. Leur étude a consisté en une analyse de la littérature associée à des d'expérimentations.

La première expérimentation concerne l'influence de l'utilisation d'un menu circulaire sur les différentes modalités d'interaction. Le but est de déterminer dans quelle mesure le suivi du regard comme modalité d'interaction peut être intégrée à des interfaces déjà existantes.

La seconde expérimentation teste l'adéquation de telles modalités avec une tâche de réponse à des événements visuels attendus, demandant une concentration visuelle importante.

La troisième expérimentation teste l'adéquation avec une tâche de réponse à des événements visuels inattendus. Il s'agit d'évaluer les capacités de perception d'alertes visuelles en vision périphérique lors d'interactions au regard. Le but est de vérifier que les modalités d'interaction au regard ne compromettent pas la diffusion d'informations critiques à l'utilisateur.

### 4.1 Intégration de l'interaction au regard sur des menus linéaires et circulaires

Cette première section présente une expérimentation visant à comparer l'adéquation de plusieurs modalités d'interaction au regard sur un menu linéaire et sur un menu circulaire. Dans un premier temps, nous expliquons pourquoi nous avons étudié cette caractéristique. Nous présentons ensuite le design expérimental suivi des résultats et d'une discussion.

### 4.1.1 Besoin opérationnel et état de l'art

Parmi l'ensemble des caractéristiques du modèle présenté dans le Chapitre 3, il nous a semblé pertinent d'étudier la forme des widgets composant les menus d'une interface graphique. La raison est que lors de l'intégration de technologies comme le suivi du regard pour l'interaction dans des systèmes optroniques, la conception de l'interface n'est pas focalisée sur la modalité d'interaction. Si on intègre cette modalité d'interaction à des systèmes existants, ceux-ci possèdent par exemple déjà des interfaces graphiques auxquelles les utilisateurs sont habitués. Même lors de la conception d'un nouveau système, il est possible que celui-ci s'intègre à un ensemble de systèmes désireux de conserver une homogénéité entre leurs interfaces graphiques. Pour cela, il n'est pas toujours possible de modifier les interfaces intégrées dans les systèmes. Il est alors nécessaire d'adapter la modalité d'interaction à l'interface, et non l'inverse.

Parmi l'ensemble des modalités d'interaction au regard, un des enjeux est alors de déterminer quelle est la modalité la mieux adaptée à un système déjà conçu et mature possédant ses propres interfaces. Etudier le comportement des différentes modalités d'interaction au regard sur des menus-types semblables à ceux utilisés actuellement permettrait de déterminer quelle est la manière la plus adaptée d'intégrer le suivi du regard à des systèmes existants. Il est possible dans un premier temps que le suivi du regard cohabite avec les modalités d'interaction actuelles pour permettre une intégration facilitée pour les utilisateurs.

Lors de l'utilisation de la souris comme modalité d'interaction, il a été montré que les menus circulaires comme les pie-menus correspondent mieux à certaines tâches utilisateur (Callahan, Hopkins, Weiser, & Shneiderman, 1988; Samp & Decker, 2010) et présentent de meilleures performances. Par exemple, le placement équidistant de tous les items à la même distance du centre d'un cercle optimise la distance parcourue par le regard. Si tous les items ont la même importance, ce type de menu est très adapté. Dans les *marking-menu* (Kurtenbach G. P., 1993), la sélection d'un item peut même être considérée comme une interaction gestuelle représentant la somme des directions empruntées lors de la navigation dans les menus. Cela permet de se dispenser de la représentation graphique.

Des études ont été réalisées avec des modalités d'interaction au regard pour la sélection dans des menus. Urbina et ses collègues (Urbina, Lorenz, & Huckauf, 2010) comparent deux modalités d'interaction au regard, le *DwellTime* et une méthode appelée *sélection border*, équivalent des *marking-menus* au regard, mais conclut sur l'avantage du *DwellTime*. Kammerer (Kammerer, Scheiter, & Beinhauer, 2008) compare des modalités d'interaction au regard sur des menus linaires, circulaires et semi-circulaires et conclut sur une meilleure utilisabilité des deux derniers.

L'expérience présentée ici vise à comparer les performances des différentes modalités d'interaction sur un menu linéaire et un menu circulaire disposant par ailleurs de caractéristiques similaires comme la taille des items et la distance inter-items.

## 4.1.2 Méthode et design expérimental

### *Participants*

L'expérience a été réalisée avec 14 participants âgés de 24 à 43 ans. Cinq participants avaient déjà utilisé un système de suivi du regard pour interagir avec des menus linéaires lors d'une expérience préliminaire comportant une tâche différente. Tous les participants disposaient d'une vision normale ou totalement compensée. Tous les participants étaient habitués à l'utilisation d'un ordinateur dans leur travail ou dans leur temps libre.

### *Design des menus linéaire et circulaire*

Deux menus sont proposés : un menu linéaire et un menu circulaire (cf. Figure 32).

*Menu linéaire* : Le menu linéaire est composé d'une suite horizontale d'items carrés avec une séparation entre chaque item. L'ouverture d'un sous-menu se fait au-dessus du menu principal. Le sous-menu est centré horizontalement par rapport à l'item ouvert. Deux lignes partent de l'item du menu principal afin de simplifier la visualisation du sous-menu effectivement ouvert. Le choix d'ouverture vers le haut est arbitraire, il rappelle la présentation des applications ouvertes sur un environnement Windows par exemple.

*Menu circulaire* : Le menu principal circulaire est composé de quatre items disposés dans des orientations N, S, E, O. Les sous-menus sont dans une configuration NE, NO, SO, SE. Le centre des sous-menus se situe dans la continuité de l'orientation de l'item sélectionné, de telle sorte qu'un recouvrement de l'item principal s'effectue lors de leur ouverture, comme dans un *marking-menu*.

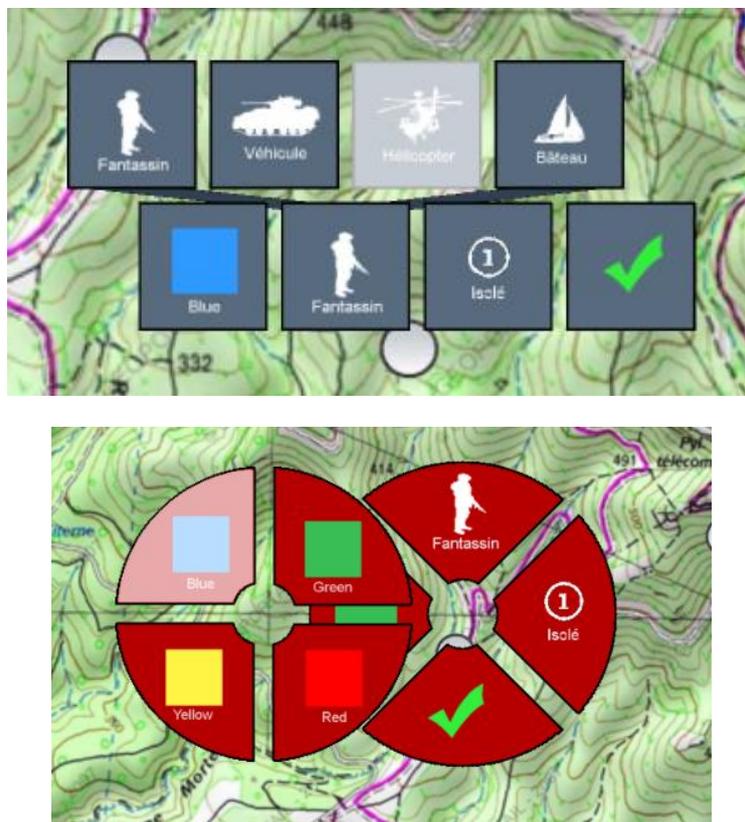


Figure 32. Menu linéaire (gauche) et menu circulaire (droite).

L’ensemble des sous-menus est présenté dans l’annexe B.

Les menus sont conçus avec des items de tailles semblables (cf. Figure 33), approximativement égale à  $1,5^\circ$  d’angle, afin d’assurer l’utilisabilité du regard comme médium d’interaction indépendamment de la distance des yeux de l’utilisateur à l’écran. Choisir des éléments de tailles supérieures à  $1^\circ$  d’angle est une des bonnes pratiques de l’interaction au regard. Cela permet de contrebalancer la précision d’ $1^\circ$  due aux micromouvements de l’œil (Jacob, 1993). L’écartement minimum entre deux items adjacents est identique et égal à environ  $0,2^\circ$  d’angle. Chaque caractéristique est représentée par une icône associée à la valeur textuelle de la caractéristique. La fermeture d’un sous-menu se fait lors de la sélection d’un sous-item, dès lors que l’utilisateur ne regarde pas le sous-menu ouvert pendant 0.5s ou lorsqu’un autre sous-menu est ouvert.

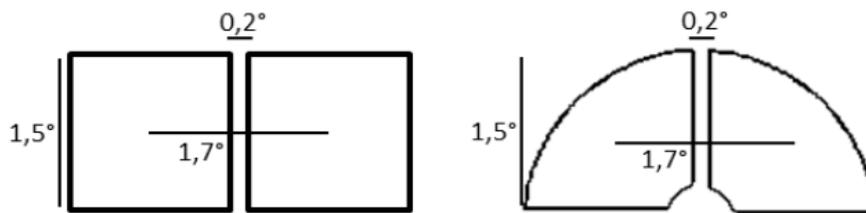


Figure 33. Dimensions des éléments des menus linéaire et circulaire.

## Modalités d’interaction

L’utilisabilité de quatre modalités différentes d’interaction a été testée dans cette expérimentation. Les modalités étudiées sont les méthodes *DwellTime*, *DwellTime* délocalisé, ouverture instantanée et une méthode multimodale regard + bouton. A part le *DwellTime* délocalisé, ces modalités d’interaction ont toutes été décrites dans la section 2.5. Nous précisons ci-dessous le design de ces différentes modalités d’interaction pour notre expérimentation :

*DwellTime* (DT) : Lors de l’interaction *DwellTime* un feedback visuel est présenté à l’utilisateur. Une image semi-transparente dont la forme dépend du type de menu apparait au centre de l’item et grossit linéairement en fonction du temps de telle sorte que l’item soit totalement recouvert lorsque le temps paramétré est atteint. L’item est alors activé (cf. Figure 34).

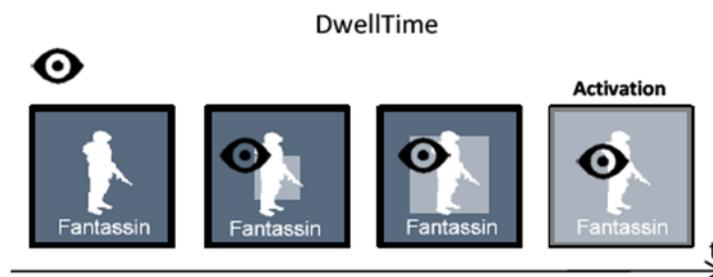


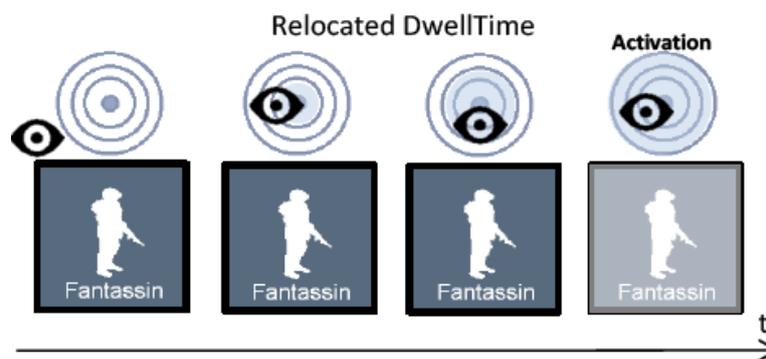
Figure 34. Interaction *DwellTime* et évolution du feedback visuel au cours du temps lorsque l’item est fixé.

Le temps de fixation cumulé sur un item n’est remis à zéro que si ce dernier n’est pas regardé pendant 0,4s, et cela pour deux raisons : la première est de permettre à l’utilisateur d’effectuer un *square wave jerk*, c’est-à-dire un aller-retour rapide sur une autre partie de l’écran sans pénaliser l’interaction (Leigh & Zee, 2015) ; la seconde est d’éviter la remise à zéro du temps d’attente si la

#### 4.1 - Intégration de l'interaction au regard sur des menus linéaires et circulaires

position du regard est détectée hors de l'item pendant un temps très court qui peut être dû à une erreur de précision du système ou aux micromouvements de l'œil. Le temps d'attente est paramétré individuellement entre 300 ms et 1000 ms par chaque utilisateur lors de la phase d'apprentissage à l'aide d'une procédure détaillée plus bas.

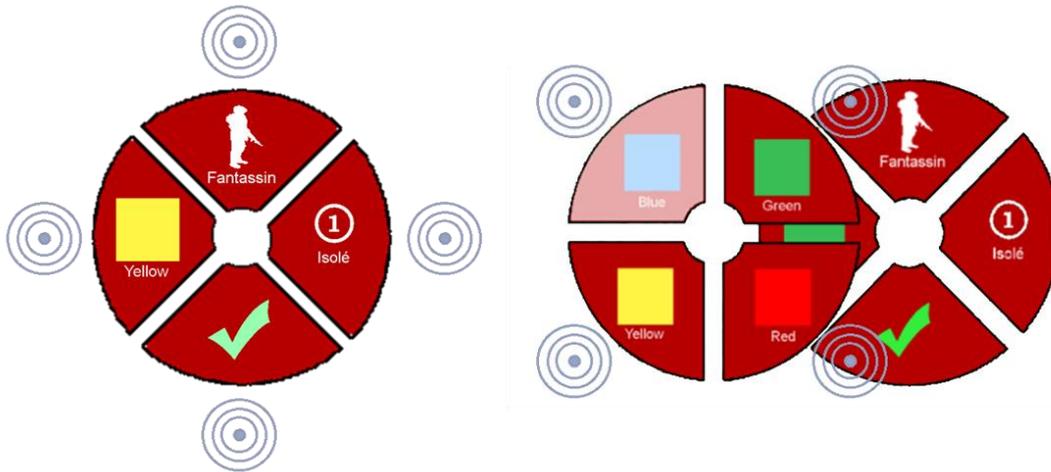
*DwellTime délocalisé* (RD pour *relocated DwellTime* en anglais) : Parmi les modalités d'interaction de la littérature que nous avons étudiées, nous avons décidé d'ajouter une nouvelle modalité d'interaction. Cette modalité d'interaction consiste en une interaction *DwellTime* sur une cible située à côté de l'item à activer. Donc, au lieu de fixer directement l'item à sélectionner, l'utilisateur signifie sa volonté d'activer l'item en fixant la cible liée à cet item. Sur la Figure 35, l'item à activer est le bouton « Fantassin ». Une fixation sur la cible au-dessus a pour effet d'activer cet item. La cible est toujours visible, contrairement à la modalité présentée par Tall (Tall, 2008) où la cible apparaît lorsque l'item est regardé.



**Figure 35. Interaction *DwellTime* délocalisé et feedback visuel. L'élément à fixer est une cible disposée à côté de l'item à activer.**

Cette modalité permet à l'utilisateur d'analyser le menu pendant autant de temps qu'il le souhaite sans risquer d'activer les différents éléments comme cela pourrait être le cas lors d'une interaction *DwellTime*. Le coût de cet avantage par rapport au *DwellTime* est un chemin visuel plus complexe. En effet, une fois l'item à activer trouvé, la modalité nécessite encore une saccade de la part de l'utilisateur. Cependant, nous pensons qu'un chemin expert est possible si, lorsqu'il connaît bien l'interface, l'utilisateur n'a pas besoin de regarder l'item à activer et peut directement se diriger vers la cible. Dans ce cas, la complexité du chemin est la même que pour la modalité *DwellTime*.

La cible à fixer est constituée de cercles concentriques afin de faciliter l'accrochage du regard sur l'item. Pour le menu linéaire, ces cibles sont situées directement au-dessus de chaque item. Pour le menu circulaire, elles sont situées à l'extérieur du cercle, dans l'axe de l'item (cf. Figure 36).



**Figure 36. Positionnement des cibles lors de la modalité d’interaction DwellTime délocalisé sur un menu circulaire.**

Lors de l’ouverture d’un sous-menu, les cibles déportées du menu principal disparaissent, ce qui ne permet plus de fermer un sous-menu en en ouvrant un autre. Le feedback visuel lié à cette méthode d’interaction est le même que le feedback de la méthode précédente. Comme pour le *DwellTime*, le temps est paramétré par l’utilisateur lors d’une phase d’apprentissage.

*Multimodal bouton-regard* (Bouton) : Tant que l’item est regardé, il est mis en surbrillance. L’utilisateur peut alors appuyer sur la touche espace pour sélectionner cet item. Afin de limiter l’impact dû à des problèmes de synchronisation œil-main, l’item peut encore être activé en appuyant sur la touche pendant 0.2s si le regard de l’utilisateur ne fixe pas un autre item du menu.

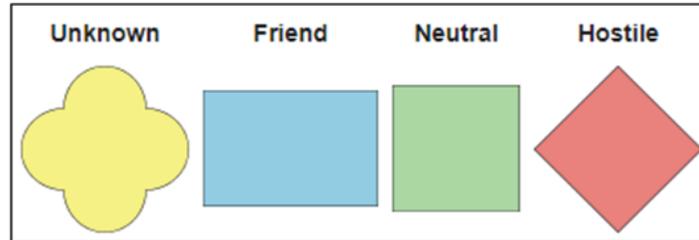
*Ouverture instantanée* (Instant) : L’activation instantanée est plutôt rejetée dans la littérature car elle est à l’origine des effets regroupés sous le nom de *Midas Touch*. Cependant, lorsque le coût de correction d’une erreur de manipulation est faible (comme lors de l’ouverture d’un mauvais menu, où il suffit d’ouvrir un autre menu pour corriger l’erreur), cette modalité d’interaction paraît tout de même pertinente. De fait, cette modalité est uniquement utilisée pour l’ouverture des sous-menus et est combinée à la méthode précédente pour la sélection de l’item du sous-menu. Dans la pratique, un temps seuil de 20 ms a été implémenté afin d’éviter l’ouverture du menu lors du passage du regard sur un item pendant une saccade. Aucun feedback visuel spécifique n’est appliqué à cette interaction, le principal indice attestant du succès de l’activation est l’ouverture du menu.

### ***Tâche expérimentale***

La tâche expérimentale est représentative d’une tâche opérationnelle réellement effectuée dans le domaine militaire appelée caractérisation de points. Comme le théâtre des opérations militaires est en constante évolution, il est nécessaire pour les utilisateurs terrain de pouvoir caractériser les éléments observés. Cette tâche consiste à spécifier un certain nombre de caractéristiques d’une unité, d’un bâtiment, de véhicules, alliés ou ennemis afin de faciliter le partage et la visualisation d’information. Nous avons utilisé le standard APP-6A, présenté par Kourkolis (Kourkolis, 1986), qui est le standard OTAN de représentation cartographique des unités. Trois dimensions sont alors à caractériser :

#### 4.1 - Intégration de l'interaction au regard sur des menus linéaires et circulaires

- L'affiliation, c'est la relation entre l'unité caractérisée et l'utilisateur caractérisant. Elle a classiquement l'une des valeurs suivantes : inconnue (par défaut), alliée, neutre ou hostile. En réalité elle peut prendre 14 valeurs différentes. Dans le standard APP-6A, cela est représenté par la forme et la couleur (cf. Figure 37).



**Figure 37. Affiliations classiques du standard APP-6A.**

- Le type d'unité, c'est la catégorie disposant du plus grand nombre de classes. Elle donne des précisions sur la nature de l'unité caractérisée. Dans le standard APP-6A, cela est représenté par un ou plusieurs symboles à l'intérieur du cadre donné par l'affiliation (e.g. Figure 38). Le nombre de symboles est assez conséquent (plus de 200) mais peut être réduit en fonction du contexte opérationnel. On a par exemple peu de chances de trouver des unités maritimes lors d'une opération en montagne. Cela réduit drastiquement le nombre de type d'unités observable dans un contexte opérationnel donné.

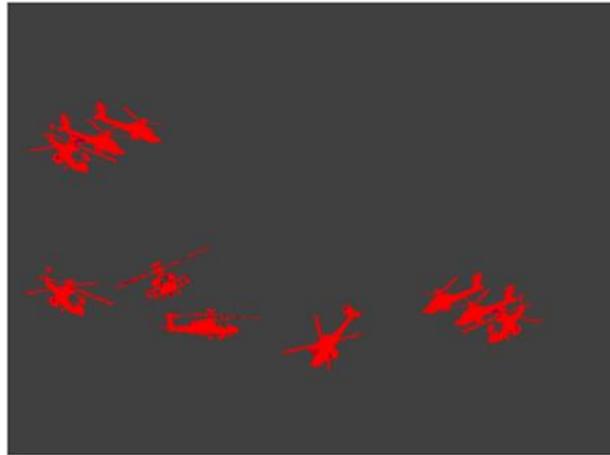


**Figure 38. Exemples de représentation de type d'unité dans le standard APP-6A.**

- La taille de l'unité, c'est le nombre d'unités, plus précisément l'ordre de grandeur de l'unité. Elle peut prendre 13 valeurs allant de l'équipe (environ 2 personnes) à la région militaire (environ 250.000 individus). De même que pour le type d'unité, les tailles d'unité rencontrés lors des missions sont plutôt de l'ordre de 4-5. Dans le standard APP-6A, cela est représenté par des symboles au-dessus du cadre de l'affiliation.

La caractérisation d'un point est bien adaptée à l'utilisation de menus contextuels.

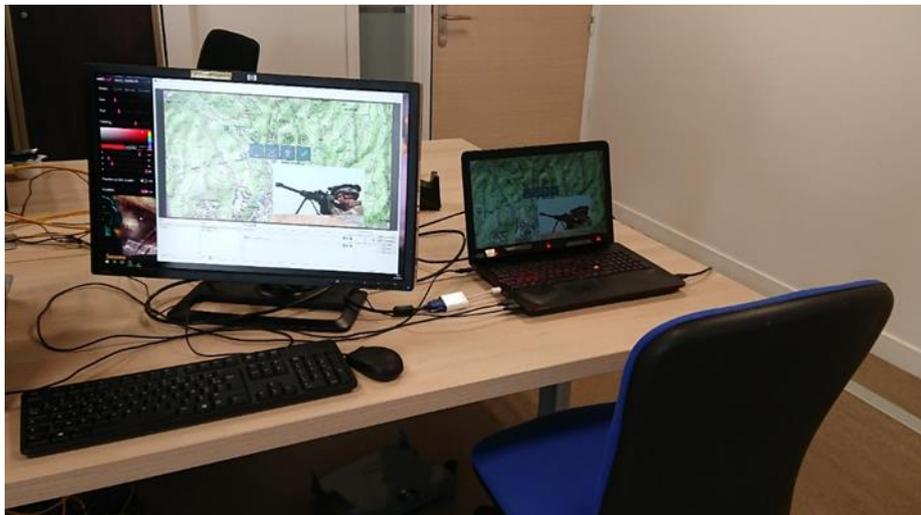
La tâche utilisateur consiste en la sélection de trois caractéristiques simplifiées pour un public non-militaire, à savoir la couleur, le type d'unité, et le nombre approximatif d'individus. Une image était présentée à l'utilisateur représentant un certain nombre d'unités de couleur (cf. Figure 39). Et un des deux menus contextuels présentés s'ouvrait au milieu de l'écran. L'utilisateur devait alors sélectionner les caractéristiques présentes sur l'image à l'aide de trois sous-menus disposant chacun de quatre items puis valider son choix à l'aide de l'item « valider ». Lorsqu'une caractéristique est sélectionnée, l'item du menu principal prend cette valeur afin de permettre une vérification à tout moment des caractéristiques actuellement sélectionnées.



*Figure 39. Exemple d'une image présentée à l'utilisateur pour être caractérisée. Pour cette image, il fallait spécifier « Rouge, Hélicoptère, 10 ou + ».*

### ***Dispositif expérimental***

L'étude a été réalisée dans un open-space très calme avec peu de passage, faute de pouvoir utiliser une salle totalement isolée. L'affichage se faisait sur un écran mat 15 pouces d'ordinateur portable avec une résolution de 1920x1080. Ni l'utilisateur, ni l'écran, n'étaient affectés par des reflets ou un excès de luminosité. Un Tobii Eye-X et le clavier de l'ordinateur portable ont été utilisés pour interagir. Un deuxième écran, hors du champ de vision de l'utilisateur permettait à l'opérateur de visualiser les mouvements du regard sur l'écran à l'aide d'un logiciel de streaming (« OBS ») ainsi que du « Streaming gaze overlay » fourni par Tobii permettant d'afficher la trace du passage du regard en temps réel (cf. Figure 40). Les retours verbaux utilisateurs ont été enregistrés à l'aide d'une petite caméra Panasonic.



*Figure 40. Dispositif expérimental. L'opérateur peut voir les mouvements du regard de l'utilisateur sur l'écran de gauche. L'utilisateur, assis sur la chaise, ne voit pas ce qui se passe sur cet écran.*

### ***Design expérimental et procédure***

Tout d'abord, un questionnaire a été présenté aux participants afin de collecter des données individuelles. Ces données contenaient l'âge, le genre, la qualité de leur vue, la fréquence d'utilisation de systèmes informatiques, l'utilisation préalable de systèmes de suivi du regard, la fréquence de la pratique de sport, de jeux vidéo. Le formulaire à remplir afin de recueillir leurs retours sur leur expérience a aussi été présenté à ce moment afin d'éviter qu'ils ne découvrent les questions qu'après coup.

Ensuite, il était demandé à l'utilisateur de réaliser la calibration du système de suivi du regard. La tâche à effectuer était présentée lors de l'apprentissage de la première modalité d'interaction.

L'expérience a suivi un modèle 4x2, c'est-à-dire composé de deux facteurs indépendants avec respectivement 4 et 2 valeurs. La première variable indépendante était la modalité d'interaction pouvant prendre comme valeur une des modalités d'interaction présentées précédemment {*DwellTime*, *DwellTime* délocalisé, Instant, Button} alors que la seconde variable dépendante représentait le type de menu et pouvait prendre pour valeur Linéaire ou Circulaire. Chaque participant a passé l'ensemble de 8 tests dont l'ordre a été rendu aléatoire de la manière suivante : en premier lieu, l'ordre des modalités était tiré au hasard puis pour chaque modalité, le premier type de menu auquel l'utilisateur était confronté était choisi aléatoirement. Aucun ordre de passage n'a été identique.

Chacun des 8 tests comportait une phase d'apprentissage de l'utilisation de la modalité d'interaction sur la tâche expérimentale pendant 1 minute. A la fin de ce temps, si l'utilisateur était capable d'effectuer correctement deux fois la tâche à la suite, le test commençait, sinon il disposait d'une minute supplémentaire pour s'habituer à la modalité d'interaction et au type de menu. La première phase d'apprentissage était plus longue étant donné que la tâche expérimentale était expliquée à ce moment. Le test consistait ensuite en 12 réalisations de la tâche expérimentale à la suite, quel que soit les résultats obtenus par l'utilisateur. Tous les 2 tests, c'est-à-dire à la fin des tests sur une modalité, il était demandé à l'utilisateur de verbaliser ses ressentis vis-à-vis de la modalité puis de remplir un questionnaire permettant de récupérer des valeurs subjectives sur l'intuitivité, la rapidité perçue, l'effort nécessaire et la fiabilité de la modalité. Une note sur 20, représentative de l'appréciation générale, était aussi demandée pour chacun des menus avec chaque modalité d'interaction. Aucune consigne sur la stratégie de réalisation de la tâche n'était fournie aux participants. Ils pouvaient spécifier les différentes caractéristiques demandées par la tâche expérimentale dans l'ordre qu'ils souhaitaient. Il leur était demandé seulement d'essayer d'être « efficace », l'interprétation de ce terme étant laissée à leur seule interprétation.

La Figure 41 résume la chronologie de l'expérimentation.

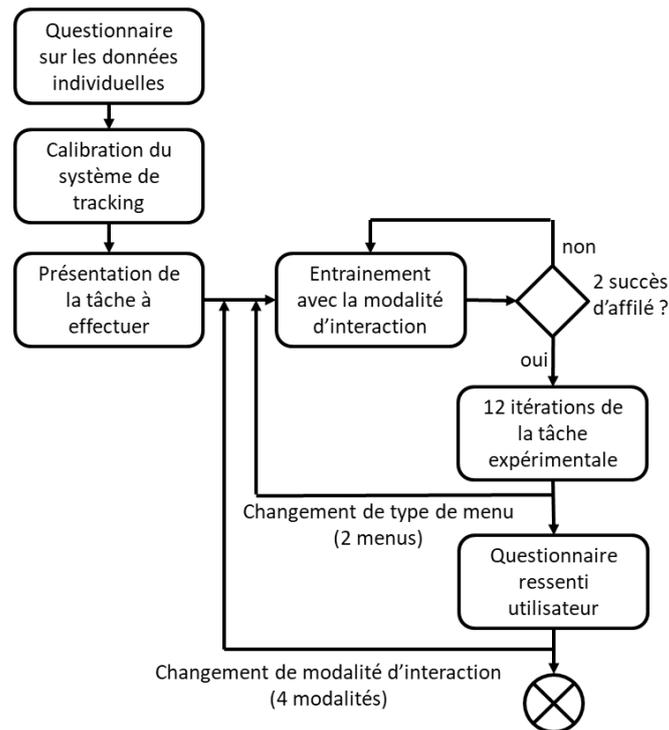


Figure 41. Diagramme représentant la chronologie de l’expérimentation 1.

### Variables mesurées

Comme il a été mentionné précédemment, il a été demandé aux participants de verbaliser leurs impressions à la fin de l’utilisation de chaque modalité d’interaction. Ces retours ont permis de comprendre le ressenti des utilisateurs vis-à-vis de chaque modalité d’interaction ainsi que leur adéquation à chaque type de menu. Dans un premier temps, les participants étaient libres de donner leurs retours, puis ils étaient éventuellement encouragés par l’évaluateur.

Un formulaire (présenté en annexe A) a permis la collecte de données qualitatives telles que l’intuitivité, la rapidité perçue, la fiabilité et l’effort demandé. Les utilisateurs notaient les différentes caractéristiques sur une échelle de Likert à 7 niveaux où 1 signifie, par exemple, pas du tout intuitif et 7 signifie très intuitif. Le choix d’une échelle à 7 niveaux s’est fait en accord avec les travaux de Symonds (Symonds 1924). Pour chaque couple <modalité ; type de menu>, une note globale sur 20 était demandée. Les données sont présentées dans la section suivante au travers de leur moyenne et leur écarts type.

D’un point de vue quantitatif, seul le temps passé à regarder l’interface par activation a été analysé. Nous supposons que ce temps est représentatif de l’efficacité des modalités, sans tenir compte de la complexité des différentes images présentées aux utilisateurs.

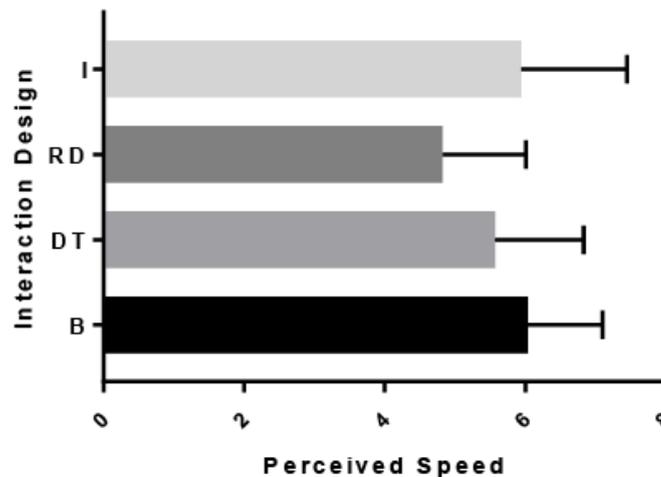
Les résultats subjectifs ont été analysés en utilisant des tests non-paramétriques. Dans un premier temps, les résultats ont été analysés par paires, en faisant varier seulement la dimension « Menu », en utilisant le test du rang signé de Wilcoxon. Le but de cette analyse était de conclure sur un effet d’interaction du type de menu sur les performances des modalités d’interaction. Dans un second temps, seule la dimension « Modalité » a été étudiée en utilisant le test de Friedman pour chercher un effet principal de la modalité sur les performances obtenues. Enfin, une analyse post hoc a été effectuée en utilisant le test de Dunn. Les données quantitatives ont été analysées à l’aide d’une

ANOVA pour mesures répétées à deux facteurs. Les ANOVAS ont été appliquées sur le logarithme des temps, cependant les résultats sont présentés avec des échelles linéaires pour plus de clarté.

### 4.1.3 Résultats

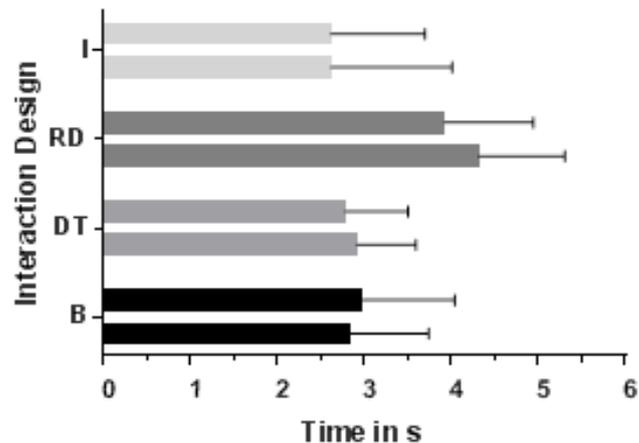
#### *Vitesse d'activation et vitesse perçue.*

La Figure 42 présente la rapidité des modalités perçue par les utilisateurs. Aucun des tests du rang de Wilcoxon n'a montré un effet du type de menu sur la rapidité perçue par les utilisateurs (tous les  $p > 0.05$ ). Aucune des modalités n'est considérée plus rapide sur un type de menu que sur l'autre de la part des utilisateurs. La différence de performance entre les différentes modalités a alors été étudiée en fusionnant les données obtenues sur les deux types de menus. Le test de Friedman a montré un effet significatif de la modalité d'interaction utilisée sur la rapidité perçue de l'interaction. ( $p < 0.001$ ). Les analyses post hoc montrent que le *DwellTime* délocalisé est significativement différent de chacune des autres modalités. Cette différence est particulièrement flagrante par rapport aux modalités Bouton et Instantanée ( $p < 0.001$ ) alors que c'est moins évident par rapport à l'interaction *DwellTime* ( $p=0.02$ ). Le *DwellTime* délocalisé obtient une note moyenne plus basse de 1.2 points (0.8 points par rapport au *DwellTime*). Les utilisateurs ont indiqué que rester concentré à fixer un seul point semblait long et qu'ils avaient l'impression d'attendre que l'interface réponde. Cela leur a semblé d'autant plus long avec le *DwellTime* délocalisé qu'il y a une étape supplémentaire.



**Figure 42. Moyenne et écart type des notes données à la rapidité perçue des modalités d'interaction.**

La Figure 43 présente les temps d'exécution (moyenne + ET) obtenu lors des tests. Lors de l'analyse des temps d'activation effectifs, le test RM-ANOVA a montré un effet d'interaction du type de menu sur les performances des modalités d'interaction. Le *DwellTime* sur un menu linéaire est significativement plus rapide de 0.5s ( $p < 0.05$ ). Cela représente environ un huitième du temps d'activation de cette modalité. Mais ce changement n'a pas été perçu par les utilisateurs. Il semble aussi que la modalité *DwellTime* délocalisé est significativement plus lente que les trois autres modalités d'environ 1.3s (3 valeurs de  $p < 0.001$ ). Cela est cohérent avec la rapidité perçue de la part des utilisateurs. Contrairement à nos attentes, les modalités Bouton et Instantanée ne sont pas plus rapides que l'interaction *DwellTime*.



**Figure 43. Temps d’activation : temps passé à regarder l’interface jusqu’à ce qu’une caractéristique soit sélectionnée par l’utilisateur. Pour chaque modalité, la barre du haut représente un menu linéaire, celle du bas un menu circulaire.**

### ***Intuitivité***

La Figure 44 présente l’intuitivité des modalités d’interaction perçue par les utilisateurs. Aucun des tests du rang de Wilcoxon n’a montré un effet du type de menu sur l’intuitivité perçue par les utilisateurs (tous les  $p > 0.05$ ). Aucune des modalités n’est considérée plus intuitive sur un type de menu que sur un autre. En mettant la dimension « Menu » de côté, le test de Friedman a montré un effet significatif ( $p < 0.0001$ ) de la modalité sur l’intuitivité perçue, montrant que les modalités d’interaction au regard ne sont pas perçues de la même manière de la part des utilisateurs. Les modalités d’interaction Bouton et *DwellTime* ne montrent pas de différences significatives entre elles ( $p > 0.05$ ). Il en est de même pour les modalités d’ouverture instantanée et de *DwellTime* délocalisé. Cependant, ces 2 groupes (Bouton-*DwellTime* vs. Instantané-*DwellTime* délocalisé) montrent une différence significative entre eux (4 valeurs de  $p < 0.05$ ).

Contrairement à nos attentes, les menus circulaires ne semblent pas plus intuitifs que des menus linéaires dans le cadre de notre expérimentation. Six utilisateurs ont trouvé inconfortable de devoir regarder un élément non-sémantique lors de l’utilisation du *DwellTime* délocalisé. Ils ont tous mentionné qu’ils préféreraient (ou préféreraient dans le cas où ils n’avaient pas encore utilisé la modalité d’interaction *DwellTime*) interagir directement avec l’item. L’ouverture instantanée a gêné certains utilisateurs, surtout à cause de la position centrale du menu qui impliquait un survol régulier de ce dernier. Cela a montré une fois de plus l’effet du *Midas Touch*.

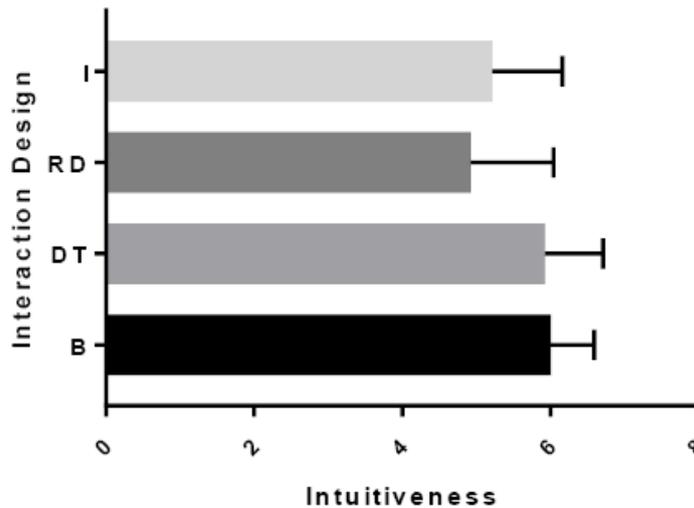


Figure 44. Moyenne et écart type des notes données à l'intuitivité des modalités d'interaction.

### Fiabilité

La Figure 45 présente la fiabilité des modalités d'interaction perçue par les utilisateurs. Contrairement aux autres critères étudiés, le test de Wilcoxon montre un effet significatif du type de menu sur la fiabilité perçue ( $p < 0.01$ ) de la modalité *DwellTime*. Cela n'est pas le cas pour les trois autres modalités ( $p > 0.05$ ). Une explication de ce phénomène a été rapidement mise en avant. En effet, le menu circulaire a été conçu dans l'expérience de telle façon que le sous-menu vient se superposer au menu principal. Cela implique que le temps d'attente de la modalité *DwellTime* sur l'un des éléments du sous menu va commencer dès lors que le menu s'ouvre, avant même que l'utilisateur ne puisse prendre connaissance de cette ouverture. Cela n'a pas d'impact sur les autres modalités d'interaction puisque celles-ci demandent soit une confirmation extérieure (Bouton), soit une fixation sur un élément que le regard ne fixait pas initialement. Ce résultat semble être lié plus à un problème de conception qu'à un véritable impact du type de menu.

En tenant compte de ce résultat, les tests suivants ont été réalisés seulement en tenant compte des données obtenues avec la version du menu linéaire pour la modalité d'interaction *DwellTime*. Les tests de Friedman et de Dunn n'ont alors montré aucun effet significatif de la modalité sur la fiabilité perçue. Notre hypothèse était que les modalités Bouton et *DwellTime* délocalisé donneraient un sentiment de fiabilité plus important grâce à une étape d'activation plus explicite. Les résultats ne reflètent pas ses hypothèses et toutes les modalités ont été perçues comme aussi fiables, mis à part le point mentionné précédemment.

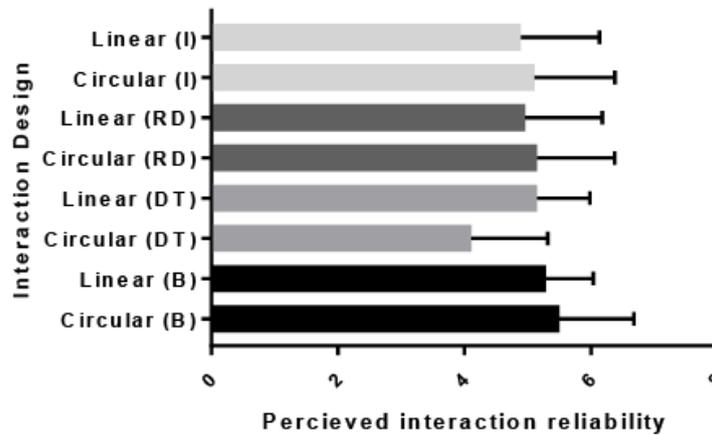
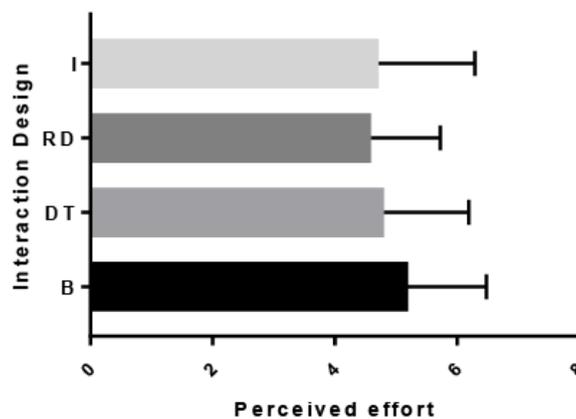


Figure 45. Moyenne et écart type des notes données à la fiabilité des modalités d’interaction.

### Effort demandé

La Figure 46 présente l’effort perçu par les utilisateurs lors de l’utilisation des modalités d’interaction. De même que pour l’intuitivité, les tests de Wilcoxon n’ont pas montré d’impact significatif du type de menu sur l’effort perçu (toutes les valeurs de  $p > 0.05$ ). La modalité utilisée n’a pas montré d’effet principal sur l’effort perçu lors de son utilisation. De plus, pratiquement toutes les comparaisons par paires n’ont montré aucun effet significatif à l’exception de la comparaison entre la modalité Bouton et la modalité *DwellTime* délocalisé qui est significative ( $p = 0.05$ ) avec une différence de 0.6 points à l’avantage de la modalité Bouton.

Cela a été verbalisé par les utilisateurs qui regrettaient de devoir effectuer une étape supplémentaire avec le *DwellTime* délocalisé. Fixer un item non sémantique pour en activer un autre demandait un certain effort. Pour seulement 3 participants, la modalité instantanée a été perçue comme difficile à utiliser. Ces utilisateurs ne réussissaient pas à s’abstraire de l’ouverture d’un sous-menu et ressentaient le besoin de regarder à l’intérieur d’un sous-menu ouvert par inadvertance. Cela leur a en particulier rendu la tâche de vérification de la sélection d’une caractéristique difficile par exemple. Cela a été encore plus verbalisé lors de l’utilisation de menus linéaires puisque l’identification du menu effectivement ouvert était plus difficile étant donné que tous les menus s’ouvraient à peu près au même endroit.



*Figure 46. Moyenne et écart type des notes données à l'effort demandé pour l'utilisation des différentes modalités d'interaction.*

#### 4.1.4 Discussion

Contrairement à ce qui était attendu, nous n'avons observé aucun impact du type de menu sur les préférences des utilisateurs ou sur les performances des modalités d'interaction. Le seul effet significatif dû au type de menu a été observé sur la modalité *DwellTime*. La diminution de la fiabilité ressentie par les utilisateurs est cependant très probablement dû à la superposition des sous menus au-dessus du menu principal. Comme le regard est prépositionné au-dessus d'un élément lors de l'ouverture du sous-menu, l'impact est le même que lors du choix d'un temps de *DwellTime* trop court, il engendre des activations intempestives.

Cependant, il est important de rappeler que la conception des menus a été effectuée afin qu'ils disposent de caractéristiques similaires, comme la taille ou la distance inter-items. En pratique, les menus circulaires possèdent des caractéristiques qui n'ont pas été exploités ici. La meilleure répartition spatiale peut par exemple permettre de réduire la taille des items, ou des modalités d'interaction pourraient être développées pour ce type de menu particulier. Cela peut mener à des propositions de modalités d'interaction telles que celles proposées par Kammerer (Kammerer, Scheiter, & Beinhauer, 2008). Si nous n'avons pas choisi des menus circulaires plus classiques, c'est que notre intérêt était de comparer des modalités d'interaction au regard sur des menus types, déjà utilisés dans les produits optroniques Safran. Ainsi, les résultats obtenus sont directement applicables aux menus tels qu'ils sont actuellement conçus dans les produits, mais ne sont pas généralisables. Il serait aussi important de faire varier le nombre d'items ou la profondeur des menus pour obtenir des informations supplémentaires sur le comportement des modalités dans les menus.

Un autre résultat intéressant a été la disparité dans les appréciations des utilisateurs. De manière générale, l'interaction multimodale ajoutant un bouton à l'utilisation du regard et l'interaction *DwellTime* ont été autant appréciées. Mais des ressentis très différents ont été observés chez les utilisateurs, certains aiment le *DwellTime* et n'aimaient pas l'interaction multimodale, et inversement. Ces différences n'ont pas été expliquées par les données personnelles recueillies comme l'âge, l'habitude des jeux-vidéo ou de l'utilisation de technologies ou par quelque différence de caractéristiques visuelles.

Pendant l'expérience pilote, l'ouverture instantanée s'était révélée prometteuse et semblait être rapide et facile d'utilisation (les menus étaient alors disposés en bas de l'écran à ce moment-là). Ici, elle n'a pas montré de tels résultats et n'est notamment pas plus rapide que l'interaction multimodale, ce qui prouve une fois encore les effets du *Midas Touch*. De par son positionnement central, les ouvertures intempestives sont plus fréquentes et donc plus gênantes. Nous pensons que la structure en étoile du positionnement sous-menu circulaires rendraient l'ouverture instantanée plus incompréhensible que sur les menus circulaires, ce qui n'a cependant pas été le cas.

La modalité *DwellTime* délocalisé a été conçue pour donner plus de contrôle à l'utilisateur lors de l'interaction. Cependant, la modalité n'a pas été perçue comme telle. Le contrôle ajouté à la modalité a été contrebalancé par la complexité d'utilisation. Il est cependant probable que cette modalité aurait un intérêt pour une tâche demandant une analyse plus profonde du menu, comme lors de l'apprentissage d'un menu plus complexe.

L'expérimentation a été conduite avec des utilisateurs novices. La tâche expérimentale était suffisamment simple pour pouvoir considérer l'utilisateur comme expert dans cette tâche, mais il n'était pas expert dans l'interaction au regard. Les utilisateurs militaires quant à eux, se forment de manière intensive à l'utilisation de systèmes, ils seraient alors considérés experts dans l'interaction au regard.

Pour une tâche telle que la caractérisation de point, faisant appel à des menus contextuels, l'interaction multimodale peut sembler très légèrement plus adaptée que les autres, apportant rapidité et contrôle aux utilisateurs pour un coût en complexité dû à la synchronisation œil-main assez faible. Il semble que pour les modalités testées, la forme du menu n'a eu que peu d'impact sur les performances des utilisateurs. Nous pensons ainsi que le type de menu doit être choisi en fonction d'autres critères comme la tâche à accomplir plutôt qu'en fonction de la modalité qui sera utilisée.

Résultats principaux de l'expérimentation :

- Aucun effet constaté du type de menu sur les performances des modalités testées
- La modalité *DwellTime* et la multimodalité Bouton-Regard présentent des performances similaires
- L'interaction *DwellTime* délocalisé est considérée comme trop complexe
- L'interaction instantanée est peu appréciée sur des menus positionnés au centre de l'interface graphique

## **4.2 Effet de la réorientation de l'attention visuelle sur l'interaction au regard**

Dans cette seconde section, nous commençons à nous intéresser à l'impact que peut avoir la tâche utilisateur sur les performances des différentes modalités d'interaction au regard. Plus particulièrement l'impact que peut avoir une tâche faisant appel à de l'attention visuelle divisée. Comme dans la section 1, nous commençons par expliquer pourquoi nous avons étudié cette caractéristique Puis nous présentons ensuite le design expérimental suivi des résultats et d'une discussion.

### **4.2.1 Besoin opérationnel**

Lors d'une interaction au regard, l'œil de l'utilisateur n'est plus seulement utilisé à des fins de perception, il joue aussi un rôle moteur. La modification de l'utilisation classique du regard vers une fonction sensorimotrice bouleverse son utilisation habituelle. Afin de pouvoir intégrer le regard comme modalité d'interaction, il est nécessaire de comprendre les implications de ce changement. Comme nous l'avons mentionné dans la section 3.3, il est possible que la tâche effectuée par l'utilisateur influe sur les performances de l'interaction au regard. De même, il est possible qu'interagir avec le regard perturbe certaines tâches en modifiant le comportement nominal de l'œil.

Lors de l'utilisation d'un système optronique tel que des jumelles ou une lunette de visée, une des tâches principales consiste en l'observation et la surveillance d'une zone. C'est une tâche nécessitant une attention visuelle forte de l'utilisateur. Il est nécessaire que l'utilisation du regard comme modalité d'interaction ne vienne pas perturber les performances de cette tâche. L'utilisateur doit pouvoir rester attentif aux modifications de la zone d'observation, et ce, même lors de l'interaction. Cette compétence correspond à une délocalisation, voire une division de l'attention visuelle.

Ce chapitre vise dans un premier temps à analyser le comportement des différentes modalités d'interaction déjà présentées lors d'un effort pour conserver son attention visuelle sur une scène. C'est-à-dire lorsque l'utilisateur interagit avec un menu en le regardant, tout en étant concentré visuellement sur une autre partie de son champ visuel. Dans un second temps, c'est la capacité effective à délocaliser son attention lors d'interaction au regard qui sera évaluée. La différence entre les deux concepts est la priorité mise tantôt sur la tâche d'interaction, tantôt sur la tâche d'observation nécessitant une attention visuelle particulière en région périphérique du champ visuel.

L'attention visuelle a été décrite dans la section 2.2.3, ce que nous visons à étudier dans ce chapitre est l'attention visuelle endogène, c'est-à-dire issue de la concentration de l'utilisateur. Notre étude se base notamment sur les travaux de Posner et ses collègues (Posner, Snyder, & Davidson, 1980) qui ont permis la création d'un protocole expérimental. L'attention visuelle exogène sera étudiée au chapitre suivant.

### **4.2.2 Méthode et design expérimental**

Deux expériences ont été menées en faisant varier la priorité donnée entre la tâche d'interaction et la tâche de fond. Cette sous-section présente ces deux expériences.

## ***Participants***

Cette expérience a été conduite avec 20 participants âgés de 21 à 46 ans. Les 20 participants disposaient tous soit d’une vision parfaite, soit d’une vision totalement compensée par des lunettes ou des lentilles de contact. Tous étaient des utilisateurs fréquents de systèmes informatiques, soit dans le milieu professionnel, soit pendant leur temps libre. Trois sujets (non inclus dans les 20) n’ont pas pu effectuer l’expérimentation car la calibration du système n’a pas pu être obtenue avec une qualité suffisante.

## ***Dispositif expérimental***

L’expérimentation a été conduite dans une salle totalement isolée où la luminosité et le bruit ambiant ont été contrôlés. Les utilisateurs faisaient face à un ordinateur portable 15 pouces avec une résolution de 1920x1080 pixels disposé à environ 50 cm d’eux. Le système de suivi du regard était un système Tobii Eye-X disposé sous l’écran. La touche « espace » de l’ordinateur était utilisée comme activateur pour l’interaction multimodale. Un second clavier permettait à l’opérateur de gérer le passage des différentes phases de l’expérimentation ainsi que d’appuyer sur une touche pour enregistrer le fait que l’utilisateur ait détecté un stimulus (cf. Tâche d’attention divisée n°2). Une souris était disponible lorsque nécessaire.

En dehors du champ de vision des participants, un second écran a été mis en place pour permettre à l’opérateur de visualiser les mouvements du regard des utilisateurs, notamment pour les phases d’apprentissage et pour la phase de calibration. Cela a été réalisé en utilisant le logiciel OBS Studio en association avec le *Streaming gaze overlay*<sup>7</sup> fourni par Tobii.

## ***Conception de l’interface***

Le type d’interface utilisée est le même que pour les menus linéaires de la première expérimentation. Les différences sont que le menu est disposé en bas de l’écran plutôt qu’au centre, et qu’une aide visuelle permet de mieux discerner quel menu est actuellement ouvert (cf. Figure 47).

## ***Modalité d’interaction***

Les modalités d’interaction testées dans cette expérience ont été le *DwellTime*, le *DwellTime* délocalisé, l’interaction multimodale regard + bouton. Ces modalités sont présentées dans la sous-section 4.1.2. Une interaction à la souris a aussi été effectuée pour servir de référence pour les comparaisons.

## ***Tâches expérimentales***

L’expérimentation est composée de trois tâches différentes : deux tâches de fond, appelées tâches d’attention délocalisées, et une tâche d’interaction avec un menu.

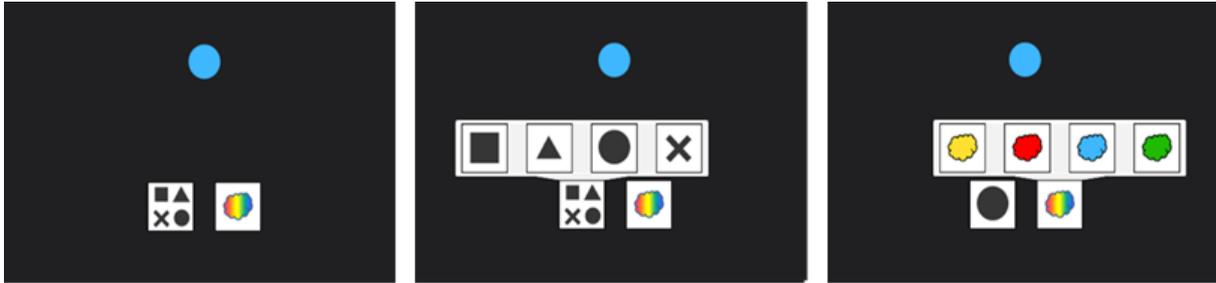
### Tâche d’interaction :

La tâche d’interaction est utilisée dans les deux parties de l’expérimentation. Elle consiste en la sélection de deux caractéristiques d’une figure simple présentée à l’utilisateur. Cette sélection se fait

---

<sup>7</sup> Logiciel fourni par Tobii permettant de visualiser la position du regard sur une vidéo enregistrée sans que celle-ci apparaissent à l’écran lors de l’utilisation du système de suivi du regard.

à l'aide de menus, contrôlés tour à tour par les différentes modalités d'interaction présentées. Une figure apparaît au-dessus des menus, disposant d'une forme et d'une couleur, chacune choisie aléatoirement parmi quatre. Dans l'ordre qu'il souhaite, l'utilisateur doit ouvrir le menu permettant de sélectionner la forme ou la couleur de la figure présentée. Lorsqu'une caractéristique est sélectionnée, elle apparaît sur l'item permettant d'ouvrir le menu pour permettre à l'utilisateur de vérifier sa sélection. Si les deux caractéristiques ont bien été sélectionnées, une nouvelle figure est présentée à l'utilisateur lui permettant d'effectuer la tâche à nouveau.



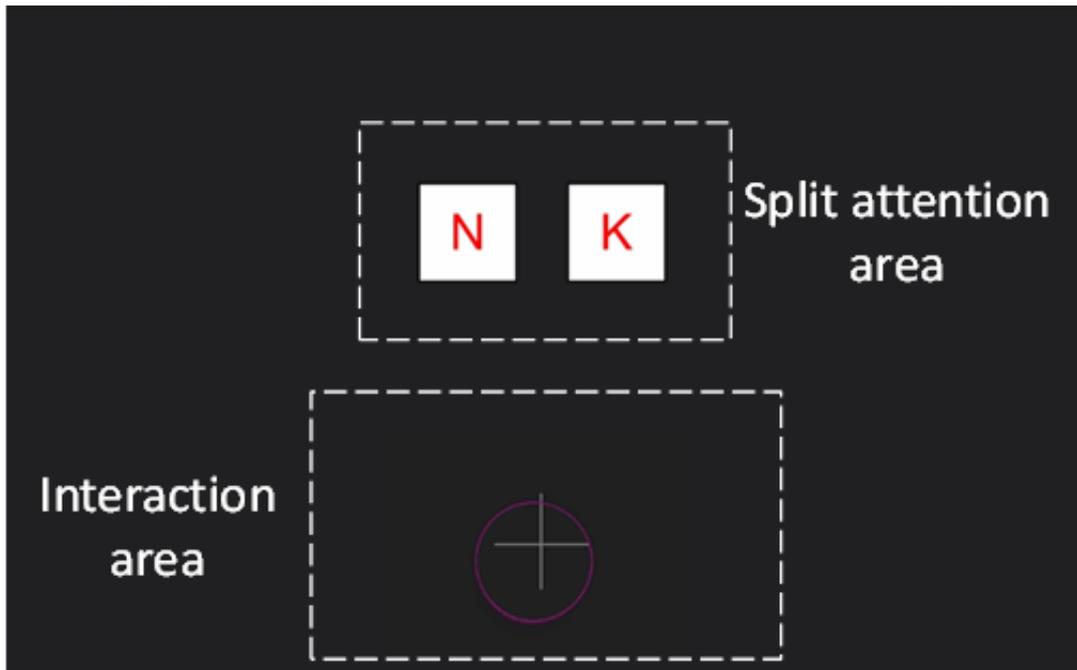
**Figure 47. Tâche d'interaction. Le menu de gauche permet de sélectionner la forme de la figure au-dessus alors que le menu de droite permet d'en sélectionner la couleur.**

Tâche d'attention divisée n°1 :

Cette tâche est utilisée pour inciter l'utilisateur à diviser son attention durant la première partie de l'expérimentation. Elle consiste à comparer deux symboles (lettre ou chiffre) en vision périphérique. Deux symboles sont présentés à l'utilisateur au-dessus de la zone dans laquelle il réalise la tâche d'interaction. Le symbole de gauche reste identique alors que le symbole de droite change régulièrement toute les 1-2 secondes. Si les symboles sont identiques, l'utilisateur spécifie à haute voix qu'il a détecté la similitude. Le symbole de gauche change alors aussi lors du prochain changement de symbole. Le symbole de droite a été forcé à être identique au symbole de gauche dans une fourchette de 7 secondes. Certains symboles ont été retirés (0, O, F et I)<sup>8</sup> car ils ont présenté des difficultés à être différencié d'autres symboles durant les pré-tests.

Cette tâche doit être effectuée en vision périphérique pour obliger l'utilisateur à concentrer son attention visuelle à cet endroit. Pour cela, dès lors que l'utilisateur sort son regard de la zone d'interaction (cf. Figure 48), les cases contenant les symboles disparaissent, obligeant l'utilisateur à retourner dans la zone d'interaction. Cela a été mis en place pour éviter à l'utilisateur d'alterner chronologiquement son attention visuelle entre les deux zones plutôt que de la séparer spatialement.

<sup>8</sup> 0 et O étaient confondus facilement avec Q, F avec E et I avec 1.



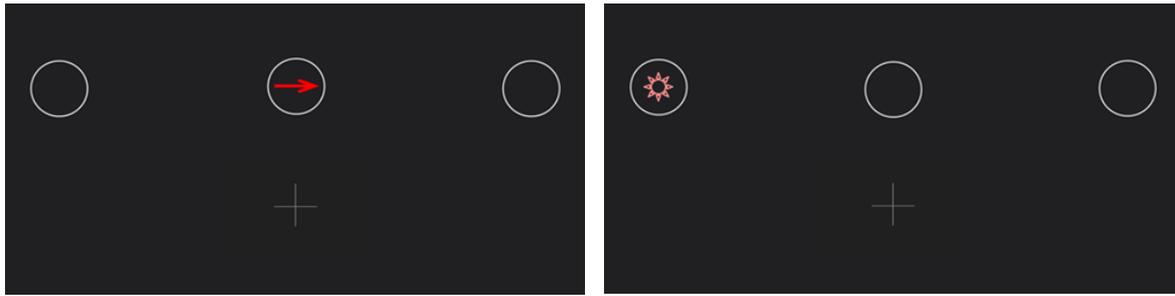
**Figure 48. Tâche d'attention divisée n°1. L'utilisateur doit signifier à haute voix quand les symboles sont identiques. Si le regard sort de la zone d'interaction, les symboles disparaissent.**

Comme présenté sur la Figure 48, la zone d'interaction est située dans la partie inférieure de l'écran alors que la tâche d'attention divisée s'effectue dans la partie supérieure.

Tâche d'attention divisée n°2 :

La seconde tâche d'attention divisée vise à mettre en évidence la capacité à localiser l'attention visuelle. Elle reprend pour cela le protocole de Posner (Posner, 1980). La zone d'attention déportée dispose de trois cercles situés au-dessus de la zone d'interaction, à gauche, au milieu et à droite (cf. Figure 49). L'écart entre les cercles représente environ 6°. Notons bien que Posner a montré que cette distance n'avait pas d'impact. Toutes les trois secondes, une flèche (appelée indice) apparaît dans le cercle central et indique la position où doit apparaître un stimulus quelques instants après. L'indice reste en place pendant un temps aléatoire entre 0.5 et 1 seconde et l'utilisateur doit alors orienter son attention visuelle vers le cercle pointé par la flèche. Après un temps variant entre 0.5 et 2.5 secondes, un stimulus apparaît d'un côté ou de l'autre pendant un temps très court (0.2s). Dès que l'utilisateur détecte un stimulus visuel, il le signale vocalement en disant « vu ». Il effectue cette action que l'indice soit présenté dans le cercle précédemment visé par l'indice ou dans l'autre cercle.

Lorsque l'indice pointe vers l'endroit où apparaît effectivement le stimulus, on dit que l'indice est valide, dans le cas contraire, l'indice est invalide. Afin de garder la confiance de l'utilisateur dans les indices, 70 % des indices sont valides.



*Figure 49. Tâche d'attention divisée n°2. L'utilisateur oriente son attention vers le cercle pointé par la flèche tout en continuant d'interagir avec les menus en bas de l'écran (aux alentours de la croix).*

## **Design expérimental et procédure**

### Phase d'apprentissage :

La première étape de l'expérimentation a consisté en une phase d'apprentissage. Comme cette étape durait environ une heure, elle était effectuée la veille de l'expérimentation pour laisser à l'utilisateur le temps de se reposer les yeux. Durant cette phase, l'utilisateur se familiarisait avec les différentes modalités d'interaction sur une tâche similaire à la tâche principale. 20 minutes ont été utilisées pour l'apprentissage de chaque modalité. Les bases de chaque modalité lui étaient présentées et, comme l'opérateur pouvait voir à côté les mouvements du regard de l'utilisateur, l'utilisateur était sensibilisé aux différentes difficultés de chaque modalité.

Des problèmes comme la synchronisation œil-main ont été abordés afin de sensibiliser l'utilisateur à une erreur courante qui était d'appuyer sur le bouton physique avant que le regard ne soit effectivement sur l'élément à activer ou juste après que le regard l'ait quitté. L'utilisateur apprenait aussi à ressentir un offset dans la position donnée par le système et à le compenser en regardant légèrement à côté de l'item à sélectionner.

C'est durant cette phase d'apprentissage que les temps de fixation pour les modalités *DwellTime* et *DwellTime* délocalisé étaient individuellement calibrés. Ce temps était initialisé à 0.5s et l'opérateur avait la possibilité de le modifier par palier de 25ms. Durant cette phase, l'utilisateur était régulièrement sollicité pour verbaliser son ressenti sur le temps de fixation et pouvait demander à l'augmenter ou à le réduire. L'opérateur, qui était habitué à l'interaction au regard, pouvait visualiser des indices montrant un temps trop court ou trop long pour chaque utilisateur. Les sélections intempestives d'éléments, souvent verbalisées par les utilisateurs (ou détectable par un grognement de frustration) étaient représentatives d'un temps trop court. Au contraire, une anticipation de la part de l'utilisateur pour aller regarder autre part avant d'avoir accompli l'activation d'un élément voulu (là aussi souvent accompagné d'une preuve auditive de frustration) était représentatif d'un temps trop long. Dans la plupart des cas, l'utilisateur avait une bonne sensation quant au temps qui lui convenait et l'opérateur n'a pas modifié ce temps de plus de 50ms.

### Expérimentation 1 :

La première partie de l'expérimentation consistait en la réalisation simultanée de la première tâche d'attention divisée et de la tâche d'interaction. Quatre sessions ont été conduites, une pour chaque modalité d'interaction dont l'ordre était défini aléatoirement. Chaque session était une succession de

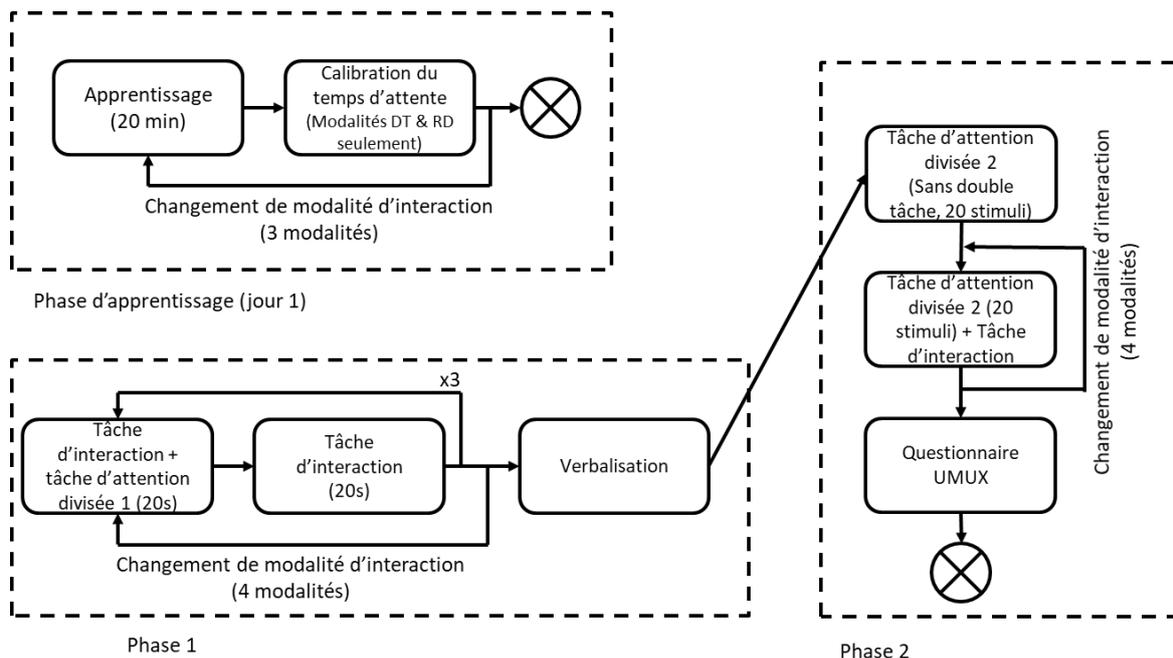
segments de 20 secondes avec ou sans tâche d’attention divisée à réaliser pour un total de 2 minutes, soit 3 cycles avec/sans double tâche. L’intérêt était d’éviter une session trop longue de double tâche car celle-ci demandait un effort particulier. Il était demandé à l’utilisateur de prioriser la tâche d’attention divisée autant que possible.

Expérimentation 2 :

Après 5 minutes de pause et de verbalisation non-dirigée sur les différentes modalités d’interaction, la seconde partie de l’expérimentation consistait en la réalisation simultanée de la seconde tâche d’attention divisée avec la tâche d’interaction. Cinq sessions ont été réalisées. La première session ne comportait pas de tâche d’interaction afin de vérifier les résultats obtenus par Posner, puis les quatre autres sessions ont été effectuées respectivement avec les quatre modalités d’interaction étudiées dans un ordre aléatoire. Chaque session consistait en 20 stimuli pour une durée d’environ 3 minutes. Dans cette partie, il était demandé à l’utilisateur de prioriser autant que possible la tâche d’interaction au dépend de la tâche d’attention divisée.

A la fin de ces deux parties, un questionnaire papier était présenté aux utilisateurs. La durée totale de l’expérimentation était d’environ 35 minutes.

La Figure 50. Diagramme représentant la chronologie de l’expérimentation 2. résume la chronologie du protocole expérimental.



**Figure 50. Diagramme représentant la chronologie de l’expérimentation 2.**

**Variables mesurées**

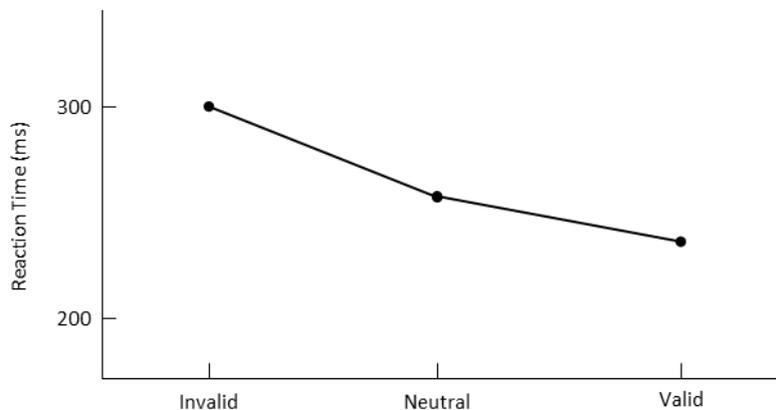
La première expérimentation visait à évaluer les performances des différentes modalités d’interaction alors que l’utilisateur essayait activement de diviser spatialement son attention visuelle. Les variables observées sont le temps d’exécution et le taux d’erreurs.

#### 4.2 - Effet de la réorientation de l'attention visuelle sur l'interaction au regard

- Le temps d'exécution est le temps écoulé entre l'apparition de la figure à caractériser et la sélection des deux bonnes caractéristiques. Ce qui nous importe est l'évolution du temps d'exécution avec et sans double tâche. Nous avons donc comparé statistiquement les temps d'exécution en utilisant des ANOVAs à deux facteurs pour étudier l'impact global, puis des ANOVAs pour étudier chaque modalité indépendamment. Les tests statistiques ont été effectués sur le logarithme du temps d'exécution, plus enclin à suivre une loi normale, bien que les résultats soient présentés sur une échelle linéaire par souci de lisibilité.
- Le taux d'erreur est le ratio de mauvaises sélections (sélection d'une mauvaise caractéristique) sur le nombre de sélections total. En réalité, il existe d'autres moyens d'effectuer une erreur (ne pas réussir à activer un élément par exemple). Des ANOVAs à deux facteurs sur le type de menu et la modalité d'interaction suivis d'ANOVA simples sur la modalité d'interaction uniquement ont été utilisées pour tester la significativité des résultats.

La seconde expérimentation visait à évaluer la capacité de l'utilisateur à effectivement déporter son attention pendant une interaction au regard.

- Les résultats de Posner (Posner, 1980) montrent que, pour la même tâche que la tâche d'attention divisée n°2, le temps de réaction à un stimuli après un indice valide est plus court que le temps de réaction après un indice invalide (cf. Figure 51). On peut observer une différence de 50ms sur le temps de réaction des sujets. Si ce profil se présente, nous pourrions conclure que l'utilisateur a effectivement réussi à diviser son attention visuelle.



**Figure 51. Temps de réaction à un stimuli après des indices valides et invalides obtenus par Posner (Posner, *Orienting of attention*, 1980).**

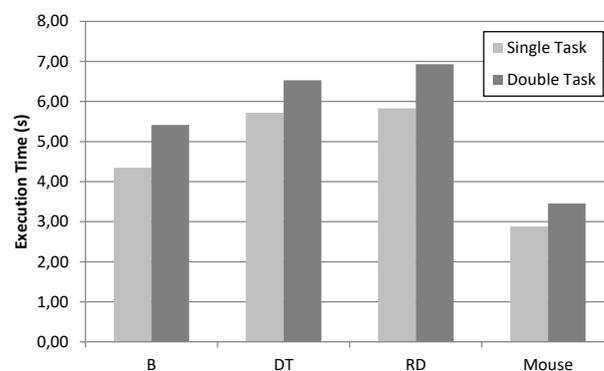
- Un questionnaire UMUX (Bosley, 2013) a été présenté aux utilisateurs à la fin de l'expérimentation (cf. Annexe C). Ce questionnaire permet d'obtenir une note sur 100 représentative de l'utilisabilité d'un système (ici d'une modalité d'interaction). Cette note est représentative de l'utilisabilité, de la fiabilité et de l'adéquation avec la tâche. Ces notes sont observées par la représentation de leur répartition. Le questionnaire UMUX juge normalement de l'utilisabilité de systèmes complets. Ici, la modalité d'interaction peut être associée au système, nous avons remplacé les phrase contenant « le système » par « la manière d'interagir ».

## 4.2.3 Résultats

### ***Temps d’exécution (Expérimentation 1)***

L’ANOVA sur mesures répétées à deux facteurs n’a pas montré d’effet principal significatif de la modalité d’interaction sur l’augmentation du temps d’exécution lors du passage en double tâche ( $p > 0.05$ ). Cela veut dire que l’augmentation du temps d’exécution est indépendante de la modalité d’interaction.

Cependant, en considérant chaque modalité d’interaction indépendamment, une différence significative entre le temps d’exécution avec et sans double tâche a été mis en avant pour chaque modalité (quatre  $p < 0.05$ , cf. Figure 52). L’augmentation du temps d’exécution est d’environ 20 % pour les quatre modalités d’interaction.



**Figure 52. Moyenne des temps d’exécution de la tâche d’interaction avec et sans tâche d’attention divisée.**

La modalité d’interaction a un effet significatif sur le temps d’exécution. Les analyses post-hoc montrent une différence significative ( $p < 0.05$ ) du temps d’exécution entre la modalité Bouton et les deux modalités *DwellTime* et *DwellTime* délocalisé (RD pour relocated *DwellTime* sur la Figure 52). L’utilisation d’un bouton est 20 % plus rapide que l’interaction au regard seul. Aucune différence significative n’a été observée entre les deux interactions utilisant le regard seul ( $p > 0.05$ ). L’interaction à la souris est significativement plus rapide que les trois autres modalités ( $3 p < 0.01$ ).

### ***Taux d’erreur (Expérimentation 1)***

L’ANOVA à deux facteurs montre un effet d’interaction de la modalité d’interaction sur l’évolution du taux d’erreur de la simple tâche à la double tâche ( $p < 0.01$ ). En accord avec la Figure 53, les interactions *DwellTime* et *DwellTime* délocalisé ont subi un impact significatif ( $p < 0.01$ ) du passage de la simple tâche à la double tâche. Cela signifie que les performances de ces 2 modalités étaient moins bonnes lorsque l’utilisateur était visuellement concentré autre part. Cela a été verbalisé par les utilisateurs : lorsqu’ils se concentrent pour orienter leur attention, ils effectuent involontairement une pause pouvant conduire à une sélection involontaire. Pour ces deux modalités, le taux d’erreur a doublé.

En revanche, pour les modalités Bouton et Souris, on ne constate pas d’évolution significative du taux d’erreur entre la simple tâche et la double tâche (2 valeurs de  $p > 0.05$ ).

#### 4.2 - Effet de la réorientation de l'attention visuelle sur l'interaction au regard

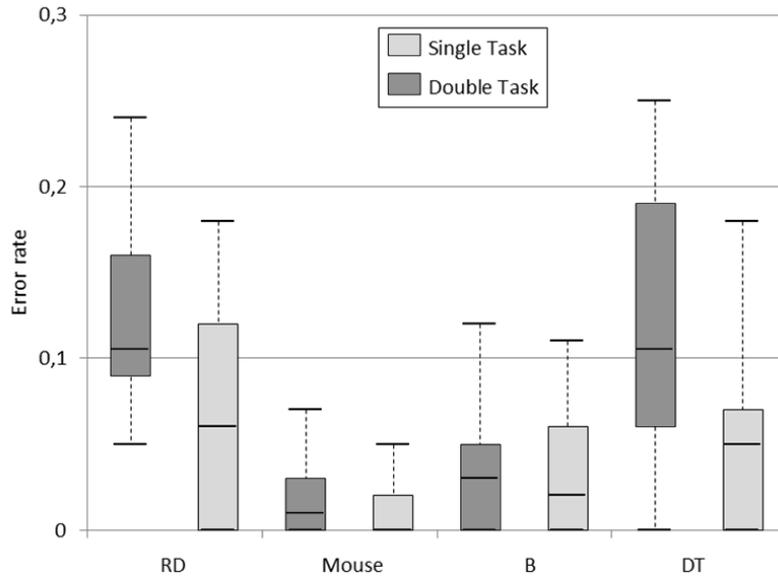
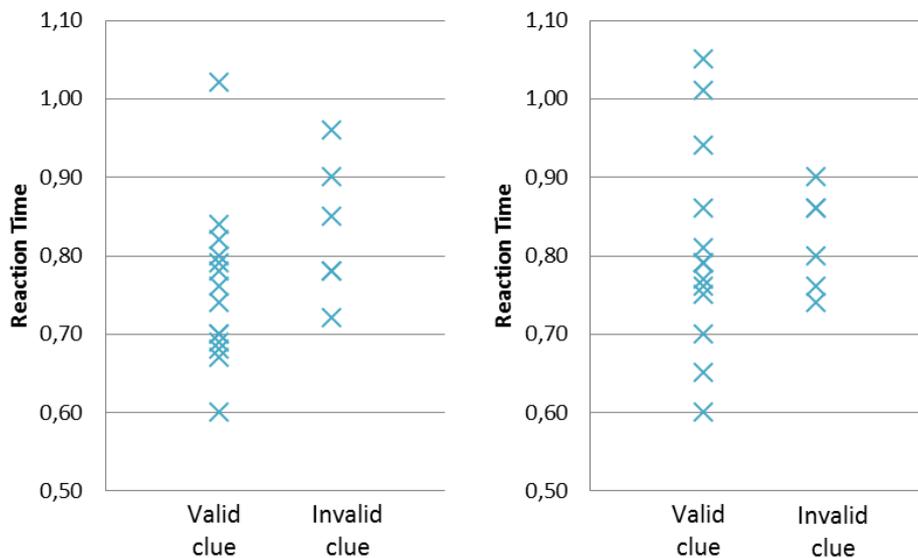


Figure 53. Taux d'erreur par modalité d'interaction pour une simple et double tâche.

Il est aussi notable que certains utilisateurs à l'aise ont réussi à obtenir un score de 0% d'erreur en double tâche avec les modalités Bouton et Souris et n'ont pas réussi avec les deux autres modalités.

#### Capacité à orienter son attention (Expérimentation 2)

Deux types de profils concernant les temps de réaction aux stimuli ont été observés durant la seconde phase de l'expérimentation. Les utilisateurs qui ont réussi à orienter leur attention tout en interagissant ont eu un temps de réaction plus court à l'apparition du stimulus lorsque l'indice était valide, en cohérence avec les résultats de Posner (Posner, 1980) (cf. Figure 54, gauche). Dans le cas contraire (cf. Figure 54, droite), on considère que l'utilisateur n'a pas réussi à orienter son attention tout en interagissant avec le regard sur une autre partie de l'écran. Nous avons considéré que l'utilisateur arrivait à orienter son attention lorsque la différence des moyennes des temps de réaction après indices valides et invalides dépassait le seuil empirique de 0.8s.



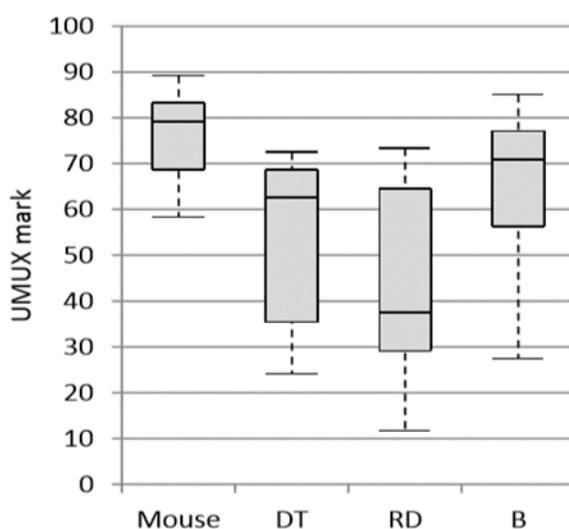
**Figure 54. Temps de réaction à un stimulus après un indice valide et invalide. A gauche un utilisateur réussissant à orienter son attention, à gauche un utilisateur n'y arrivant pas.**

Lors de l'utilisation de la souris, 15 % des utilisateurs ont réussi à orienter leur attention tout en interagissant, 20 % pour la modalité Bouton, et respectivement 40 et 45 % pour les modalités *DwellTime* délocalisé et *DwellTime*. Dans tous les cas, les utilisateurs réussissant à déporter leur attention avec une des modalités Souris ou Bouton ont forcément réussi à déporter leur attention avec les deux autres modalités.

Ces résultats sont cohérents avec les retours donnés par les utilisateurs. Ils ont dit avoir plus de facilité à « élargir » ou changer la localisation de leur attention visuelle lors du temps de fixation des modalités basées sur un temps de fixation. Ce temps d'attente représente environ 50% du temps total passé à regarder l'interface.

### **Appréciation utilisateur (Expérimentation 2)**

L'analyse des formulaires UMUX montrent que l'expérience utilisateur des différentes modalités d'interaction est assez éparse. La Figure 55 montre que l'interaction à la souris a obtenu de meilleurs scores que les modalités au regard. Parmi celles-ci, l'interaction multimodale se démarque des deux modalités à base de temps de fixation. La modalité *DwellTime* délocalisé a été considérée très difficile à utiliser en même temps que l'utilisateur devait délocaliser son attention visuelle puisque la modalité d'interaction est déjà demandante à ce niveau. Cela a été mis en évidence par le fait que la modalité a obtenu un score inférieur à 40 avec plus de 50% des sujets.



**Figure 55. Distribution statistique des notes obtenues à l'aide du questionnaire UMUX. L'UMUX est représentatif de l'usabilité, de la fiabilité ainsi que de l'adéquation avec la tâche.**

### **4.2.4 Discussion**

Afin de pouvoir inclure des modalités d'interaction au regard pour améliorer l'utilisabilité de systèmes, il est important d'étudier l'adéquation de l'interaction au regard avec les tâches

#### 4.2 - Effet de la réorientation de l'attention visuelle sur l'interaction au regard

effectivement utilisées dans ces systèmes. Cette expérimentation a cherché à étudier le comportement de différentes modalités d'interaction lorsque l'utilisateur délocalise son attention visuelle pendant une interaction au regard, ce qui est une action courante lors d'une tâche d'observation ou de surveillance utilisant les produits optroniques.

Lors de l'expérimentation, les modalités d'interaction au regard semblent avoir été affectées de la même manière que l'interaction à la souris en ce qui concerne le temps d'exécution. Une augmentation d'environ 20% du temps pour réaliser la tâche est observée sur l'ensemble des modalités testées. Cette variable n'est pas limitante quant au choix des méthodes à utiliser lors d'une tâche simultanée exigeant de réorienter son attention visuelle.

En ce qui concerne le taux d'erreur, les mesures directes semblent indiquer que le taux d'erreur augmente en double tâche lors de l'utilisation de modalités d'interaction au regard basées sur un temps de fixation alors que ce n'est pas le cas pour une interaction multimodale. Cependant, la seconde partie de l'expérimentation laisse penser que les utilisateurs ont beaucoup plus réussi à réorienter leur attention avec les modalités à base de *DwellTime*. Nous pensons que les utilisateurs n'ont, en fait, pas vraiment réussi à déporter leur attention avec la souris et avec la modalité Bouton lors de la première phase de l'expérimentation. Aucune métrique sur la double tâche de reconnaissance des symboles n'a été mise en œuvre pour éventuellement mesurer cela.

Le *DwellTime* délocalisé a été conçu pour améliorer l'utilisabilité d'une interaction basée sur un temps de focalisation, cette amélioration n'a pas été constatée. Cette modalité est en réalité plus difficile à utiliser selon les retours utilisateurs.

En revanche, le questionnaire UMUX met en avant l'adéquation de l'interaction Bouton avec la double tâche. Encore une fois, il est nécessaire de confirmer que les utilisateurs arrivaient réellement à réorienter leur attention dans ce cas.

L'expérience a été conduite avec des novices. Le fait de se concentrer visuellement sur une zone ou le regard n'est pas focalisé n'est pas une tâche triviale. Ainsi des résultats différents pourraient être obtenus avec des utilisateurs experts.

Cette expérimentation concerne uniquement l'analyse de l'orientation endogène de l'attention visuelle, c'est-à-dire issue d'un processus conscient de l'utilisateur. Une autre expérimentation a été effectuée sur l'attention exogène dans la section suivante.

Bien que l'expérimentation ait été conçue pour diviser spatialement l'attention visuelle de l'utilisateur, il est très difficile de maintenir une telle concentration pendant toute la durée de l'expérimentation. Ainsi, comme les utilisateurs l'ont évoqué, la séparation de l'attention n'a pas été constante et les tâches ont parfois été traitées séquentiellement plutôt que simultanément. Néanmoins, cela est représentatif du comportement des utilisateurs en situation d'observation ou de surveillance mais peut être limitant quant à l'interprétation du processus cognitif effectivement mis en œuvre par les utilisateurs

Les deux parties de l'expérimentation ont demandé une priorisation différente entre la tâche d'attention divisée et la tâche d'interaction. Cependant, certains utilisateurs disent ne pas avoir réussi à prioriser la bonne tâche. Il est difficile d'évaluer quelle fraction des ressources cognitives est

allouée à telle ou telle tâche. Cette division est un compromis approché différemment par les différents utilisateurs.

Nous sommes partis du postulat que les résultats de Posner étaient applicables dans un cas plus complexe que lors de ses travaux. De plus, nous avons distingué les utilisateurs qui ont réussi à orienter leur attention des autres à l'aide d'un seuil empirique sur la moyenne. Bien que cette méthode ait donné des résultats encourageants, elle pourrait être améliorée en récoltant plus de données par utilisateurs et en utilisant des ANOVAs par exemple.

Résultats principaux de l'expérimentation :

- Les modalités d'interaction au regard et d'interaction à la souris sont impactées de la même façon par une tâche d'attention visuelle divisée
- Il est plus facile de diviser son attention visuelle avec les modalités à base de temps de fixation, mais cela augmente le taux d'erreur

## 4.3 Effet de l'interaction au regard sur la capacité à détecter des signaux en vision périphérique

Dans cette troisième section, nous continuons à nous intéresser à l'impact de la tâche utilisateur sur les performances des différentes modalités d'interaction au regard. Le type de tâche étudié est la détection d'alertes visuelles non attendues. Nous commençons par expliquer pourquoi nous avons étudié cette caractéristique Puis nous présentons ensuite le design expérimental suivi des résultats et d'une discussion.

### 4.3.1 Besoin opérationnel et état de l'art

Interagir avec le regard demande une certaine concentration et, comme cela a déjà été évoqué, cela change l'utilisation normale de l'œil dont la fonction habituelle est d'analyser. Dans la section précédente, nous avons vu que la tâche de fond effectuée par l'utilisateur était en lien direct avec l'interaction. Elle est modifiée par le fait d'interagir au regard, et les modalités d'interaction au regard ne se comportent pas de la même manière en fonction de la tâche de fond.

Lors de l'utilisation d'un système tel que des lunettes de visée ou des jumelles infrarouges, dans la plupart des cas, l'interaction est induite par l'utilisateur et le système est réactif des demandes de l'utilisateur. Cependant, il arrive aussi que système doive communiquer des informations à l'utilisateur qui n'ont pas été demandées. Considérons trois situations :

- Le premier cas de figure serait un besoin de communiquer sur l'état même du système comme la batterie, la mémoire restante ou un défaut dans l'un des composants. Ces cas sont souvent prévus lors de la conception du produit, et, à l'instar des voyants d'un tableau de bord de voiture, ces informations sont représentées par des symboles chacun lié à un problème différent.
- Le second cas de figure est l'affichage d'informations issues d'un processus automatique du système, ou la communication d'informations provenant de systèmes extérieurs. Un exemple pourrait être la mise en mouvement d'unités adverses surveillées. Si des unités sortent de la zone dans laquelle elles étaient statiques, il est nécessaire de communiquer rapidement cette information aux fantassins concernés. En fonction de la criticité, on peut vouloir interrompre la tâche actuelle de l'utilisateur, ou bien seulement lui signaler l'information.
- Le troisième cas est l'apparition succincte d'un élément d'intérêt dans une zone d'observation supposée vide, où l'utilisateur ne s'attend pas à voir quelque chose.

Quel est le point commun entre ces trois cas ? Ces trois cas sont associables à de la détection d'alertes en vision périphérique. Contrairement au chapitre précédent où une attention de l'utilisateur était impliquée, nous parlons ici de perception pure.

La détection d'alerte en vision périphérique est liée au concept d'attention exogène. L'attention de l'utilisateur est dans ce cas dirigée par un élément nouveau et/ou saillant (e.g. un changement rapide de luminance (Jonides & Yantis, 1988)) plutôt que par le processus conscient de l'utilisateur. Berger et ses collègues (Berger, Henik, & Rafal, 2005) donnent plus d'informations sur les différences entre attention endogène et exogène.

Les performances de détection d'alerte en vision périphérique sont directement liées à la charge mentale lors de l'apparition des alertes. Plus la concentration de l'utilisateur sur la tâche en cours est grande, plus les performances de détection ou de reconnaissance d'alertes visuelles sont faibles. De fait, l'analyse des performances de détection est utilisée comme un moyen de mesure de la charge mentale induite par certaines situations ou certaines tâches. Martens et Van Winsum par exemple (Martens & Van Winsum, 2000), donnent un exemple de cette utilisation à différentes étapes de la conduite d'un véhicule. Ils utilisent le temps de réaction à un stimuli pour conclure sur la charge mentale.

D'après Jonides (Jonides J. , 1981), il semble que la détection de stimuli visuels en vision périphérique soit moins affectée par une augmentation de la charge cognitive que peut l'être une tâche d'orientation endogène de l'attention. Cependant, d'autres aspects impactant les performances de l'orientation de l'attention exogène ont été étudiés. Par exemple, Mackworth (Mackworth, 1965) a montré que du bruit visuel diminue les capacités de détection en vision périphérique. C'est notamment le cas lors de la réalisation d'une tâche, même peu coûteuse en termes de charge mentale dans la vision centrale. Berger montre aussi la priorité de l'orientation de l'attention endogène sur l'attention exogène.

En reprenant la métaphore du faisceau lumineux (cf. section 2.2.3), nous pensons que des modalités d'interaction au regard basées sur un temps de fixation donnent l'occasion à l'utilisateur d'élargir momentanément la largeur du faisceau d'attention, et ainsi d'être plus enclin à détecter des alertes en vision périphérique.

Cette section vise à comprendre l'impact que peut avoir le fait d'interagir au regard sur les capacités de détection d'alertes en vision périphérique. Deux expérimentations sont présentées : la première concerne effectivement la mesure des performances de détection de stimuli visuels en vision périphérique. La seconde mesure la charge cognitive liée aux différentes modalités d'interaction afin de vérifier si les mesures observées par la première expérimentation ne sont pas simplement dues à l'effet tunnel.

### **4.3.2 Méthode et design expérimental**

#### ***Participants***

Cette expérience a été conduite avec 20 volontaires participants âgés de 21 à 60 ans. Les 20 participants disposaient tous soit d'une vision parfaite, soit d'une vision totalement compensée par des lunettes ou des lentilles de contact. Comme pour les autres expérimentations, tous étaient des utilisateurs fréquents de systèmes informatiques, soit dans le milieu professionnel, soit pendant leur temps libre. Ces 20 participants ont réalisé les deux expérimentations présentées dans ce chapitre mais n'avaient pas participé aux précédentes expériences.

#### ***Dispositif expérimental***

Pour cette expérimentation visant à faire apparaître des stimuli en vision périphérique, un écran 4k de 58" a été utilisé. Lorsque l'utilisateur était à environ 50 cm de cet écran, cela correspondait à une ouverture visuelle de 100°, représentative du champ de vision des casques de réalité virtuelle actuels (cf. Figure 56). Un système de suivi du regard Tobii EyeX est disposé en-dessous de l'écran. Un clavier

#### 4.3 - Effet de l'interaction au regard sur la capacité à détecter des signaux en vision périphérique

était aussi à disposition de l'utilisateur, notamment pour la touche espace utilisée pour la méthode d'interaction multimodale. Un clavier externe permettait à l'opérateur de gérer l'avancement de l'expérimentation ainsi que de spécifier la détection de stimuli de la part de l'utilisateur.

Le Tobii EyeX n'étant pas utilisable sur un écran de cette taille, l'ensemble des interactions est effectué en bas et au centre de l'écran, au plus proche du système de suivi du regard.

Comme précédemment, un second écran a été mis en place en dehors du champ de vision des participants pour permettre à l'opérateur de visualiser les mouvements du regard des utilisateurs, notamment pour les phases d'apprentissage et pour la phase de calibration. Cela a été réalisé en utilisant le logiciel OBS Studio en association avec le *Streaming gaze overlay* fourni par Tobii.



*Figure 56. Dispositif expérimental. L'opérateur peut voir les mouvements du regard de l'utilisateur sur son écran.*

### **Conception de l'interface**

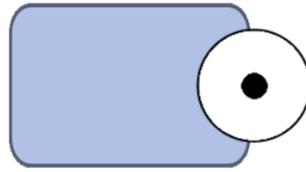
Pour ces expérimentations, des boutons virtuels rectangulaires ont été utilisés. Cette forme est plus naturelle lors de l'affichage de texte. Ces boutons faisaient  $3^\circ$  de large pour  $2.25^\circ$  de haut afin de respecter la bonne pratique nécessitant des items plus grands que  $1^\circ$  d'angle. Des items suffisamment larges ont été choisis pour compenser les problèmes de précisions dus à l'utilisation du système Tobii EyeX sur un écran large.

Deux écrans se succédaient (cf. Figure 58). Le premier écran contient un seul élément alors que le second contient une grille de  $4 \times 3$  items. Ces items étaient plus espacés que lors des autres expérimentations, pour un écart de  $1^\circ$  entre les bords de chaque item.

### **Modalités d'interaction**

Les modalités d'interaction testées dans cette expérience sont le *DwellTime*, le *DwellTime* délocalisé, l'interaction multimodale regard + bouton. Ces modalités sont présentées dans le chapitre 4.1.2. Une interaction à la souris a aussi été effectuée pour servir de référence pour les comparaisons.

Pour l’interaction *DwellTime* délocalisée, l’item à fixer est un disque blanc avec un point noir au centre disposé à droite de l’item auquel il est rattaché, suivant ainsi le sens de la lecture (cf. Figure 57).



*Figure 57. Item utilisé pour la modalité DwellTime délocalisé lors de cette expérimentation. L’item déporté est sur la droite de l’item principal.*

## **Tâches expérimentales**

Cette expérimentation présente deux tâches réalisées simultanément : une tâche d’interaction consistant à interagir avec le regard, et une tâche de détection d’alerte visuelle en vision périphérique.

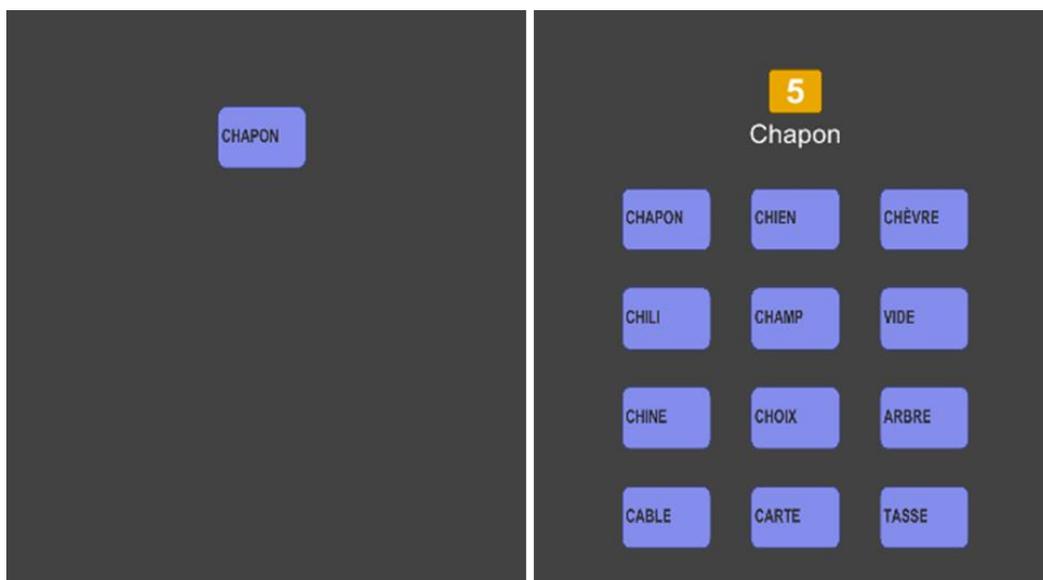
### Tâche d’interaction :

La tâche d’interaction est divisée en deux étapes présentées sur la Figure 58 :

- La première étape consiste à prendre connaissance d’un mot et d’activer l’item sur lequel il est inscrit. L’activation se fait avec l’une des modalités d’interaction présentées précédemment.
- La deuxième étape consiste en une phase de recherche du mot parmi une grille de 12 mots et de l’activation de l’item lié à ce dernier.

Afin de maintenir la priorité sur la tâche d’interaction et de la rendre stimulante, la phase de recherche est limitée à 5 secondes. Cette durée est largement suffisante pour réaliser la tâche si l’utilisateur reste concentré.

Lorsque les 5 secondes sont écoulées ou qu’un mot a été sélectionné par l’utilisateur, un nouveau mot apparait seul, et la tâche recommence à zéro.



#### 4.3 - Effet de l'interaction au regard sur la capacité à détecter des signaux en vision périphérique

**Figure 58. Tâche d'interaction de l'expérimentation. A gauche, l'utilisateur prend connaissance du mot et l'active avec la modalité d'interaction testée. A droite l'utilisateur recherche ce mot parmi 12 et active l'item associé.**

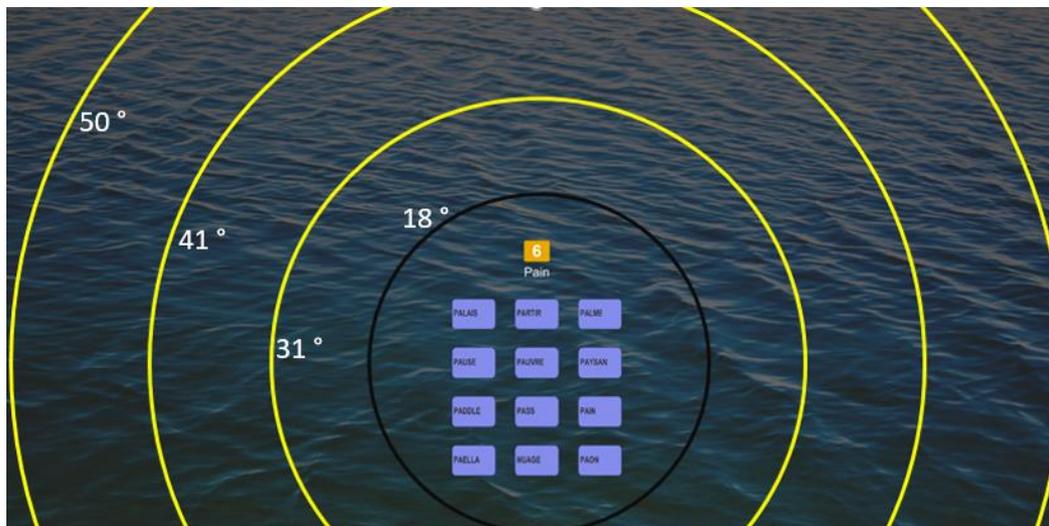
##### Tâche de détection de stimuli :

Pour l'utilisateur, la tâche de détection de stimuli consiste simplement à dire « vu » lors de l'apparition d'un stimulus sur l'écran. Ce stimulus est un disque bleu de diamètre 0,6°. Ces caractéristiques ainsi que le fond d'écran sombre représentant la surface d'une étendue d'eau ont été choisies de manière empirique afin de rendre cette détection réalisable mais non-triviale dans l'ensemble des parties de l'écran.

L'apparition de ces stimuli se fait dans trois zones (cf. Figure 59) représentant des ouvertures angulaires différentes :

- La zone 1 représente les ouvertures angulaires par rapport au centre entre 18° et 31°.
- La zone 2 représente les ouvertures angulaires par rapport au centre entre 31° et 41°.
- La zone 3 représente les ouvertures angulaires par rapport au centre entre 41° et 50°.

Ces zones sont en réalité centrées sur le regard de l'utilisateur lors de l'apparition du stimulus. L'ouverture angulaire à laquelle apparait un stimulus est calculée par rapport à la position du regard. Le stimulus est statique durant toute sa durée d'apparition, il ne bouge pas en fonction des mouvements du regard.



**Figure 59. Représentations des zones dans lesquelles apparaissent les stimuli. Ces cercles sont invisibles lors de l'expérimentation et sont affichés ici à des fins explicatives.**

Pour chaque modalité d'interaction testée, 24 stimuli apparaissent (8 par zone). Un stimulus apparait toutes les 3 à 5 secondes et est affiché durant 0,5 secondes.

##### Tâche de mesure de la charge mentale :

La tâche de mesure de la charge mentale est une tâche de calcul mental simple, c'est le PASAT (Paced Auditory Serial Addition Test) (Gronwall, 1977). À intervalle de 3 secondes, des chiffres sont énoncés de façon orale à l'utilisateur qui doit donner à haute voix les sommes intermédiaires des

chiffres entendus. C'est-à-dire la somme des deux derniers chiffres énoncés et non pas la somme cumulée. La Figure 60 schématise ce processus. Si l'utilisateur décroche, il lui est demandé de recommencer comme si la série reprenait depuis le début. Un total de 21 chiffres sont énoncés, impliquant 20 sommes intermédiaires. L'énoncé des chiffres n'est pas fait par l'opérateur mais à l'aide d'un enregistrement. Différents enregistrements ont été utilisés pour chaque itération de ce test.

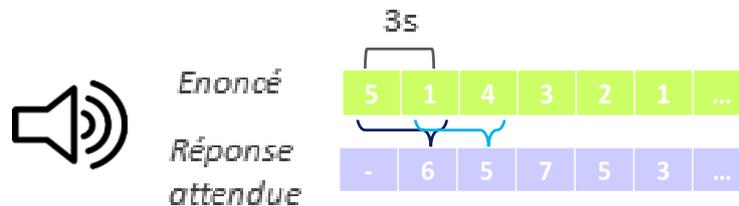


Figure 60. Représentation des sommes intermédiaires du PASAT.

## Design expérimental et procédure

### Phase d'apprentissage :

La phase d'apprentissage des modalités d'interaction est la même que celle présentée dans l'expérimentation présentée précédemment mais a été réalisée avec la tâche d'interaction présentée ci-dessus. Elle est composée d'une explication de la tâche, d'une habitude aux modalités d'interaction et d'une phase de calibrage des temps d'attente pour les méthodes *DwellTime* et *DwellTime* délocalisé.

### Expérimentation 1 :

L'expérimentation 1 regroupe la tâche d'interaction avec la tâche de détection de stimuli en vision périphérique. Dans un premier temps, seule la tâche de détection de stimuli a été réalisée. L'utilisateur devait alors garder son regard au centre de l'écran. Dans un second temps, l'expérimentation a consisté en la réalisation de la tâche d'interaction avec les différentes modalités d'interaction testées tout en réalisant la tâche de détection de stimuli. L'ordre de passage des différentes modalités d'interaction a été tiré aléatoirement pour chaque participant. Une pause d'une minute a été observée entre chaque rush.

### Expérimentation 2 :

L'expérimentation 2 regroupe la même tâche d'interaction avec la tâche de mesure de la charge mentale. Le PASAT a été réalisé 5 fois par participant. Dans un premier temps, seul le PASAT était réalisé pour servir de référence pour les résultats. L'utilisateur devait alors aussi garder son regard au centre de l'écran. Dans un second temps, l'expérimentation a consisté en la réalisation de la tâche d'interaction avec les différentes modalités d'interaction testées tout en réalisant le PASAT. L'ordre de passage des modalités d'interaction été aléatoire pour chaque participant.

La chronologie de l'expérimentation est résumée sur la Figure 61.

#### 4.3 - Effet de l'interaction au regard sur la capacité à détecter des signaux en vision périphérique

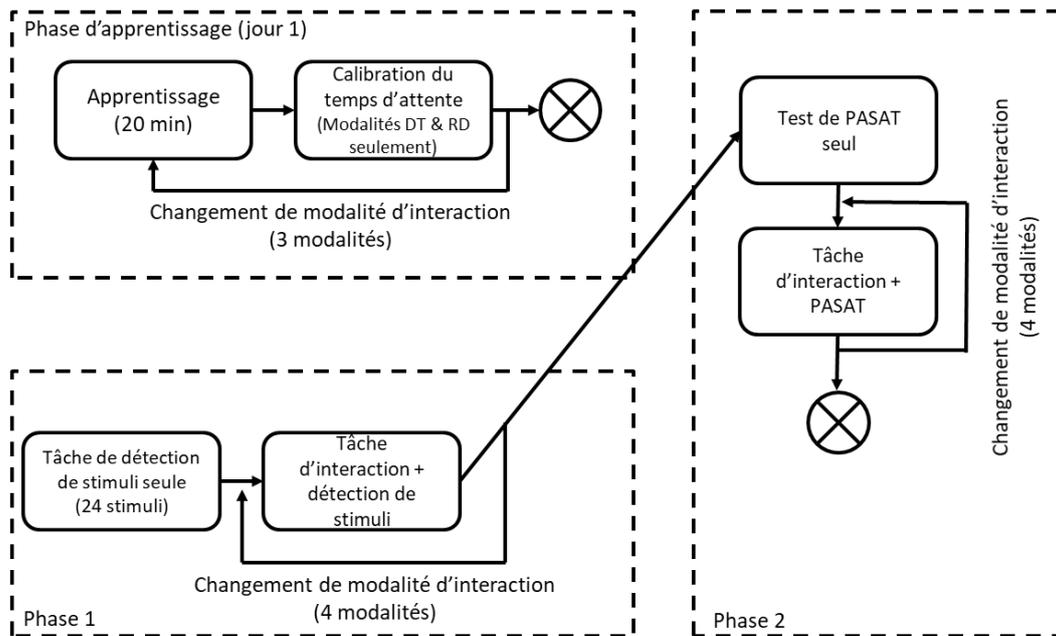


Figure 61. Diagramme représentant la chronologie de l'expérimentation 3.

### Variables mesurées

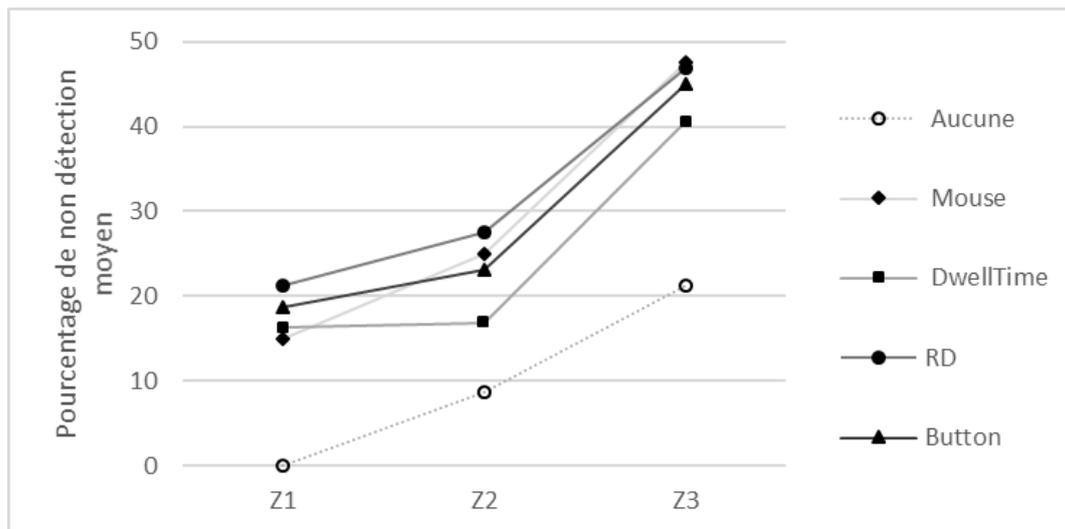
Lors de l'expérimentation 1, la variable mesurée est la moyenne du taux de détection de stimuli par zone. Les tests statistiques sont effectués en utilisant le test de Friedman.

Lors de l'expérimentation 2, la variable étudiée est la note sur 20 obtenue au test de PASAT. Cette note est représentative de la charge cognitive de la modalité d'interaction utilisée dans la tâche principale. Plus le score est haut, moins la modalité d'interaction est exigeante en termes de charge mentale. Les analyses statistiques sur ces notes sont aussi effectuées en utilisant le test de Friedman.

### 4.3.3 Résultats

#### Taux de non-détection en vision périphérique

L'expérimentation 1 a permis d'observer le pourcentage de non détection de stimuli visuels présentés dans le champ de vision périphérique. La Figure 62 résume ces résultats.



**Figure 62. Pourcentage d'erreur moyen de non-détection de stimuli en vision périphérique lors d'interaction avec différentes modalités d'interaction.**

L'utilisation du test de Friedman nous permet de montrer que les 4 modalités d'interaction engendrent une augmentation significative du taux de non-détection dans l'ensemble des zones de la vision périphérique (4 valeurs de  $p < 0.05$ ). La différence engendrée est d'environ 15-20% de non-détection supplémentaire.

Dans la zone de vision périphérique la plus centrale (zone 1), le test de Friedman ne montre aucune différence significative entre les résultats obtenus pour les quatre modalités d'interaction testées.

Dans les zones plus éloignées de la vision centrale, la modalité d'interaction *DwellTime* semble permettre une meilleure détection des alertes que les modalités d'interaction Souris et *DwellTime* délocalisé (RD sur la Figure 62). Une augmentation du taux d'erreur de quasiment 10% est observable sur ces deux modalités d'interaction par rapport à la modalité précitée. Ces différences sont significatives d'après les analyses post-hoc du test de Friedman (2 valeurs de  $p < 0.01$ ).

Aucune autre différence significative n'a été mise en évidence bien qu'il semble (cf. Figure 62) que l'interaction *DwellTime* permette d'obtenir de meilleurs résultats que l'interaction Bouton d'environ 5%.

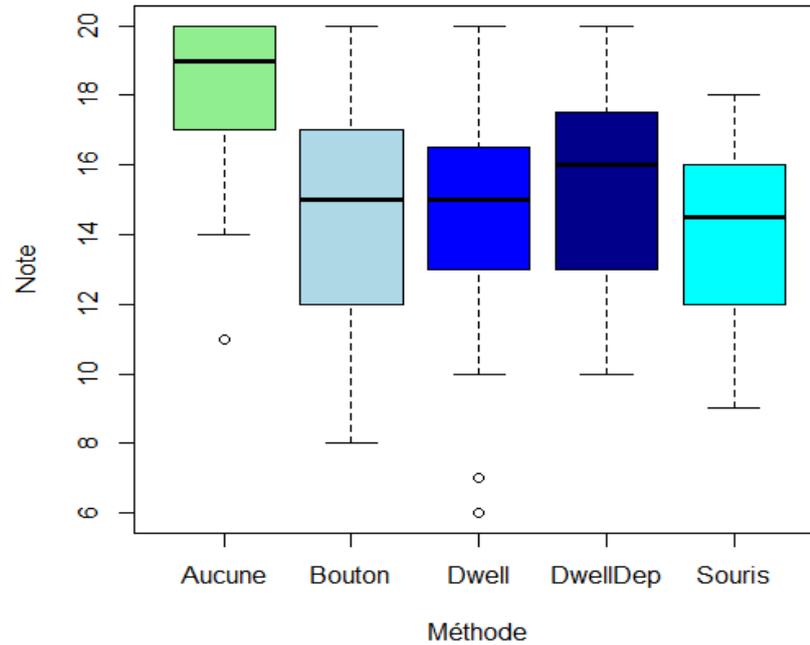
### **Charge cognitive**

L'utilisation du PASAT en double tâche permet d'obtenir des notes sur 20 représentatives de la charge cognitive associée à l'utilisation des différentes modalités d'interaction.

Lors de la réalisation de ce test sans interaction parallèle, les utilisateurs ont obtenu un score moyen de  $17,9 \pm 1,8$  sur 20. Les moyennes obtenues (cf. Figure 63) lors de la réalisation d'une tâche d'interaction avec l'une des quatre modalités d'interaction sont respectivement 14,2 pour la souris, 14,35 pour le *DwellTime*, 14,85 pour le Bouton et 15,3 pour le *DwellTime* délocalisé.

Le test de Friedman ne montre aucune différence significative entre ces résultats ( $p > 0.05$ ). On peut conclure à une charge cognitive relativement similaire pour l'utilisation des quatre modalités.

#### 4.3 - Effet de l'interaction au regard sur la capacité à détecter des signaux en vision périphérique



*Figure 63. Répartition des résultats obtenus au PASAT pour chaque modalité d'interaction. Aucun écart significatif entre les 4 modalités d'interaction.*

#### 4.3.4 Discussion

Afin de pouvoir inclure des modalités d'interaction au regard pour améliorer l'utilisabilité de systèmes, il est important d'étudier l'adéquation de l'interaction au regard avec les tâches effectivement utilisées dans ces systèmes. Cette expérimentation a cherché à étudier le comportement de différentes modalités d'interaction lorsque l'utilisateur délocalise son attention visuelle pendant une interaction au regard, ce qui est courant lors d'une tâche d'observation ou de surveillance.

Afin de valider la réalisation de certaines tâches avec le regard, il était nécessaire d'observer si l'utilisateur était toujours réactif à des alertes visuelles en vision périphérique lors de l'interaction. La première expérimentation mesurait cela alors que la seconde visait à mesurer la charge mentale pour comparer les résultats de la première à ce que l'on pourrait obtenir simplement à cause de l'effet tunnel.

La première conclusion concerne la deuxième expérimentation. Il semble que les trois modalités d'interaction au regard engendrent la même charge mentale lors de leur utilisation. De plus, cette charge cognitive est comparable à l'utilisation d'une modalité d'interaction classique, la souris.

Conclusion induite : le PASAT a été choisi comme méthode d'évaluation car la tâche associée ne faisait pas intervenir le regard. Ainsi, les résultats obtenus lors de l'autre expérimentation ne sont pas dus à la charge cognitive mais uniquement à des mécanismes visuels et attentionnels.

Dans la zone de vision périphérique la plus proche de la vision fovéale (zone 1 dans l'expérimentation), aucune différence significative entre les différentes modalités d'interaction n'a été observée. Etant donné que sans double tâche simultanée, les utilisateurs détectent 100% des stimuli visuels dans cette zone, Les 20% de non-détection sont bien dus à la tâche d'interaction.

Dans les zones 2 et 3, les résultats montrent que la modalité d'interaction *DwellTime* semble permettre d'obtenir un meilleur taux de détection d'alertes en vision périphériques que les autres modalités. Nous pensons que lors de la fixation nécessaire à l'interaction, le champ d'attention (la largeur du faisceau dans la métaphore du faisceau de lumière, cf. sous-section 2.3.3) s'élargit, permettant d'être plus « à l'écoute » (visuellement) de stimuli visuels. Ce résultat pourrait être mis en évidence par une expérimentation supplémentaire faisant apparaître les stimuli à l'instant précis de l'interaction, quand l'œil est sur l'item à activer.

Notamment, elle obtient de meilleurs résultats que la modalité *DwellTime* délocalisé. Contrairement au *DwellTime*, le *DwellTime* délocalisé demande de fixer un item qui n'est pas directement l'item à activer. Ainsi, le focus d'attention lors de l'interaction semble être l'item principal, réduisant la capacité de la méthode à élargir le champ attentionnel.

Résultats principaux de l'expérimentation :

- La modalité d'interaction *DwellTime* permet plus efficacement d'« ouvrir » inconsciemment son champ d'attention visuelle. Ceci n'est pas dû à un effet tunnel car :
- Les différentes modalités d'interaction engendrent des charges mentales similaires.

## 4.4 Conclusion

A travers trois expérimentations visant à étudier l'impact de certaines caractéristiques du contexte sur les modalités d'interaction au regard, nous avons pu mettre en évidence l'influence du contexte. Nous avons montré une influence limitée du type de menu sur les modalités testées, et mis en évidence une supériorité des modalités à base de temps de fixation (*DwellTime* notamment) lorsque la tâche utilisateur demandait de maximiser la perception visuelle en vision périphérique.

# Chapitre 5

## Conclusion

L'interaction au regard a fait l'objet de divers travaux depuis plusieurs années. Pour dépasser les problèmes de précision du regard et pour mettre au point des moyens techniques permettant de différencier analyse et intention d'interaction de la part de l'utilisateur, de nombreuses modalités d'interaction ont été proposées. Certaines études, réalisées dans des situations spécifiques, ont mis en évidence l'influence du contexte sur le fonctionnement des modalités d'interaction. Cependant, nous n'avons pas identifié de travaux s'intéressant précisément aux divers éléments du contexte de l'interaction au regard. Pourtant, lors de l'intégration d'une technologie dans des systèmes, l'analyse détaillée des situations qui vont être mises en œuvre réellement avec le système et de la façon dont celui-ci s'intégrera dans son environnement constitue une des premières étapes de réflexion des concepteurs industriels.

Cela nous a conduit à analyser les facteurs d'influence du contexte des modalités d'interaction au regard et la façon dont ils influencent les comportements des utilisateurs vis-à-vis de ces modalités. Poussé par la volonté industrielle d'une intégration du suivi du regard dans des systèmes déjà existants, nous nous sommes intéressés plus précisément au cas d'une interaction au regard avec des menus, éléments déjà présents dans les systèmes actuels.

Après une analyse du problème, nous avons adopté deux approches, menant à nos deux types de contributions, pour comprendre l'influence du contexte sur les modalités d'interaction au regard. La première approche a été de proposer un modèle du contexte des modalités d'interaction au regard. La seconde approche, expérimentale, s'est centrée sur des points précis de ce contexte, justifiée par les besoins industriels immédiats.

## 5.5 Contributions

Nous résumons nos contributions selon nos deux approches : descriptive et expérimentale.

### 5.5.1 Proposition d'un modèle du contexte des modalités de l'interaction au regard

A l'instar de modèles du contexte de systèmes ou de modèles du contexte de la tâche, nous avons proposé un modèle du contexte de l'interaction active au regard. Nous avons découpé ce modèle selon 4 axes qui sont : l'utilisateur, la tâche utilisateur, la plateforme et l'environnement. Chacun de ces axes correspond à plusieurs caractéristiques. Bien que des liens existent entre les différentes caractéristiques proposées, notre approche est d'essayer de les considérer de manière indépendante.

L'intérêt scientifique de cette modélisation est de proposer une formalisation détaillée qui vise à mieux comprendre comment une interaction au regard s'intègre à des situations. Chaque caractéristique abordée ouvre un axe de recherche nouveau sur l'influence de cette caractéristique sur l'interaction au regard.

Pour modéliser l'utilisateur, nous avons choisi une modélisation par caractéristique plutôt que par stéréotype afin de différencier les différents aspects qui caractérisent l'utilisateur et de les traiter indépendamment. Nous avons abordé l'expérience utilisateur, son état physique et psychologique et ses caractéristiques individuelles (talents propres).

La tâche utilisateur, qui est distincte de la tâche d'interaction, représente les actions réalisées en dehors du fait d'interagir simplement au regard. L'interaction n'est souvent qu'un moyen pour atteindre un but qui est la tâche utilisateur. Nous avons abordé plusieurs types de tâches visuelles comme la supervision, la réponse à des événements visuels et le contrôle visuel, ainsi que des tâches motrices non visuelles.

La plateforme, parfois directement appelée système dans certains modèles, représente l'ensemble des ressources logicielles et matérielles dont dispose le système pour l'interaction. Aborder les implications des performances des systèmes de suivi du regard ainsi que les choix d'interfaces graphiques nous a permis de montrer que les utilisateurs ne sont pas passifs vis-à-vis des modalités d'interaction au regard et qu'ils peuvent faire des choix qui influencent les performances des modalités d'interaction au regard.

Enfin, l'environnement regroupe les caractéristiques les plus extérieures au système comme la luminosité ou le contexte social dans lequel la tâche utilisateur est réalisée. L'environnement a été moins étudié dans la mesure où il paraît plus éloigné du système et de l'interaction, mais son importance est loin d'être négligeable.

Ce modèle structure le contexte et permet d'isoler une caractéristique à la fois, ce que nous avons fait au cours de nos expérimentations.

## 5.5.2 Expérimentations sur des critères pertinents pour les systèmes optroniques

Les besoins directs de l'intégration de modalité d'interaction au regard dans des systèmes optroniques nous ont conduits à étudier plus précisément certains aspects du contexte. Nous avons étudié l'adéquation des modalités d'interaction au regard avec des menus linéaires et circulaires, puis nous nous sommes intéressés aux capacités perceptives en vision périphérique.

La première expérimentation a été motivée par le design existant des menus dans les systèmes optroniques et a visé à évaluer les performances de différentes modalités sur des menus linéaires et circulaires semblables à ceux des produits Safran. Nous avons montré que, dans ce cas, la forme du menu n'avait pas d'impact sur les différentes modalités testées. Nous avons également constaté la complexité d'utilisation du *DwellTime* délocalisé, une modalité que nous avons proposée. Enfin, nous avons mis en évidence que l'interaction *DwellTime* n'est pas adaptée à des menus dont les différents niveaux se superposent.

La seconde expérimentation que nous avons réalisée a permis de mettre en opposition le fait d'interagir au regard avec notre capacité à déporter notre attention visuelle en vision périphérique. Un protocole expérimental en deux parties (nécessitant de la part de l'utilisateur respectivement la priorisation de la tâche d'attention divisée puis de la tâche d'interaction puis) nous a permis de montrer que les modalités d'interaction à base de temps de fixation semblent plus aptes à permettre à l'utilisateur ce déport. Ces modalités paraissent donc plus adaptées aux tâches réalisées dans les systèmes optroniques.

La troisième expérimentation concernait elle aussi la vision périphérique lors d'interactions au regard, mais l'attention était cette fois attirée de manière exogène, par des stimuli inattendus apparaissant en zone périphérique. Nous avons mis en évidence que l'interaction *DwellTime* laissait l'utilisateur plus apte à capter ce genre de stimuli que les autres modalités d'interaction testées. Nous avons aussi montré que ce résultat n'était pas directement lié à l'effet tunnel dû à la charge cognitive induite par les modalités d'interaction.

## 5.5.3 Contribution industrielle

Les expérimentations ont permis d'obtenir des résultats directement exploitables pour l'intégration dans des systèmes optroniques. Ils mettent en évidence une meilleure capacité de perception en vision périphérique lors d'interactions *DwellTime*. Cette modalité serait en effet à privilégier si cette faculté était prioritaire par rapport à l'efficacité dans la tâche d'interaction.

Le modèle développé dans ce manuscrit peut être utilisé de deux manières. Dans un premier temps, il permet de lister un ensemble de critères pouvant avoir une influence sur les modalités d'interaction au regard. Lors de l'établissement du cahier des charges d'un produit, celui-ci peut être comparé au modèle du contexte pour en déduire la modalité d'interaction au regard qui serait la plus adéquate. Par ailleurs, il constitue aussi un guide sur les différents aspects à étudier lors de la réalisation de tests sur les produits futurs.

## 5.6 Perspectives

Bien que nous ayons considéré les critères du contexte de manière indépendante, un des premiers travaux à réaliser serait de vérifier les résultats obtenus avec des utilisateurs entraînés. Il serait intéressant d'étudier si les tendances obtenues avec des utilisateurs novices seraient transférables à ce cas plus représentatif de la réalité d'utilisation des produits.

Pour l'étude de modalités d'interaction au regard sur des menus linéaires et circulaires, nous nous sommes limités à des modalités « classiques » utilisables sur les deux types de menus à des fins de comparaison. Cependant les menus circulaires permettent des modalités d'interaction originales telles que celle qui a été proposée par Urbina et ses collègues (Urbina, Lorenz, & Huckauf, 2010), qui peuvent présenter des avantages par rapport aux modalités d'interaction au regard que nous avons étudiées dans nos travaux.

La complexité des menus (nombre d'items, profondeur) a été fixée durant nos expérimentations. Il serait intéressant d'envisager des complexités différentes pour généraliser les résultats obtenus. Encore une fois, nous avons considéré ce critère de manière indépendante, il nous semble pourtant que certains feedbacks ou techniques comme l'utilisation d'une taille d'item dynamique comme feedback visuel (Miniotas, Spakov, & MacKenzie, 2004) ne s'applique pas de la même manière sur les deux types de menus.

Concernant le nouveau protocole de mesure de la capacité à séparer son attention visuelle que nous avons proposé, les résultats sont encourageants. Cependant, notre méthode pour savoir si un utilisateur réussissait ou non à diviser son attention nous semble encore relativement sommaire. Il faudrait l'améliorer en utilisant des méthodes statistiques par exemple.

D'autres critères évoqués dans le modèle nous semblent encore pertinents à étudier, notamment, l'expertise utilisateur ainsi que l'effet de la fatigue et du stress, les missions opérationnelles étant souvent longues et parfois pleines de situations inattendues voire dangereuses.

De manière générale, négliger les effets croisés des différents critères est une approche trop restrictive. Une des forces de notre modèle est son utilisation générative afin d'appréhender la tâche complexe d'analyse du contexte et de son influence sur les modalités d'interaction au regard. Considérer des couples, ou des trios de critères permet de se rapprocher de situations réelles tout en ayant le contrôle des éléments testés.

Enfin, intégrer une nouvelle modalité d'interaction dans des produits déjà utilisés peut susciter une certaine réticence de la part des utilisateurs finaux. Et si nos résultats nous semblent prometteurs, faire accepter ce changement est un enjeu aussi complexe que de concevoir le système.

Comme nous l'avons souligné, les différentes modalités d'interaction disposent d'avantages et d'inconvénients spécifiques à chaque situation. Nous voudrions donc étudier les possibilités de cohabitation de plusieurs modalités dans un même système. Cette cohabitation pourrait être optionnelle, laissée au choix de l'utilisateur, ou le système pourrait analyser dynamiquement le contexte pour déterminer les modalités d'interaction utilisables ou les paramètres de ces modalités. Par exemple, le temps d'attente nécessaire à l'interaction *DwellTime*, pourraient être déterminé en fonction des caractéristiques de fatigue ou de stress de l'utilisateur, à l'instar des travaux initiés par

Spakov et Miniotas (Spakov & Miniotas, 2004). Ou encore, le type de feedback visuel pourrait être fonction de la complexité visuelle de la scène observée, afin de permettre à l'utilisateur de garder un certain niveau de concentration sur l'interface. La modification dynamique d'une interface ou d'une interaction peut être une décision de l'utilisateur ou une décision automatique du système. On rejoint dans ce cas les problématiques d'interface plastique (Calvary G. , et al., 2003).

Un autre axe que nous aimerions voir étudié serait l'impact de la transformation de l'œil en organe sensorimoteur sur les fonctions cognitives de l'utilisateur : comment ce changement modifie-t-il les capacités de mémorisation de l'utilisateur ou de récupération d'informations mémorisées ? Ou encore comment ce changement agit-il sur les fonctions motrices simultanées : le fait d'interagir au regard ne modifie-t-il pas notre perception de l'espace et par là-même nos actions motrices dans celui-ci ? Gardons un œil là-dessus...



# Glossaire

**Activité** : le LAROUSSE propose cette définition : « Ensemble de phénomènes par lesquels se manifestent certaines formes de vie, un processus, un fonctionnement. » Dans certains cas, l'activité est équivalente à la tâche utilisateur, c'est-à-dire l'ensemble des actions entreprises par un utilisateur. Dans certains cas, le terme activité est utilisé dans un sens plus large pour décrire l'ensemble <Utilisateur, Tâche, Système>, c'est-à-dire la tâche, l'entité qui la réalise et les outils nécessaires à sa réalisation. Dans le document, le terme activité est utilisé au sens large, en opposition avec la tâche utilisateur.

**Axe optique de l'œil** : centre de symétrie du globe oculaire, aligné sur le centre de la pupille, de la lentille.

**Axe visuel** : ne pas confondre avec l'axe optique de l'œil. L'axe visuel est l'axe passant par le centre de la pupille et le centre de la fovea.

**Contexte** : le contexte d'une entité est l'ensemble des éléments extérieurs à cette entité qui sont pertinents pour sa description.

**Environnement** : il existe souvent un abus de langage entre contexte et environnement. La définition « Ce qui entoure de tous côtés » ne restreint pas l'environnement aux éléments pertinents à l'analyse d'une entité. Contrairement au contexte qui, on l'a vu est limité par sa pertinence vis-à-vis de l'entité étudiée. En IHM, l'environnement est l'ensemble des éléments extérieurs à l'activité au sens large, c'est-à-dire tout sauf l'utilisateur, le système et la tâche. Souvent, une partie de l'environnement est considéré dans le contexte comme la localisation ou l'heure de la journée.

**Fovea** : zone de l'œil disposant du plus grand nombre de récepteurs photosensibles. Correspond à la zone d'acuité visuelle maximale.

**Gaze aversion** : phénomène impliquant que lors de l'utilisation de la mémoire, le regard se dirige instinctivement vers les zones de moindre complexité visuelle.

**Medium d'interaction** : dispositif physique permettant la communication entre un système et un utilisateur. Pour permettre cette communication, le medium d'interaction est associé à un langage d'interaction (cf. Modalité d'interaction).

**Midas Touch** : fait de ne pas pouvoir discerner les phases d'analyse de l'œil des phases comportant une intention d'interaction. Implique des difficultés d'interaction.

**Modalité d'interaction** : d'après Vernier (Vernier & Nigay, 2000), une modalité d'interaction est définie comme un couple <ystème physique, langage d'interaction>. Le système physique (medium d'interaction) représente évidemment les outils utilisés pour l'interaction (souris, clavier, système de suivi du regard) alors que le langage d'interaction décrit comment le système est utilisé pour générer de l'information, le plus souvent pour un système informatique (ex : double-clic, appui sur une touche, clin d'œil). Dans ce document, lorsque ce n'est pas précisé, le terme modalité peut à la fois déterminer une modalité simple ou une modalité multiple (on parle de multimodalité), c'est-à-dire utilisant une combinaison de plusieurs systèmes physiques et de plusieurs langages d'interaction.

**Tâche d'interaction** : ensemble des actions effectuées pour communiquer avec le système.

**Tâche utilisateur** : en opposition à la tâche d'interaction, la tâche utilisateur représente l'ensemble des actions effectuées excepté celles de la tâche d'interaction. Aussi appelée tâche de fond.

**Plateforme** : défini par Calvary (Calvary G. , et al., 2003) comme les moyens physiques et logiciels qui permettent à un utilisateur de communiquer avec un système. La plateforme fait partie du système.

**Précision (*accuracy*)** : variance des échantillons de positions obtenus à l'aide d'un système de suivi du regard.

**Précision (*precision*)** : une faible *precision* est associée à un décalage entre la moyenne des échantillons de position obtenus à l'aide d'un système de suivi du regard et la position réellement fixée.

# Annexes

Annexe A : Questionnaire d'utilisabilité de l'expérimentation 1 concernant l'impact du type de menu.

Annexe B : Menus utilisés dans la première expérimentation

Annexe C : Adaptation du questionnaire UMUX à l'évaluation d'une modalité d'interaction au regard.



## **Annexe A : Questionnaire d'utilisabilité de l'expérimentation 1 concernant l'impact du type de menu.**

Le questionnaire d'utilisabilité que nous avons soumis à l'utilisateur lors de la première expérimentation concernant l'impact du type de menu sur les modalités d'interaction au regard est présenté page suivante. Ce questionnaire nous a permis de collecter des informations personnelles sur l'utilisateur ainsi que son ressenti sur l'ensemble des couples <Modalité-Type de Menu>. Le questionnaire était rempli par le superviseur.

Parallèlement à ce questionnaire, l'utilisateur était poussé à verbaliser son ressenti à la fin de l'utilisation de chaque modalité d'interaction.

Annexe C Adaptation du questionnaire UMUX à l'évaluation d'une modalité d'interaction au regard.

ID Sujet : \_\_\_\_\_ Nom : \_\_\_\_\_ Heure : \_\_\_\_\_  
 Age : \_\_\_\_\_ Sexe : \_\_\_\_\_  
 Avez-vous déjà utilisé un système de commandes au regard ? Oui / Non  
 Pratiquez-vous du sport ? Lesquels ? \_\_\_\_\_  
 Jouez-vous aux jeux vidéo ? Quel genre ? \_\_\_\_\_  
 Lisez-vous beaucoup ? \_\_\_\_\_  
 Quelles autres activités pratiquez-vous ?? \_\_\_\_\_

---

Ordre des méthodes : \_\_\_\_\_

Quelles sont vos remarques sur la méthode que vous venez d'utiliser ? Décrivez-la.

Sur une échelle de 1 à 7, la méthode vous à t-elle parue intuitive ?

Très intuitive

Pdt intuitive

Sur une échelle de 1 à 7, quelle quantité d'effort vous a demandé la méthode à la prise en main ?

Peu d'effort

Bcp d'effort

Sur une échelle de 1 à 7, la méthode vous à t-elle parue rapide ?

Très rapide

Pdt rapide

Sur une échelle de 1 à 7, quelle quantité d'effort vous a demandé la méthode une fois la prise en main terminée ?

Peu d'effort

Bcp d'effort

Sur une échelle de 1 à 7, la méthode vous à t-elle parue fiable (peu d'erreur) ?

Très fiable

Pdt fiable

Note d'appréciation (sur 20) ? \_\_\_\_\_

De manière générale, étiez-vous à l'aise avec le fait d'utiliser votre regard pour interagir ?

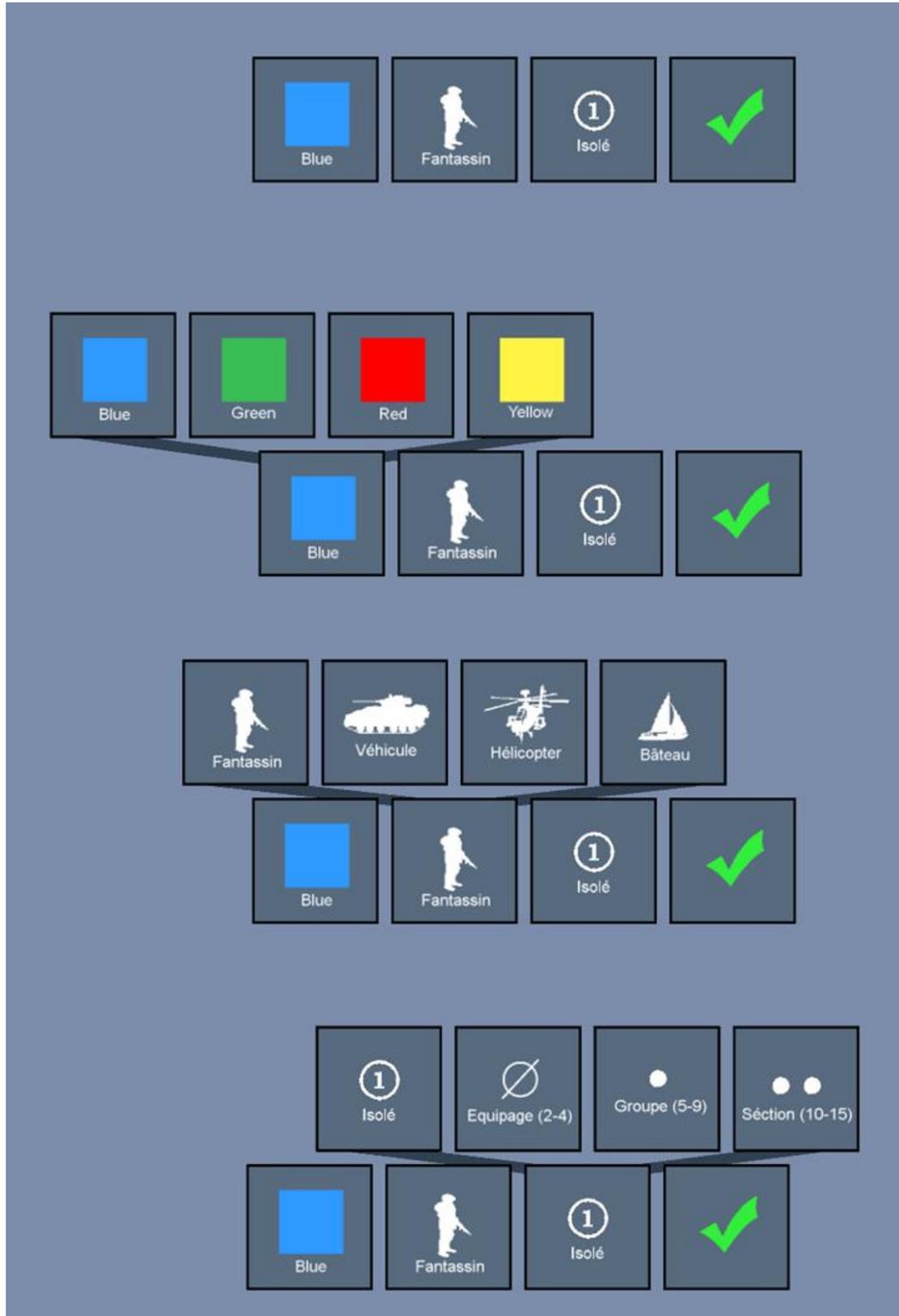
Classement

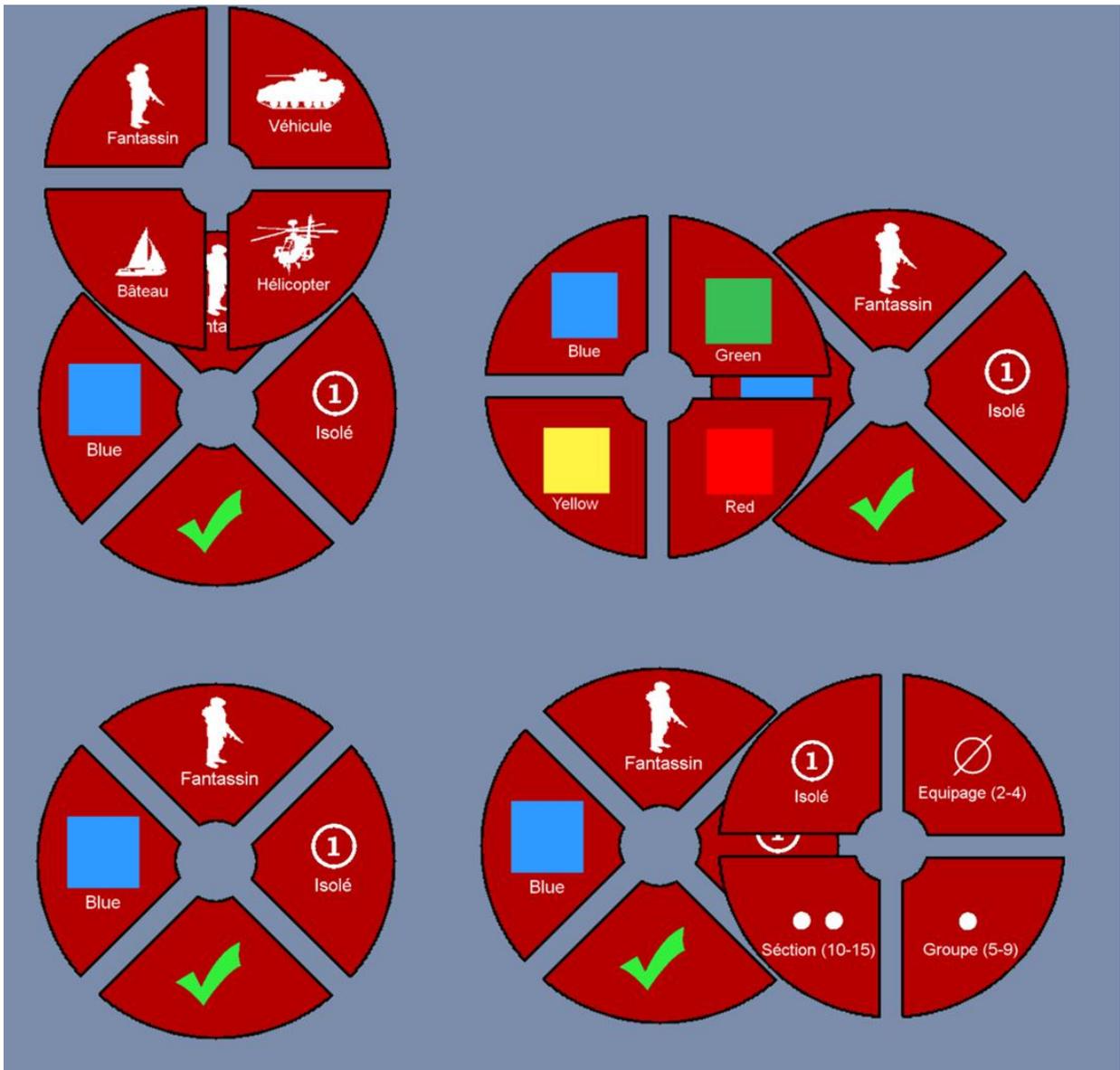
	C DD	C Dep	C Bou	C Inst	L DD	L Dep	L Bou	L Inst
Utilisable								
Classement								

Figure A-1 : Questionnaire d'utilisabilité de l'expérimentation 1 sur l'impact du type de menu.



## Annexe B : Menus utilisés dans la première expérimentation







## Annexe C : Adaptation du questionnaire UMUX à l'évaluation d'une modalité d'interaction au regard.

Nom :		Id :	
Grille	Temps <i>DwellTime</i> :	ColorShape	Temps <i>DwellTime</i> :
	Temps <i>DwellTime</i> Délocalisé :		Temps <i>DwellTime</i> Délocalisé :
Distance œil-écran			
Remarques			
Phase d'apprentissage			
Xp1 : Lettres			
Xp2 : Flèches			

EST-CE QUE LA DOUBLE TACHE VOUS A POSEE PROBLEME ? (VERBALISATION)

OBSERVATIONS REMPLISSAGE QUESTIONNAIRE

**Consigne générale :** Gardez à l'esprit que vos réponses portent sur les méthodes d'interaction au regard testées dans le cadre d'une tâche prioritaire de délocalisation de l'attention, et non sur le simple fait d'utiliser le regard pour interagir

## Annexe C Adaptation du questionnaire UMUX à l'évaluation d'une modalité d'interaction au regard.

Tâche de délocalisation de l'attention et tâche d'interaction avec la méthode « Souris »		1=Pas du tout d'accord 7=Tout à fait d'accord						
		1	2	3	4	5	6	7
1.	La méthode « Souris » me permet d'interagir tout en déportant mon attention.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2.	Utiliser la méthode « Souris » tout en déportant mon attention est une expérience désagréable.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3.	Avec une double tâche de déportation d'attention, la méthode « Souris » est facile à utiliser.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
4.	En déportant mon attention, je n'ai pas été efficace sur ma tâche d'interaction avec la méthode « Souris ».	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
5.	Avec la méthode « Souris », j'ai eu confiance dans ma capacité à interagir tout en déportant mon attention.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
6.	Tout en déportant mon attention, mon interaction m'a semblée plutôt lente avec la méthode « Souris ».	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Tâche de délocalisation de l'attention et tâche d'interaction avec la méthode « Dwell-Time »		1=Pas du tout d'accord 7=Tout à fait d'accord						
		1	2	3	4	5	6	7
1.	La méthode « Dwell-Time » me permet d'interagir tout en déportant mon attention.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2.	Utiliser la méthode « Dwell-Time » tout en déportant mon attention est une expérience désagréable.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3.	Avec une double tâche de déportation d'attention, la méthode « Dwell-Time » est facile à utiliser.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
4.	En déportant mon attention, je n'ai pas été efficace sur ma tâche d'interaction avec la méthode « Dwell-Time ».	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
5.	Avec la méthode « Dwell-Time », j'ai eu confiance dans ma capacité à interagir tout en déportant mon attention.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
6.	Tout en déportant mon attention, mon interaction m'a semblée plutôt lente avec la méthode « Dwell-Time ».	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Tâche de délocalisation de l'attention et tâche d'interaction avec la méthode « Dwell-Time Délocalisé »		1=Pas du tout d'accord 7=Tout à fait d'accord						
		1	2	3	4	5	6	7
1.	La méthode « Dwell-Time délocalisé » me permet d'interagir tout en déportant mon attention.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2.	Utiliser la méthode « Dwell-Time délocalisé » tout en déportant mon attention est une expérience désagréable.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3.	Avec une double tâche de déportation d'attention, la méthode « Dwell-Time délocalisé » est facile à utiliser.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
4.	En déportant mon attention, je n'ai pas été efficace sur ma tâche d'interaction avec la méthode « Dwell-Time délocalisé ».	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
5.	Avec la méthode « Dwell-Time délocalisé », j'ai eu confiance dans ma capacité à interagir tout en déportant mon attention.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
6.	Tout en déportant mon attention, mon interaction m'a semblée plutôt lente avec la méthode « Dwell-Time délocalisé ».	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Tâche de délocalisation de l'attention et tâche d'interaction avec la méthode « Bouton »		1=Pas du tout d'accord 7=Tout à fait d'accord						
		1	2	3	4	5	6	7
1.	La méthode « Bouton » me permet d'interagir tout en déportant mon attention.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2.	Utiliser la méthode « Bouton » tout en déportant mon attention est une expérience désagréable.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3.	Avec une double tâche de déportation d'attention, la méthode « Bouton » est facile à utiliser.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
4.	En déportant mon attention, je n'ai pas été efficace sur ma tâche d'interaction avec la méthode « Bouton ».	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
5.	Avec la méthode « Bouton », j'ai eu confiance dans ma capacité à interagir tout en déportant mon attention.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
6.	Tout en déportant mon attention, mon interaction m'a semblée plutôt lente avec la méthode « Bouton ».	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

## Annexe D : Implémentation.

Cette annexe précise certains points sur la méthodologie employée pour l'implémentations des interactions au regard lors de nos travaux. Elle n'a pas vocation à être exhaustive, mais à justifier l'implémentation réalisée ainsi que spécifier des patterns utilisés.

### 1. Les données d'entrée

L'implémentation de nos expérimentations, et donc de nos interfaces a été effectuée sur l'environnement de développement Unity (v 5.5) en C#. L'environnement de développement permet une mise en place des éléments graphiques, ainsi que des composants liés à chaque item. Le comportement logique des éléments est spécifié dans des scripts ensuite liés aux items de l'environnement de développement.

Unity fournit un SDK permettant l'intégration du Tobii Eye-X avec Unity.<sup>9</sup> Parmi l'ensemble de fonctions permettant de contrôler le système de suivi du regard, nous avons principalement utilisé la méthode permettant de savoir si un élément est actuellement regardé ou non. Tous les objets susceptibles d'être interactifs au regard disposent d'un composant « GazeAware » qui permet d'accéder à la propriété « HasGazeFocus », booléen représentant le fait que l'objet est actuellement celui vers lequel se dirige le regard de l'utilisateur. Afin de fonctionner, ce composant nécessite aussi que l'objet dispose d'un collider 3D représentant physiquement la taille et la forme de l'objet, indépendamment de la texture utilisée. Notre interface 2D a donc été associée à une dimension z arbitraire pour permettre l'utilisation de cette fonctionnalité.

Cependant, notre développement des expérimentations nécessitait aussi une interaction à la souris. Afin d'homogénéiser l'architecture, un niveau d'abstraction supplémentaire a été ajouté permettant de savoir si un item a le « focus », que cela soit à la souris, ou au regard. L'information est alors déléguée au composant de Bouton représentant le type de bouton représenté.

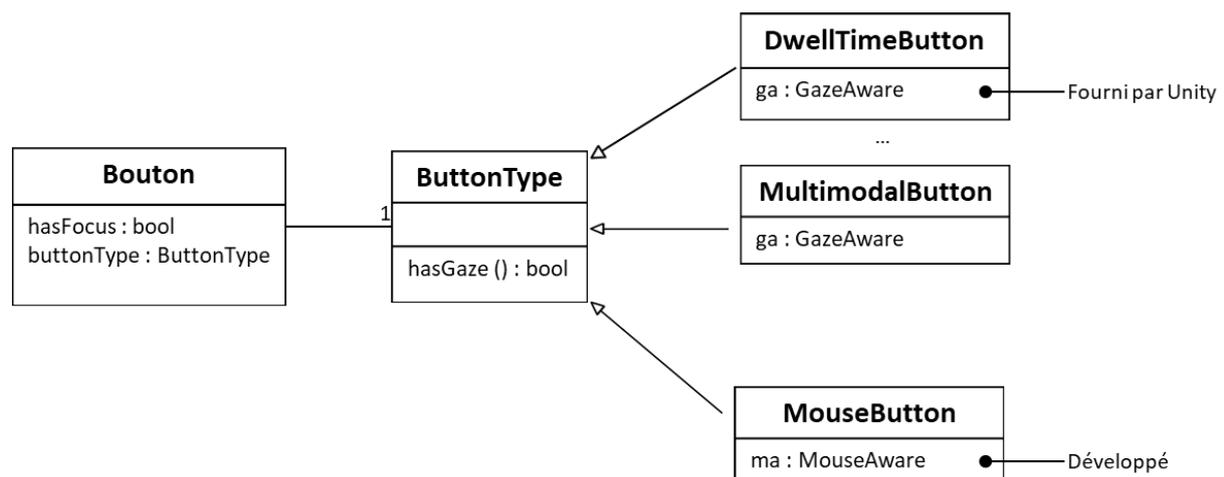


Diagramme UML spécifiant la manière d'obtenir le « focus » d'un item.

<sup>9</sup> <https://developer.tobii.com/tobii-unity-sdk/>

## 2. Implémentation des types d'items et de la logique d'interaction

Chaque modalité d'interaction est représentée par un item disposant d'une logique d'interaction différente. Pour chaque logique d'interaction (c.a.d chaque type de modalité d'interaction), une classe a été développée. Ce sont les classes *DwellTimeButton*, *MultimodalButton*, *InstantButton*, *RelocatedDwellTimeButton* (basée elle-même sur la classe *DwellTimeButton* pour l'élément déporté), et *MouseButton*.

Ces classes gèrent les notions de feedback visuels, de leur reset lorsque l'élément n'est plus regardé pendant un certain temps, de temps de fixation, d'erreur acceptable de synchronisation œil-main, etc. En fonction des actions de l'utilisateur, un bouton peut se retrouver dans des états complexes. Par exemple il a été regardé pendant un temps, laissé de côté, puis le regard s'est reposé sur l'item. Pour simplifier l'écriture et la lecture du code, nous avons utilisé le « State Pattern »<sup>10</sup>. Comme les actions à prendre en compte sont dépendante de l'état de l'item (neutre, regardé, activé, survolé), une classe état permet de déléguer le comportement de manière dynamique.

Une pseudo-implémentation de la classe *ButtonType* ressemble à ça :

```
class ButtonType
{
    state : AbstractState

    // Other code...

    void handleFocus (bool hasGaze)
    {
        state-> handleFocus (&this, hasGaze);
    }

    // Other code...
}
```

De manière générique, une classe état suit la logique suivante :

```
class State : AbstractState
{
    // variables et fonctions propres à un état

    void handleFocus (ButtonType* buttonType, bool hasGaze)
    {
        // Logic code...

        buttonType->state = new OtherState() ;

        // Logic code...
    }
}
```

---

<sup>10</sup> <http://www.gameprogrammingpatterns.com/state.html>

Le changement d'état se fait par la création d'une nouvelle instance d'état, pouvant disposer de paramètres propres à l'état (comme le temps de fixation par exemple). De plus, certains états disposent de comportements similaires pour plusieurs types de boutons, la classe état est alors réutilisable tel quel.

### 3. Implémentation de la logique de l'interface et des actions liées aux différents items.

La fonctionnalité « hasGaze » fourni par Tobii (et par extension « hasFocus » développée par nos soins ne suffit pas à l'intégration facile dans une interface. Par exemple, si l'on regarde un item d'un sous-menu, on désire que ce menu reste ouvert. Avec la fonctionnalité de Tobii, un seul élément peut avoir le focus à la fois, il a donc été nécessaire d'implémenter les mécanismes liés à l'interface permettant la logique entre les différents niveaux de menus, l'apparition et disparition des items en fonction etc.

La classe Bouton (cf. Diagramme 1) dispose de la logique liée au reste de l'interface et aux actions qu'engendre l'activation du bouton. Elle gère l'intégration et le lien avec les autres items d'un même niveau de menu, elle possède une liste de pointeurs de fonction à activer lors de son activation, gère le changement potentiel de type du bouton etc.

<b>Bouton</b>
menuLayer : Menu boutonType : ButtonType UnityEvent : command enableFeedBack : bool isActive : bool hasFocus : bool
Activate() setActive(bool) changeButtonType() getMenuLayer() setMenuLayer() Reset()

La complexité de la gestion de l'implémentation de l'expérimentation a ensuite été de définir l'ensemble des commandes liées à l'activation d'un élément, c'est-à-dire à la définition complète de la variable *command*. L'ensemble de la logique liée à l'activation d'un élément pouvait alors être effectuée via l'environnement de développement Unity.



# Bibliographie

- Awh, E., Pashler, H., & others. (2000). Evidence for split attentional foci. *Journal of Experimental Psychology-Human Perception and Performance*, 26(2), (pp. 834-846).
- Baas, M., Bernonville, S., Bricon-Souf, N., Hassler, S., Kolski, C., & Boy, G. A. (2014). Towards a context model for human-centered design of contextual data entry systems in healthcare domain. *International Conference on Engineering Psychology and Cognitive Ergonomics*, (pp. 223-233).
- Bailly, G. (2009). *Techniques de menus: caractérisation, conception et évaluation*. Ph.D. dissertation, Université Joseph-Fourier (Grenoble).
- Bazire, M., & Brézillon, P. (2005). Understanding context before using it. *International and Interdisciplinary Conference on Modeling and Using Context* (pp. 29-40). Springer, Berlin, Heidelberg.
- Bednarik, R., Gowases, T., & Tukiainen, M. (2009). Gaze interaction enhances problem solving: Effects of dwell-time based, gaze-augmented, and mouse interaction on problem-solving strategies and user experience. *Journal of Eye Movement Research*, 3(1), 1-10.
- Bee, N., & André, E. (2008). Writing with your eye: A dwell time free writing system adapted to the nature of human eye gaze. *International Tutorial and Research Workshop on Perception and Interactive Technologies for Speech-Based Systems* (pp. 111-122). Springer, Berlin, Heidelberg.
- Berger, A., Henik, A., & Rafal, R. (2005). Competition between endogenous and exogenous orienting of visual attention. *Journal of Experimental Psychology-General*, 134(2), (pp. 207–221).
- Bolt, R. A. (1980). "Put-that-there": Voice and gesture at the graphics interface (Vol. 14, No. 3, pp. 262-270). ACM.
- Bolt, R. A. (1982). Eyes at the interface. *Proceedings of the 1982 conference on Human factors in computing systems* (pp. 360-362). ACM.
- Bordwell, D., Thompson, K., & Smith, J. (2016). *Film art: An introduction*. New, York: McGraw-Hill Education.

## Bibliographie

- Bosley, J. J. (2013). Creating a short usability metric for user experience (UMUX) scale. *Interacting with Computers*, 25(4), (pp. 317-319).
- Bowman, D. (2000). The science of interaction design. *ACM SIGGRAPH Course Notes*.
- Boy, G. A. (2011). *The handbook of human-machine interaction: a human-centered design approach*. Ashgate.
- Brandt, T., Dichgans, J., & Koenig, E. (1973). Differential effects of central versus peripheral vision on egocentric and exocentric motion perception. *Experimental brain research*, 16, (pp. 476-491).
- Bransford, J. D., Brown, A. L., & Cocking, R. R. (2004). *How people learn (Expanded ed.)*. Washington, DC: National Academy of Sciences.
- Brézillon, P. (2007). Context modeling: Task model and practice model. *International and Interdisciplinary Conference on Modeling and Using Context* (pp. 122-135). Springer, Berlin, Heidelberg.
- Brusilovsky, P., & Millan, E. (2007). User models for adaptive hypermedia and adaptive educational systems. Dans *The adaptive web* (pp. 3-53). Springer, Berlin, Heidelberg.
- Callahan, J., Hopkins, D., Weiser, M., & Shneiderman, B. (1988). An empirical comparison of pie vs. linear menus. *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 95-100). ACM.
- Calvary, G., Coutaz, J., & Thevenin, D. (2001). A unifying reference framework for the development of plastic user interfaces. Dans *Engineering for human-computer interaction* (pp. 173-192). Springer, Berlin, Heidelberg.
- Calvary, G., Coutaz, J., Thevenin, D., Limbourg, Q., Bouillon, L., & Vanderdonckt, J. (2003). A unifying reference framework for multi-target user interfaces. *Interacting with computers*, 15(3), (pp. 289-308).
- Carpenter, R. H. (1988). *Movements of the Eyes, 2nd Rev.* Pion Limited.
- Cazzoli, D., Antoniadis, C. A., Kennard, C., Nyffeler, T., Bassetti, C. L., & Müri, R. M. (2014). Eye movements discriminate fatigue due to chronotypical factors and time spent on task--a double dissociation. *PloS one*, 9(1), e87146.
- Cooper, L., Gale, A., Darker, I., Toms, A., & Saada, J. (2009). Radiology image perception and observer performance: how does expertise and clinical information alter interpretation? Stroke detection explored through eye-tracking. *Medical Imaging 2009: Image perception, observer performance, and technology assessment*. 7263. Loughborough: International Society for Optics and Photonics.
- Corbetta, M., Patel, G., & Shulman, G. L. (2008). The reorienting system of the human brain: from environment to theory of mind. *Neuron*, 58(3), (pp. 306-324).
- Coutaz, J., & Calvary, G. (2012, 5). HCI and Software Engineering for User Interface Plasticity. Dans J. A. Jacko (Éd.), *Human-Computer Interaction Handbook: Fundamentals, Evolving*

- Technologies, and Emerging Applications, Third Edition* (pp. 1195-1220). CRC Press. Récupéré sur <https://hal.archives-ouvertes.fr/hal-00752122>
- Crowley, J. L., Coutaz, J., Rey, G., & Reignier, P. (2002). Perceptual components for context aware computing. *International conference on ubiquitous computing* (pp. 117-134). Springer, Berlin, Heidelberg.
- Dey, A. K. (2001). Understanding and using context. *Personal and ubiquitous computing*, 5(1), (pp. 4-7).
- Dey, A. K., & Abowd, G. D. (2000). *Providing architectural support for building context-aware applications*. Ph.D. dissertation, College of Computing, Georgia Institute of Technology.
- Drewes, H. (2010). *Eye gaze tracking for human computer interaction*. Ph.D. dissertation, Imu.
- Drewes, H., & Schmidt, A. (2007). Interacting with the computer using gaze gestures. *IFIP Conference on Human-Computer Interaction* (pp. 475-488). Springer, Berlin, Heidelberg.
- Dubois, J., Hamker, F. H., & VanRullen, R. (2009). Attentional selection of noncontiguous locations: The spotlight is only transiently "split". *Journal of Vision*, 9(5), 1-11.
- Duchowski, A. T. (2007). Eye tracking methodology. *Theory and practice*, 328.
- Epelboim, J., & Suppes, P. (2001). A model of eye movements and visual working memory during problem solving in geometry. *Vision research*, 41(12), (pp. 1561-1574).
- Fischer, G. (2001). User modeling in human--computer interaction. *User modeling and user-adapted interaction*, 11, (pp. 65-86).
- Glenn III, F. A., Iavecchia, H. P., Ross, L. V., Stokes, J. M., Weiland, W. J., Weiss, D., & Zaklad, A. L. (1986). Eye-voice-controlled interface. *Proceedings of the Human Factors Society Annual Meeting. Vol. 30, No. 4*, pp. 322-326. Sage CA: Los Angeles, CA: SAGE Publications.
- Grauman, K., Betke, M., Lombardi, J., Gips, J., & Bradski, G. R. (2003). Communication via eye blinks and eyebrow raises: Video-based human-computer interfaces. *Universal Access in the Information Society*, 2(4), (pp. 359-373).
- Gronwall, D. M. (1977). Paced auditory serial-addition task: a measure of recovery from concussion. *Perceptual and motor skills*, 44(2), (pp. 367-373).
- Hansen, J. P., Johansen, A. S., Hansen, D. W., Itoh, K., & Mashino, S. (2003). Command without a click: Dwell time typing by mouse and gaze selections. *Proceedings of Human-Computer Interaction--INTERACT*, (pp. 121-128).
- Hansen, J. P., Lund, H., Aoki, H., & Itoh, K. (2006). Gaze communication systems for people with ALS. *ALS Workshop, in conjunction with the 17th International Symposium on ALS/MND*.
- Harrison, A. W., & Rainer Jr, R. K. (1992). The influence of individual differences on skill in end-user computing. *Journal of Management Information Systems*, 9(1), (pp. 93-111).

## Bibliographie

- Helmert, J. R., Pannasch, S., & Velichkovsky, B. M. (2008). Influences of dwell time and cursor control on the performance in gaze driven typing. *Journal of Eye Movement Research*, 2(4), 1-8.
- Hillaire, S., Lécuyer, A., Cozot, R., & Casiez, G. (2008). Using an eye-tracking system to improve camera motions and depth-of-field blur effects in virtual environments. *Virtual Reality Conference, 2008. VR'08. IEEE*, (pp. 47-50).
- Hillaire, S., Lécuyer, A., Regia-Corte, T., Cozot, R., Royan, J., & Breton, G. (2010). A real-time visual attention model for predicting gaze point during first-person exploration of virtual environments. *Proceedings of the 17th ACM Symposium on Virtual Reality Software and Technology*, (pp. 191-198).
- Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures*. OUP Oxford.
- Hyrskykari, A. (2006). *Eyes in attentive interfaces: Experiences from creating iDict, a gaze-aware reading aid*. Tampereen yliopisto.
- Istance, H., Hyrskykari, A., Immonen, L., Mansikkamaa, S., & Vickers, S. (2010). Designing gaze gestures for gaming: an investigation of performance. *Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications*, (pp. 323-330).
- Jacob, R. J. (1993). Eye movement-based human-computer interaction techniques: Toward non-command interfaces. *Advances in human-computer interaction*, 4, (pp. 151-190).
- Jang, S., & Woo, W. (2003). Ubi-UCAM: a unified context-aware application model. *International and Interdisciplinary Conference on Modeling and Using Context*, (pp. 178-189).
- Jonides, J. (1981). Voluntary vs. automatic control over the mind's eye's movement. En: Long JB, Badeley AD (Eds.). *Attention and performance*. Hillsdale, NJ: Erlbaum.
- Jonides, J., & Yantis, S. (1988). Uniqueness of abrupt visual onset in capturing attention. *Perception & Psychophysics*, 43(4), (pp. 346-354).
- Just, M. A., & Carpenter, P. A. (1976). Eye fixations and cognitive processes. *Cognitive psychology*, 8(4), (pp. 441-480).
- Kahneman, D., Beatty, J., & Pollack, I. (1967). Perceptual deficit during a mental task. *Science*, 157(3785), (pp. 218-219).
- Kammerer, Y., Scheiter, K., & Beinhauer, W. (2008). Looking my way through the menu: the impact of menu design and multimodal input on gaze-based menu selection. *Proceedings of the 2008 symposium on Eye tracking research & applications*, (pp. 213-220).
- Konstantopoulos, P., Chapman, P., & Crundall, D. (2010). Driver's visual attention as a function of driving experience and visibility. Using a driving simulator to explore drivers' eye movements in day, night and rain driving. *Accident Analysis & Prevention*, 42(3), (pp. 827-834).
- Kourkolis, M. (1986). APP-6 Military symbols for land-based systems. *Military Agency for Standardization (MAS), NATO Letter of Promulgation*.

- Kubicki, S., Lepreux, S., & Kolski, C. (2013). Distributed UI on Interactive tabletops: issues and context model. Dans *Distributed User Interfaces: Usability and Collaboration* (pp. 27-38). Springer.
- Kurtenbach, G. P. (1993). *The design and evaluation of marking menus*. Ph.D. dissertation, University of Toronto, University.
- Kurtenbach, G., & Buxton, W. (1994). User learning and performance with marking menus. *Proceedings of the SIGCHI conference on Human factors in computing systems*, (pp. 258-264).
- Land, M. F. (2009). Vision, eye movements, and natural behavior. *Visual neuroscience*, 26(1), (pp. 51-62).
- Lanthier, S., Risko, E., Smilek, D., & Kingstone, A. (2013). Measuring the separate effects of practice and fatigue on eye movements during visual search. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 35(35).
- Large, E. W., & Jones, M. R. (1999). The dynamics of attending: How people track time-varying events. *Psychological review*, 106(1), 119-159.
- Leigh, R. J., & Zee, D. S. (2015). *The neurology of eye movements* (Vol. 90). Oxford: Oxford University Press, USA.
- Liebling, D. J., & Dumais, S. T. (2014). Gaze and mouse coordination in everyday work. *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication*, (pp. 1141-1150).
- Mackworth, N. H. (1965). Visual noise causes tunnel vision. *Psychonomic science*, 3(1-12), (pp. 67-68).
- Maguire, M. (2001). Context of use within usability activities. *International Journal of Human-Computer Studies*, 55(4), (pp. 453-483).
- Majaranta, P., & R  ih  , K.-J. (2002). Twenty years of eye typing: systems and design issues. Dans ACM (  d.), *Proceedings of the 2002 symposium on Eye tracking research & applications* (pp. 15-22). ACM.
- Majaranta, P., Ahola, U.-K., & Spakov, O. (2009). Fast gaze typing with an adjustable dwell time. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, (pp. 357-360).
- Martens, M. H., & Van Winsum, W. (2000). Measuring distraction: the peripheral detection task. *TNO Human Factors*.
- Mazza, V., Turatto, M., & Umilt  , C. (2005). Foreground--background segmentation and attention: A change blindness study. *Psychological research*, 69(3), (pp. 201-210).
- Miniotas, D., Spakov, O., & MacKenzie, I. S. (2004). Eye gaze interaction with expanding targets. Dans ACM (  d.), *CHI'04 extended abstracts on Human factors in computing systems* (pp. 1255-1258). ACM.

## Bibliographie

- Mollenbach, E. (2010). *Selection strategies in gaze interaction*. Ph.D. dissertation, Loughborough University.
- Morimoto, C. H., & Mimica, M. R. (2005). Eye gaze tracking techniques for interactive applications. *Computer vision and image understanding*, *98*(1), (pp. 4-24).
- Morimoto, C. H., Koons, D., Amit, A., Flickner, M., & Zhai, S. (1999). Keeping an eye for HCI. *Computer Graphics and Image Processing, 1999. Proceedings. XII Brazilian Symposium on*, (pp. 171-176).
- Norman, D. A. (1986). *Cognitive engineering*. Hillsdale, NJ: Lawrence Erlbaum.
- Norman, J. (2002). Two visual systems and two theories of perception: An attempt to reconcile the constructivist and ecological approaches. *Behavioral and brain sciences*, *25*(1), (pp. 73-96).
- Oshinger, W., Maeder, A., & Bergmann, N. (1998). A technique for image quality assessment based on a human visual system model. *Signal Processing Conference (EUSIPCO 1998), 9th European* (pp. 1-4). IEEE.
- Paas, F., Tuovinen, J. E., Tabbers, H., & Van Gerven, P. W. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational psychologist*, *38*(1), (pp. 63-71).
- Pelz, J. B., & Canosa, R. (2001). Oculomotor behavior and perceptual strategies in complex tasks. *Vision research*, *41*(25), (pp. 3587-3596).
- Pfeuffer, K., Alexander, J., Chong, M. K., & Gellersen, H. (2014). Gaze-touch: combining gaze with multi-touch for interaction on the same surface. *Proceedings of the 27th annual ACM symposium on User interface software and technology* (pp. 509-518). ACM.
- Posner, M. I. (1980). Orienting of attention. *Quarterly journal of experimental psychology*, *32*(1), (pp. 3-25).
- Posner, M. I., Snyder, C. R., & Davidson, B. J. (1980). Attention and the detection of signals. *Journal of experimental psychology: General*, *109*(2), (pp. 160-174).
- Reingold, E. M., & Sheridan, H. (2011). *Eye movements and visual expertise in chess and medicine*. Oxford University Press.
- Rensink, R. A., Kevin O'Regan, J., & Clark, J. J. (2000). On the failure to detect changes in scenes across brief interruptions. *Visual cognition*, *7*(1-3), (pp. 127-145).
- Rey, G. (2006). Méthode pour la modélisation du contexte d'interaction. *Ingénierie des Systèmes d'Information*, *11*(5), (pp. 141-166).
- Rich, E. (1979). User modeling via stereotypes. *Cognitive science*, *3*(4), (pp. 329-354).
- Riding, R. J., & Sadler-Smith, E. (1997). Cognitive style and learning strategies: Some implications for training design. *International Journal of Training and Development*, *1*(3), (pp. 199-208).
- Robinski, M., & Stein, M. (2013). Tracking visual scanning techniques in training simulation for helicopter landing. *Journal of Eye Movement Research*, *6*(2), 1-17.

- Rodieck, R. W., & Rodieck, R. W. (1998). *The first steps in seeing* (Vol. 1). Sinauer Associates Sunderland, MA.
- Rousseau, C., Bellik, Y., Vernier, F., & Bazalgette, D. (2006). A framework for the intelligent multimodal presentation of information. *Signal Processing*, *86*(12), (pp. 3696-3713).
- Russell, M. (2005). Using eye-tracking data to understand first impressions of a website. *Usability News*, *7*(1), (pp. 1-14).
- Salvucci, D. D. (1999). *Mapping eye movements to cognitive processes*. Pittsburgh, PA: Carnegie Mellon University.
- Salvucci, D. D., & Anderson, J. R. (2000). Intelligent gaze-added interfaces. *Proceedings of the SIGCHI conference on Human Factors in Computing Systems* (pp. 273-280). ACM.
- Samp, K., & Decker, S. (2010). Supporting menu design with radial layouts. *Proceedings of the International Conference on Advanced Visual Interfaces* (pp. 155-162). ACM.
- Schmidt, J. C., & Zelinsky, G. (2006). How is eye gaze affected by cognitive load and visual complexity? *Journal of Vision*, *6*(6), (pp. 363-363).
- Skovsgaard, H. (2012). *Noise Challenges in Monomodal Gaze Interaction*. Ph.D. dissertation, IT University of Copenhagen.
- Spakov, O., & Miniotos, D. (2004). On-line adjustment of dwell time for target selection by gaze. *Proceedings of the third Nordic conference on Human-computer interaction* (pp. 203-206). ACM.
- Spakov, O., & Miniotos, D. (2005). Gaze-based selection of standard-size menu items. *Proceedings of the 7th international conference on Multimodal interfaces* (pp. 124-128). ACM.
- Stellmach, S., & Dachselt, R. (2012). Look & touch: gaze-supported target acquisition. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 2981-2990). Austin: ACM.
- Tall, M. (2008). Neovisus: Gaze driven interface components. *Proceedings of the 4rd Conference on Communication by Gaze Interaction (COGAIN 2008)*, (pp. 47-51).
- Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *nature*, *381*(6582), 520-522.
- Treisman, A., & Schmidt, H. (1982). Illusory conjunctions in the perception of objects. *Cognitive psychology*, *14*(1), (pp. 107-141).
- Tsotsos, J. K., Culhane, S. M., Wai, W. Y., Lai, Y., Davis, N., & Nuflo, F. (1995). Modeling visual attention via selective tuning. *Artificial intelligence*, *78*(1-2), (pp. 507-545).
- Urbina, M. H., Lorenz, M., & Huckauf, A. (2010). Pies with EYEs: the limits of hierarchical pie menus in gaze control. *Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications* (pp. 93-96). ACM.

## Bibliographie

- Vernier, F., & Nigay, L. (2000). A framework for the combination and characterization of output modalities. *International Workshop on Design, Specification, and Verification of Interactive Systems* (pp. 35-50). Springer, Berlin, Heidelberg.
- Vidal, M., Pfeuffer, K., Bulling, A., & Gellersen, H. W. (2013). Pursuits: eye-based interaction with moving targets. Dans ACM (Éd.), *CHI'13 Extended Abstracts on Human Factors in Computing Systems* (pp. 3147-3150). ACM.
- Ware, C., & Mikaelian, H. H. (1987). An evaluation of an eye tracker as a device for computer input2. (ACM, Éd.) *ACM SIGCHI Bulletin*, 17, No. SI, (pp. 183-188).
- Wedel, M., & Pieters, R. (2008). A review of eye-tracking research in marketing. Dans *Review of marketing research* (pp. 123-147). Emerald Group Publishing Limited.
- Wickens, C. D., Hollands, J. G., Banbury, S., & Parasuraman, R. (2015). *Engineering psychology & human performance*. Psychology Press.
- Williams, L. J. (1982). Cognitive load and the functional field of view. *Human Factors*, 24(6), (pp. 683-692).
- Wobbrock, J. O., Rubinstein, J., Sawyer, M. W., & Duchowski, A. T. (2008). Longitudinal evaluation of discrete consecutive gaze gestures for text entry. *Proceedings of the 2008 symposium on Eye tracking research & applications*, (pp. 11-18).
- Zander, T. O., Gaertner, M., Kothe, C., & Vilimek, R. (2010). Combining eye gaze input with a brain-computer interface for touchless human-computer interaction. *Intl. Journal of Human-Computer Interaction*, 27(1), (pp. 38-51).
- Zhai, S., Morimoto, C., & Ihde, S. (1999). Manual and gaze input cascaded (MAGIC) pointing. *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, (pp. 246-253).
- Zimmermann, A., Lorenz, A., & Oppermann, R. (2007). An operational definition of context. *International and Interdisciplinary Conference on Modeling and Using Context* (pp. 558-571). Springer, Berlin, Heidelberg.

