



HAL
open science

Améliorer l'expressivité des interfaces d'entrée par des interactions tactilo-kinesthésiques à multiples composantes

Gauthier Faisandaz

► **To cite this version:**

Gauthier Faisandaz. Améliorer l'expressivité des interfaces d'entrée par des interactions tactilo-kinesthésiques à multiples composantes. Interface homme-machine [cs.HC]. Université Grenoble Alpes, 2024. Français. NNT: . tel-04948639

HAL Id: tel-04948639

<https://hal.science/tel-04948639v1>

Submitted on 14 Feb 2025

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.

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ GRENOBLE ALPES

École doctorale : MSTII - Mathématiques, Sciences et technologies de l'information, Informatique

Spécialité : Informatique

Unité de recherche : Laboratoire d'Informatique de Grenoble

Améliorer l'expressivité des interfaces d'entrée par des interactions tactilo-kinesthésiques à multiples composantes

Increasing input interaction expressiveness using eyes-free multi-finger interaction

Présentée par :

Gauthier FAISANDAZ

Direction de thèse :

Laurence NIGAY

PROFESSEURE DES UNIVERSITES, Université Grenoble Alpes

Directrice de thèse

Christophe JOUFFRAIS

DR CNRS, Université Toulouse III - Paul Sabatier

Co-directeur de thèse

Alix GOGUEY

Université Grenoble-Alpes

Co-encadrant de thèse

Rapporteurs :

Isabelle PECCI

MAITRESSE DE CONFERENCES, Université de Lorraine

Céline LEMERCIER

PROFESSEURE DES UNIVERSITES, Université Toulouse

Thèse soutenue publiquement le **12 décembre 2023**, devant le jury composé de :

Isabelle PECCI

MAITRESSE DE CONFERENCES, Université de Lorraine

Rapporteuse

Céline LEMERCIER

PROFESSEURE DES UNIVERSITES, Université Toulouse

Rapporteuse

Sophie DUPUY-CHESSA

PROFESSEURE DES UNIVERSITES, Université Grenoble-Alpes

Examinatrice

Gilles BAILLY

DIRECTEUR DE RECHERCHE, CNRS Délégation Ile-de-France

Examineur

Sabrina PANEELS

DOCTEUR EN SCIENCES, CEA-List

Examinatrice



Table des matières

Table des figures	iv
Introduction	1
1 Interaction Humain-Machine en entrée pour les Personnes ayant une Déficience	
Visuelle	9
1.1 Introduction du chapitre	10
1.2 Terminologie adoptée dans nos travaux	11
1.2.1 Modalité d'interaction	11
1.2.2 Mode de communication	13
1.2.3 Paradigme d'interaction	14
1.2.4 WIMP et post-WIMP	15
1.2.5 Objets d'intérêt	15
1.3 Justification de la grille d'analyse	15
1.3.1 Modèles conceptuels des dispositifs	16
1.3.2 Modèles conceptuels de l'interaction	20
1.4 Synthèse	23
1.5 Grille d'analyse des paradigmes et modalités d'interaction	24
1.6 Analyse des paradigmes et modalités d'interaction pour les Personnes ayant une Déficience Visuelle	29
1.6.1 Le paradigme Tactile	29
1.6.2 Le paradigme Vocal	50

TABLE DES MATIÈRES

1.6.3	Le paradigme Tangible	58
1.6.4	Le paradigme Gestuel	67
1.6.5	Synthèse	79
1.7	Les combinaisons de paradigmes	80
1.7.1	Tactile + Vocal	80
1.7.2	Tactile + Tangible	83
1.7.3	Tactile + Gestuel	87
1.8	Le paradigme Micro-Gestuel	91
1.8.1	Tactile + Micro-gestuel	101
1.9	Synthèse	103
1.10	Conclusion	105
2	Faisabilité des microgestes Thumb-To-Finger en complément du touché sur un écran tactile	107
2.1	Question de recherche	108
2.2	Étude expérimentale	111
2.2.1	Protocole expérimental	111
2.2.2	Implémentation	114
2.2.3	Résultats	116
2.2.4	Analyse	121
2.2.5	Optimisation des μ Gestes TTF : TAP et DOUBLE TAP	123
2.2.6	Optimisation des μ Gestes TTF : SWIPE	126
2.3	Synthèse	128
2.4	Conclusion	130
3	Utilité et Utilisabilité des microgestes Thumb-To-Finger en complément du touché sur un écran tactile	131
3.1	Question de recherche	132
3.1.1	Dispositif de captation	137
3.2	Études expérimentales	138
3.2.1	Étude 1 : exploration de carte	138

TABLE DES MATIÈRES

3.2.2 Étude 2 : couper / coller	143
3.2.3 Étude 3 : sélection de texte	149
3.3 Conclusion du chapitre	172
Conclusion	173
Bibliographie	182
A Annexe I	219
A.1 La déficience visuelle	219
A.2 Les technologies d'assistance	222
A.3 Accessibilité des IHM	225
A Annexe II	227

Table des figures

1	Exemples de Technologies d'Assistance (TA).	3
2	Exemple de microgeste (μ Geste) Thumb-To-Finger (TTF).	5
1.1	Schéma d'un système numérique interactif.	13
1.2	Tableau de la taxonomie de Buxton.	17
1.3	Illustration de l'espace de conception de Card.	18
1.4	Illustration du modèle de l'interaction instrumentale de Beaudouin-Lafon	22
1.5	Tableau synthétisant la grille d'analyse.	27
1.6	Schéma représentant le dispositif "écran tactile" dans l'espace de conception des dispositifs en entrée de Card.	29
1.7	Schéma illustrant la transformation d'une interface graphique (2D) en une liste d'éléments (1D).	31
1.8	Illustration du clavier Braille sur tablette.	32
1.9	Illustration d'un clavier physique 3*4 de téléphone mobile	33
1.10	Exemple de trajectoire d'interaction sur téléphone portable (Android) pour ajouter un périphérique bluetooth.	34
1.11	Capture d'écran du rotor de VoiceOver, sous iOS.	38
1.12	Captures d'écrans des dispositifs de Krajnc et al. et de Pooripanyakun et al.	42
1.13	Capture d'écran du dispositif AudioDraw de Grussenmeyer et al.	44
1.14	Photographie du dispositif No-Look de Bonner et al.	46
1.15	Schéma du fonctionnement du dispositif VectorEntry de Samanta et al.	47
1.16	Photographies du dispositif QB-Gest de Ye et al.	48

TABLE DES FIGURES

1.17 Tableau synthétisant l'analyse des modalités d'interaction issues du paradigme Tactile.	49
1.18 Tableau listant les modalités d'interaction issues du paradigme Tactile.	49
1.19 Schéma représentant le dispositif "microphone" dans l'espace de conception des dispositifs en entrée de Card.	51
1.20 Tableau synthétisant l'analyse des modalités d'interaction issues du paradigme Vocal.	57
1.21 Tableau listant les modalités d'interaction issues du paradigme Vocal.	57
1.22 Schéma représentant différents dispositifs relevant du paradigme Tangible dans l'espace de conception des dispositifs en entrée de Card.	58
1.23 Photographies du dispositif Tangible Desktop de Baldwin et al.	60
1.24 Photographie et schéma illustrant le dispositif DualPanto de Schneider et al.	60
1.25 Matrice braille rafraîchissable et tactile : l'HyperBraille F de Metec.	61
1.26 Photographies des dispositifs MapSense de Brulé et al. et Tac-Trace de Jafri et al.	63
1.27 Schéma illustrant le dispositif LETSmath de Marichal et al.	64
1.28 Photographies du dispositif BrailleBuddy de Lang et al.	65
1.29 Tableau synthétisant l'analyse des modalités d'interaction issues du paradigme Tangible.	67
1.30 Tableau listant les modalités d'interaction issues du paradigme Tangible.	67
1.31 Schéma représentant différents dispositifs relevant du paradigme Gestuel dans l'espace de conception des dispositifs en entrée de Card.	68
1.32 Photographies de VirtualShelves de Li et al. et du dispositif de Costa et al.	69
1.33 Illustrations du dispositif rapMenu de Ni et al.	71
1.34 Illustrations du dispositif de Kajastila et al.	72
1.35 Illustrations du dispositif WatchIt de Perrault et al.	75
1.36 Illustrations du dispositif EarBuddy de Xu et al.	75
1.37 Illustrations du fonctionnement du dispositif ThroughHand de Jung et al.	77
1.38 Tableau synthétisant l'analyse des modalités d'interaction issues du paradigme Gestuel.	78
1.39 Tableau listant les modalités d'interaction issues du paradigme Gestuel.	79

TABLE DES FIGURES

1.40 Schéma représentant les dispositifs "écran tactile" et "microphone" dans l'espace de conception des dispositifs en entrée de Card.	81
1.41 Tableau listant les modalités d'interaction combinant les paradigmes Tactile et Vocal.	83
1.42 Schéma représentant les dispositifs "écran tactile" et tangibles dans l'espace de conception des dispositifs en entrée de Card.	84
1.43 Photographies de divers dispositifs Tactile + Tangible, de Kane et al., El-Glaly et al., Lee et al.	84
1.44 Photographies du dispositif Interactiles de Zhang et al.	85
1.45 Photographies du dispositif ClipWidgets de Visschedijk et al.	86
1.46 Tableau listant les modalités d'interaction combinant les paradigmes Tactile et Tangible.	87
1.47 Schéma représentant les dispositifs "écran tactile" et gestuel dans l'espace de conception des dispositifs en entrée de Card.	88
1.48 Schéma du dispositif TouchOver de Scoditti et al.	89
1.49 Tableau listant les modalités d'interaction combinant les paradigmes Tactile et Gestuel.	91
1.50 Photographie illustrant un TAP du pouce sur l'index.	92
1.51 Schéma représentant les dispositifs Micro-Gestuels dans l'espace de conception des dispositifs en entrée de Card.	93
1.52 Schéma représentant le dispositif DigitTap de Prätorius et al.	95
1.53 Schéma représentant les zones à atteindre dans l'expérience de Huang et al.	96
1.54 Schéma représentant le dispositif Tip-Tap de Katsuragawa et al.	96
1.55 Schéma représentant le dispositif FingerT9 de Wong et al.	97
1.56 Schéma représentant le dispositif TipText de Xu et al.	98
1.57 Tableau synthétisant l'analyse des modalités d'interaction issues du paradigme Micro-Gestuel.	100
1.58 Tableau listant les modalités d'interaction issues du paradigme Micro-Gestuel.	100
1.59 Schéma illustrant l'utilisation des μ Gestes TTF pour augmenter l'expressivité.	101
1.60 Schéma illustrant le dispositif de Tsai et al.	102

TABLE DES FIGURES

2.1 Photographies des dispositifs EyeRing de Nanayakkara et al. et AiSee de Boldu et al.	110
2.2 Un TAP effectué sur l'index lorsque celui-ci est en contact avec l'écran tactile. . .	110
2.3 Les zones sur lesquels les gestes sont effectués.	113
2.4 Exemples d'instruction affichés durant l'expérience.	115
2.5 Photographique du premier prototype.	116
2.6 Graphique illustrant le taux d'erreur pour la phase d'entraînement.	118
2.7 Graphique illustrant le taux d'erreur.	119
2.8 Graphique illustrant la note de confort et les intervalles de confiances.	120
2.9 Graphique illustrant la dérive.	121
2.10 Tableau récapitulatif pour les TAP et 2-TAP uniquement	123
2.11 Graphique illustrant le taux d'erreur avec intervalle de confiance pour les TAP et 2-TAP	124
2.12 Tableau récapitulatif pour les SWIPE uniquement.	126
2.13 Graphique illustrant le taux d'erreur avec intervalle de confiance pour les SWIPE.	127
2.14 Ensemble de μ Gestes TTF pouvant être effectués tout en touchant un écran tactile.	129
3.1 Photographie d'un utilisateur explorant une carte interactive audio-tactile.	133
3.2 Illustration schématisée d'un dispositif audio-tactile.	134
3.3 Prototype de gant de captation modulaire. Exemple de manchon pour l'index.	138
3.4 Exemple d'utilisation de la carte audio-tactile	139
3.5 Exemple d'une interaction complète.	144
3.6 Capture d'écran du rotor de VoiceOver, sous iOS.	150
3.7 Schéma du sous-menu accessible depuis la position #8 (Selection de texte) du rotor.	151
3.8 Scénario de sélection de texte.	152
3.9 Modélisation StEM des trajectoires d'interaction.	154
3.10 Capture d'écran de l'interface à destination de l'expérimentateur.	158
3.11 Photographie illustrant l'utilisation de la tablette avec la technique VoiceOver.	160
3.12 Photographie des manchons utilisés dans le troisième prototype.	160

TABLE DES FIGURES

3.13	Tableau synthétisant les résultats de l'expérience de sélection de texte.	162
3.14	Graphique montrant le temps de sélection et de pré-sélection.	163
3.15	Tableau synthétisant les résultats des questionnaire Raw Nasa-TLX et UMUX-Lite.164	
3.16	Tableau synthétisant les résultats de l'analyse des trajectoires.	167
A.1	Gauche : Echelle de vision. Droite : Tableau de l'OMS résumant la mesure de la déficience visuelle. Illustration issue de [165].	220
A.2	Gauche : Loupe électronique (aussi appelée "loupe vidéo"). Droite : Clavier Perkins permettant d'écrire et de lire en Braille sur ordinateur.	223
A.3	Exemple de mauvaise accessibilité pour PDV. Capture d'écran issue du site internet de l'université d'Art de Yale.	226

Introduction

Dans ce chapitre, nous présentons le public cible et la portée de nos travaux. Nous introduisons par ailleurs quelques définitions nécessaires à la lecture de ce manuscrit : la déficience visuelle, les technologies d'assistance et leurs usages par les personnes ayant une déficience visuelle ainsi que les défis rencontrés.

Contexte

Dans nos sociétés largement fondées sur la capacité de voir, les déficiences visuelles ont un impact significatif sur la vie des individus et peuvent constituer un handicap majeur à la réalisation des activités du quotidien. Pour mieux en appréhender les enjeux, il est important de définir les concepts mis en œuvre dans ce texte que sont la déficience visuelle, le handicap, les aides technologiques d'assistance et l'accessibilité des interfaces de systèmes interactifs. Nous présentons ici une vision synthétique de ces concepts pour poser le contexte de nos travaux de recherche. Nous invitons le lecteur à se référer à notre présentation plus exhaustive de ces concepts, en annexe A.

La déficience visuelle

Dans le cadre de cette thèse, nous utilisons le terme "déficience visuelle" tel que défini par la 11ème édition de la Classification internationale des maladies (CIM 11) [163] pour désigner des niveaux de déficience modérés, sévères, ainsi que la cécité. Nos travaux portant sur les systèmes interactifs, nous utilisons tout au long de ce document le terme "personne ayant une

déficience visuelle" (abrégé PDV) pour parler de personnes dont la déficience ne permet pas d'effectuer des tâches de nature numérique nécessitant la vision (c'est-à-dire interagir avec un écran) sans une aide supplémentaire notable (par exemple, un agrandissement du texte ou une vocalisation).

Le handicap

Dans le langage courant, le terme "handicap" est souvent utilisé comme synonyme de "déficience" [162]. Il faut cependant distinguer la déficience du handicap. Le handicap fait référence aux limitations et restrictions qu'une personne peut rencontrer dans ses interactions avec son environnement - qu'il soit physique, social ou attitudinal. Notamment, les chercheurs Sears et Young [200] ont introduit en Interaction Humain-Machine le concept de "déficiences et incapacités induites par la situation" (SIID, pour Situationally-Induced Impairments and Disabilities) en 2003. Cela implique que le niveau de handicap des PDV peut être réduit en optimisant leur fonctionnement quotidien et fournissant des ajustements pratiques pour compenser les conséquences de leur déficience [218, 191, 97]. A l'aide de dispositifs technologies par exemple, il devient ainsi possible d'améliorer significativement leur qualité de vie [210, 52, 121, 166].

Les technologies d'assistance

Tout équipement ou système, utilisé pour accroître, maintenir ou améliorer les capacités fonctionnelles des personnes en situation de handicap peut être considéré comme une Technologie d'Assistance (TA) [1]. Ces TA renforcent l'indépendance dans l'exécution d'activités de la vie quotidienne des personnes en situation de handicap. Il existe une large gamme de TA à destination des PDV : des loupes électroniques permettant de modifier la luminosité, les couleurs, le contraste . . . mais aussi des TA qui substituent la perception tactile ou auditive à la perception visuelle pour permettre aux PDV de consulter du contenu, sous forme de texte Braille par exemple (figure 1).

Néanmoins les taux d'abandon et de non-utilisation des TA traditionnels sont préoccupants. Selon certains rapports, le pourcentage d'abandon varie de 30% à 50% pour l'ensemble des



FIGURE 1 – **Gauche** : Loupe électronique (aussi appelée “loupe vidéo”). **Droite** : Clavier Perkins permettant d’écrire et de lire en Braille sur ordinateur.

dispositifs, et peut atteindre jusqu’à 75% [77]. En effet, ces dispositifs ne sont pas utilisables par tous en raison de divers facteurs tels que leur disponibilité, leur coût, leur acceptabilité, voire leur facilité d’usage [2, 172, 11, 165].

C’est pourquoi les appareils grand public tels que les téléphones portables et les tablettes, dont l’usage croît chaque année [3], sont de plus en plus utilisés par les PDV [140], car ils sont généralement plus abordables et attirent moins l’attention sur l’utilisateur [97]. Dotés de nombreuses possibilités (notamment : affichage, caméra, audio, GPS, internet), les téléphones portables modernes peuvent exécuter de nombreuses applications d’assistance pour les PDV, comparables à d’autres dispositifs traditionnels plus spécialisés [56, 167]. Cela permet aux PDV de n’utiliser que leur téléphone portable plutôt que de devoir acheter et gérer une kyrielle de TA différentes.

Accessibilité des interfaces

: L’utilisation des appareils grand public évoqués précédemment peut être une cause de handicap pour les PDV lorsque les fonctionnalités proposées ne sont pas suffisamment adaptées [113]. En particulier, les PDV sont confrontés à des défis inhérents au fonctionnement de ces appareils grand public, ceux-ci utilisant principalement des écrans tactiles [147], dont l’usage

repose essentiellement sur des interfaces visuelles [194].

Il s'agit là toutefois de problèmes liés au retour d'information en sortie, celui produit par le système à destination de l'utilisateur. La déficience visuelle handicape aussi l'usage des téléphones portables en entrée. Les entrées font référence aux moyens par lesquels les utilisateurs interagissent avec un système. Les appareils grand public sont dotés d'écrans tactiles qui requièrent principalement de manipuler directement [205, 207] des éléments appartenant à l'interface graphique (appelée GUI, pour Graphical User Interface). Cela nécessite de viser précisément les éléments interactifs qui la composent (par exemple, liens, boutons, icônes) ce qui présente un défi majeur pour les PDV - particulièrement dans le cas des téléphones portables et tablettes, dotés d'écrans tactiles de taille modeste. Par ailleurs, la taille et l'aspect mobile de ces appareils restreint l'usage de dispositifs physiques supplémentaires plus accessibles aux PDV, tel qu'un clavier par exemple. En effet, les PDV apprécient le téléphone portable pour sa multifonctionnalité et sa portabilité, or ajouter des dispositifs supplémentaires serait préjudiciable à la praticité perçue du téléphone portable (ou de la tablette).

Portée des travaux et Question de recherche

Dans ce contexte, cette thèse met au centre de ses préoccupations l'accessibilité des appareils tactiles grand public auprès des personnes ayant une déficience visuelle (PDV). Les travaux s'inscrivent donc dans le domaine de l'Interaction Humain-Machine (IHM), un domaine de l'informatique qui se concentre sur la conception et l'interaction des utilisateurs avec les systèmes informatiques. Elle vise à rendre les interfaces utilisateur conviviales, efficaces et intuitives, facilitant ainsi la communication entre les êtres humains et les ordinateurs.

Pour cerner les travaux, nous nous concentrons sur les techniques d'interaction en entrée et nous étudions l'utilisation complémentaire aux écrans tactiles d'une modalité d'interaction gestuelle particulière que sont les **microgestes** (μ Gestes, ou tout simplement, μ G).

Afin de mieux appréhender cette introduction à nos travaux, nous définissons brièvement

ici ce que sont ces microgestes, néanmoins, une définition plus complète sera donnée dans le chapitre suivant de ce document.

Il existe de nombreux μ Gestes différents dans la littérature scientifique (utilisant, par exemple, les pieds [10, 55], la tête [59], les yeux [98, 215], voire la mâchoire [13, 79, 51]). Cependant dans nos travaux nous traitons uniquement des μ Gestes dits "TTF". Les μ Gestes TTF (pour Thumb-to-Finger) sont des mouvements effectués avec le pouce sur les doigts de la même main, tels que des "Taps" (similaires à des "clics" effectués avec une souris) ou des "Swipes" (des "glissements") (par exemple, figure 2).

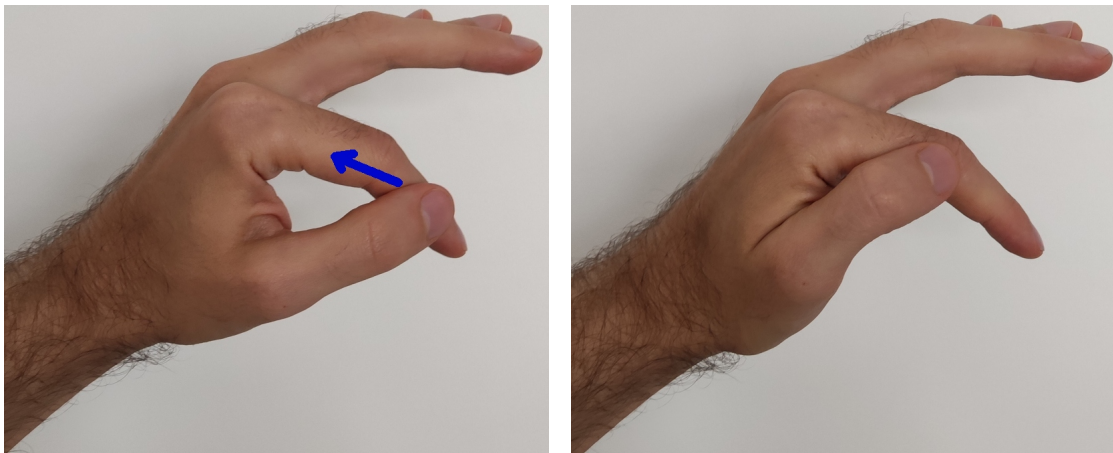


FIGURE 2 – Un exemple de microgeste Thumb-To-Finger : un Swipe Haut-Bas sur l'index. Le pouce initie le contact avec l'index sur la phalange la plus proche du pouce, puis glisse le long du doigt en direction de la paume avant de relâcher le contact.

Les μ Gestes TTF présentent une grande expressivité : ils offrent une variété de mouvements possibles qui peuvent être réalisés avec précision sur des parties distinctes des doigts en tirant parti de la dextérité de la main et de la proprioception. Les recherches actuelles sur les μ Gestes TTF montrent qu'ils sont peu exigeants sur le plan cognitif ([141, 203, 241]) et utilisables sans regarder [240] : deux caractéristiques particulièrement adaptées à l'interaction pour les PDV. Les μ Gestes TTF pourraient donc être un moyen d'améliorer l'accessibilité de divers dispositifs, en particulier des dispositifs à écran tactile, dont le principe d'interaction est intrinsèquement basé sur la vision.

Or, bien que les μ Gestes TTF aient été étudiés dans diverses situations ([49, 240, 31, 230]

pour n'en citer que quelques-unes), à notre connaissance, ils n'ont jamais été étudiés avec des PDV, malgré les opportunités qu'ils pourraient offrir à ce public. En effet, nous avons énoncé précédemment que les PDV utilisent - voire dépendent - de plus en plus des dispositifs à écran tactile pour leurs tâches quotidiennes [83, 5, 113], telles que la lecture, les interactions sociales, la navigation en extérieur ou la reconnaissance d'objets.

Aujourd'hui, les PDV accèdent à ces dispositifs à l'aide d'outils d'accessibilité, notamment des applications de lecteur d'écran telles que VoiceOver sur iOS et TalkBack sur Android¹. Ces applications posent toutefois deux problèmes majeurs :

- L'interaction avec les outils d'accessibilité repose sur l'utilisation d'un menu ou de raccourcis. Le menu nécessite généralement un geste simple pour être déclenché, puis plusieurs gestes pour sélectionner l'action ou l'élément requis, et enfin un geste final pour valider la sélection. Cette séquence complète de gestes doit souvent être répétée plusieurs fois, y compris pour des tâches courantes telles que copier/coller, ce qui augmente considérablement le temps, la complexité et la pénibilité de l'interaction.
- Les raccourcis consistent en des gestes effectués sur l'écran avec un ou plusieurs doigts. Ils peuvent être simples (par exemple, un tapotement avec un doigt pour sélectionner une cible) ou plus complexes (par exemple, dessiner un motif en forme de Z avec deux doigts pour fermer un message d'alerte). Les raccourcis sont plus directs que les menus. Néanmoins, en raison de la longue liste de gestes disponibles, les raccourcis sont longs à apprendre et à mémoriser [186, 131], sont parfois difficiles ou fatigants à réaliser pour les PDV [109], et sont mal reconnus par les dispositifs [186].

Nous souhaitons répondre à ces problèmes à l'aide des μ Gestes TTF. Nous posons donc la question de recherche suivante : *peut-on, et si oui comment, améliorer l'accessibilité et l'utilisabilité des dispositifs à écran tactile dans une situation d'usage sans retour visuel à l'aide d'une modalité d'interaction supplémentaire que sont les microgestes dit "Thumb-To-Finger" ?*

1. Talkback - Android : <https://support.google.com/accessibility/android/answer/6007100> et VoiceOver - iOS : <https://support.apple.com/en-gb/HT211899>

Contributions

Répondant à cette question de recherche, nous présentons dans ce document nos travaux dont l'objectif est d'étudier la faisabilité, l'utilité et l'utilisabilité des μ Gestes TTF couplés à la modalité tactile. Nos contributions sont les suivantes :

- Une exploration systématique de la faisabilité et la performance de 33 μ Gestes TTF parmi les plus courants dans la littérature, en situation de non-retour visuel, et ce dans un contexte d'usage d'un dispositif à écran tactile - c'est-à-dire lorsque l'extrémité de l'index touche une surface (tactile) ;
- Un vocabulaire d'entrée sous la forme d'un ensemble de 8 μ Gestes TTF utilisables dans les conditions évoquées dans le premier point ;
- Trois applications démontrant le potentiel, l'utilité et l'utilisabilité des μ Gestes TTF couplés à la modalité tactile ;
- Trois prototypes de gants électroniques permettant de capter les μ Gestes TTF.

Structure de la thèse

La structure de ce document est organisée comme suit :

Chapitre I : Interaction Humain-Machine en entrée pour les Personnes ayant une Déficience Visuelle

Dans ce chapitre, nous faisons un état des lieux des paradigmes d'interaction les plus communs employés par les PDV pour interagir avec un dispositif numérique. Nous examinons les solutions techniques existantes à l'aide d'une grille d'analyse que nous avons établie à l'aide de modèles conceptuels en IHM. Nous introduisons ensuite les μ Gestes TTF afin de montrer qu'ils sont une solution particulièrement adaptée à notre problématique.

Chapitre II : Faisabilité des microgestes Thumb-To-Finger en complément du touché sur un écran tactile

Au sein de ce chapitre, nous présentons nos travaux sur la faisabilité d'une interface multi-

modale utilisant conjointement l'écran tactile d'une tablette avec les μ Gestes TTF en l'absence de retour visuel. Ce chapitre reprend et présente des éléments issus d'un article publié et présenté dans la conférence ACM-ICMI 2022 : "Keep in Touch : Combining Touch Interaction with Thumb-to-Finger μ Gestures for People with Visual Impairment" [71].

Chapitre III : Utilité et Utilisabilité des microgestes Thumb-To-Finger en complément du touché sur un écran tactile Dans ce chapitre, nous présentons trois applications afin d'évaluer l'utilité et l'utilisabilité de ces μ Gestes TTF pour des PDV : une exploration d'une carte tactile, une tâche de copier/coller et une tâche de sélection de texte. Nous présentons aussi 2 implémentations de gants permettant de reconnaître les ensembles de μ Gestes TTF utilisés dans ces scénarios. Ce chapitre reprend et présente des éléments issus d'un article publié et présenté dans la conférence ACM-ICMI 2023 : " μ GeT : Multimodal eyes-free text selection technique combining touch interaction and microgestures" [72]

Conclusion Dans ce dernier chapitre, nous rappelons nos contributions et discutons des limitations et perspectives possibles de nos travaux à moyens et longs termes.

Chapitre 1

Interaction Humain-Machine en entrée pour les Personnes ayant une Déficience Visuelle

En introduction, nous avons présenté le public cible de nos travaux - les Personnes ayant une Déficience Visuelle (PDV) - et les définitions de concepts clefs relatives à ce contexte : la déficience visuelle, les technologies d'assistance TA) et leurs usages par les PDV ainsi que les défis rencontrés. Nous avons ensuite introduit la question de recherche que nous traitons dans nos travaux et présenté nos contributions.

Dans ce chapitre, nous examinons les solutions techniques existantes à destination des PDV. Pour les étudier, nous décrivons d'abord plusieurs modèles conceptuels du domaine de l'Interaction Humain-Machine (IHM) qui nous ont permis d'établir une grille d'analyse. Nous utilisons cette grille pour décrire ces solutions et motiver nos travaux sur les microgestes (μ Gestes) Thumb-To-Finger (TTF). En effet, les μ Gestes TTF ont été étudiés dans de nombreux contextes et possèdent des caractéristiques très intéressantes, notamment pour les PDV, pourtant il n'existe que très peu de littérature à ce sujet. Après avoir analysé les solutions existantes,

nous présentons les μ Gestes TTF puis leur utilisation conjointe à l'écran tactile afin d'illustrer leur potentiel pour un usage par des PDV.

1.1 Introduction du chapitre

Pour paraphraser David Osborne, auteur du livre Ergonomics at Work : *"La transmission de l'information avec un système est optimale lorsque les moyens d'interaction correspondent aux capacités et comportements de l'utilisateur"* [159].

Partant de ce postulat, notre objectif est de trouver un "moyen" d'interaction - terme générique que nous utilisons ici en guise d'introduction - correspondant aux "capacités" et aux "comportements" des PDV. Celui-ci doit donc respecter deux critères, il doit être utilisable :

- 1) ... en prenant en compte la déficience visuelle des utilisateurs (capacités) ;
- 2) ... sur téléphone portable / tablette, le support qu'ils préfèrent utiliser (comportement).

En effet, nous avons vu en introduction que les PDV utilisent majoritairement leur téléphone portable pour leurs activités quotidiennes plutôt que des dispositifs d'assistance plus spécialisés, généralement plus chers et plus stigmatisants. Mais bien que les téléphones portables soient préférés, les écrans tactiles dont ils sont dotés posent des défis inhérents à leur conception : ils sont avant tout fait pour afficher une multitude d'interfaces graphiques différentes et dont l'interaction s'effectue - notamment, mais pas uniquement - par le biais de la vue.

Il existe toutefois d'autres "moyens" d'interaction avec des systèmes numériques qui ne nécessitent pas la vue. L'objectif de ce chapitre est d'étudier ces différents "moyens" et de capitaliser sur le ou les plus appropriés pour compléter l'écran tactile des téléphones portables, dans le but de le rendre plus accessible aux PDV, notamment celles ayant une cécité totale. Nous partons du principe que les solutions utilisables sans la vue proposées dans cette thèse ou issues de la littérature sont aussi utiles et utilisables par les PDV pouvant utiliser une vision

dégradée.

Dans ce chapitre, nous dressons donc un état de l'art de la littérature traitant de ces "moyens" d'interaction. Nous définissons dans un premier temps la terminologie et les concepts abordés dans cette thèse. Puis, nous introduisons des modèles conceptuels issus du domaine de l'Interaction Humain-Machine (IHM) sur lesquels nous construisons une grille d'analyse. Cette grille nous permet ensuite d'étudier les solutions existantes pour les PDV : 1) les écrans tactiles mais adaptés pour fonctionner au mieux dans un contexte sans vision, ainsi que les "moyens" basés sur 2) le vocal, 3) le tangible et 4) le gestuel en l'air et sur le corps. Nous présentons ensuite les solutions basées sur une combinaison entre l'écran tactile et les autres "moyens" (vocal, tangible, gestuel). Enfin, nous concluons le panorama des solutions en soulignant les enjeux existant, pour introduire ensuite les microgestes (μ Gestes) Thumb-To-Finger (TTF) comme une piste de recherche prometteuse afin d'enrichir l'interaction tactile sur des téléphones portables et tablettes.

1.2 Terminologie adoptée dans nos travaux

Si les "moyens" d'interaction, terme utilisé dans l'introduction de ce chapitre, sont entourés de guillemets, c'est que ce terme n'a pas de fondement épistémologique. Il s'agit en réalité d'un terme générique que nous avons employé en guise d'initiation et englobant plusieurs concepts différents et connexes que nous définissons successivement : modalité d'interaction, mode de communication, paradigme d'interaction, WIMP, post-WIMP et objets d'intérêt.

1.2.1 Modalité d'interaction

La notion de modalité d'interaction - aussi appelée technique d'interaction - est bien connue des acteurs de la communauté IHM, mais elle est souvent appréhendée sous des perspectives différentes. Cette pluralité a pour effet de brouiller les contours de ce concept, comme le montre la définition qu'en fait Wikipedia ¹ :

1. Human-Computer Interaction - Wikipedia : [https://en.wikipedia.org/wiki/Modality_\(human%E2%80%93computer_interaction\)](https://en.wikipedia.org/wiki/Modality_(human%E2%80%93computer_interaction))

"In the context of human–computer interaction, a modality is the classification of a single independent channel of input/output between a computer and a human. Such channels may differ based on sensory nature (e.g., visual vs. auditory), or other significant differences in processing (e.g., text vs. image)."

Ce que cette définition nous dit, c'est qu'une modalité d'interaction est "une combinaison de canaux d'entrée et de sortie entre un système numérique et un être humain" et que ces canaux peuvent être de nature "sensorielle" ou "selon la nature de l'information traitée (par exemple, texte vs. image)".

Face à ce flou terminologique, nous choisissons d'employer la définition de Nigay et Coutaz [156], issue d'une revue de la littérature sur les notions de média et de modalité. Dans un système numérique interactif (figure 1.1), une **modalité d'interaction** représente une façon de communiquer entre l'utilisateur et le système par le biais d'une interface. Elle peut s'effectuer de l'utilisateur vers le système (on parle alors de modalité en entrée) ou du système vers l'utilisateur (on parle de modalité de sortie).

Selon ces auteurs, une modalité d'interaction est formée de deux composantes symbolisées ainsi : <le dispositif physique, le langage d'interaction>. Le **dispositif** physique fait référence à un dispositif d'entrée/sortie constituant l'interface du système comme une souris, un clavier ou un écran. Le langage d'interaction, quant à lui, fait référence au contenu informationnel et à la structure qui lui permet d'assurer sa fonction de communication. Par exemple, dans une interface en ligne de commande (CLI), une modalité d'entrée peut être décrite de la sorte <clavier, langage de commande>, ce qui signifie que le dispositif d'entrée est un clavier physique et que le langage d'interaction est un ensemble formalisé de commandes textuelles et de règles de syntaxe.

Un système peut utiliser une ou plusieurs modalités d'entrée et/ou de sortie. S'il n'en utilise qu'une seule, on le désigne comme unimodal : par exemple, un assistant vocal recevant

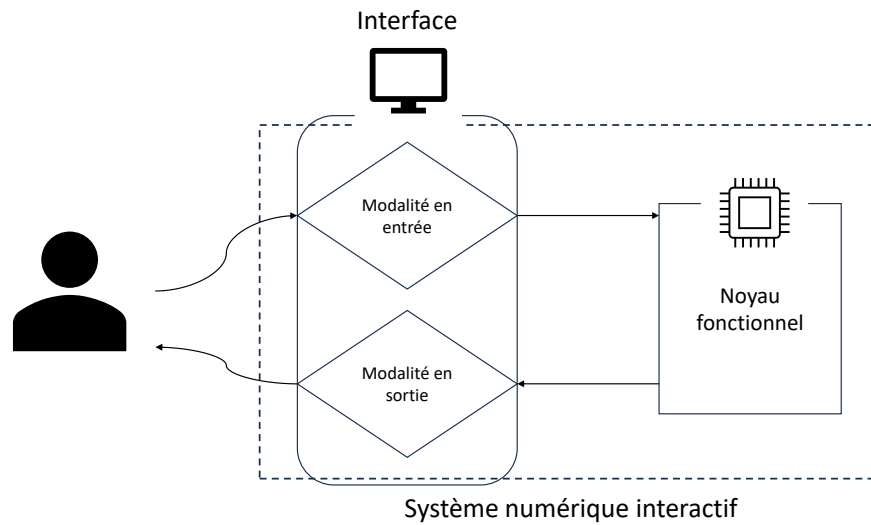


FIGURE 1.1 – Schéma d'un système numérique interactif.

uniquement des commandes vocales en entrée. Si le système utilise plusieurs modalités dans le même sens (par exemple, d'entrée), on le désigne comme multimodal : par exemple, une application GPS sur téléphone portable permettant d'entrer vocalement l'adresse souhaitée à l'aide du microphone, puis de naviguer sur la carte à l'aide de l'écran tactile.

Bien que les modalités d'entrée et de sortie soient indissociables et toutes deux indispensables à un système interactif, nous souhaitons dans le cadre de ce chapitre évaluer différentes modalités d'entrées accessibles aux PDV. De ce fait, dans la suite de ce document, nous nous focalisons exclusivement sur les modalités d'entrée lorsque nous parlons de modalités d'interaction.

1.2.2 Mode de communication

Issue d'études adoptant le point de vue anthropocentré de l'interaction, la notion de mode d'interaction a été introduite par Bellik en 1995 [26] pour spécifier l'interaction entre un utilisateur

et un système numérique. Elle concerne les organes utilisés lors d'une interaction en entrée et fait écho aux sens humains (c'est-à-dire la vue, l'ouïe, le toucher, l'odorat, le goût) utilisés lors d'une interaction en sortie.

Bellik distingue deux modes : le gestuel et l'oral. Le premier correspond aux organes qui permettent d'interagir avec le système par le mouvement de tout ou parties du corps (par exemple, tête, bras, main, yeux). Le second correspond aux organes qui permettent d'interagir avec le système par la production de sons (par exemple, parole, chant, bruits).

1.2.3 Paradigme d'interaction

D'après Dragicevic [62], *“un paradigme d'interaction est un ensemble cohérent de techniques d'interaction qui coopèrent de façon étroite, ou qui reposent sur les mêmes principes techniques ou conceptuels”*. Par exemple, le paradigme d'interaction “Tactile” utilise les écrans tactiles, les gestes et la manipulation directe, que l'on retrouve sur les téléphones portables, les smartwatches, mais aussi certains micro-ondes ou thermostats.

C'est un concept proche (mais distinct) de celui de modèle d'interaction proposé par Beaudouin-Lafon [24], qu'il décrit comme étant un *“ensemble de principes, de règles et de propriétés guidant la conception d'une interface, permettant de combiner des modalités d'interaction de manière congruente afin de définir l'aspect et le ressenti de l'interface”*. Pour cet auteur [25], les paradigmes d'interaction sont des modèles à haut niveau d'abstraction qui définissent la place qu'occupe la machine dans son interaction avec l'utilisateur et dont les trois principaux sont les suivants : la machine comme outil, la machine comme partenaire, la machine comme médium. Dans nos travaux nous nous intéressons au premier : la machine comme un outil contrôlé par le PDV.

Dans le cadre de cette thèse, nous utilisons la définition de Dragicevic pour désigner les paradigmes d'interaction que nous étudions. Nous précisons la nature de chaque paradigme analysé avant de les détailler.

1.2.4 WIMP et post-WIMP

L'acronyme WIMP (c'est-à-dire de l'anglais "Windows, Icons, Menus and Pointing device") est un paradigme d'interaction. Plus spécifiquement, il s'agit d'un type d'interface graphique utilisant la métaphore du bureau et dont le principe est basé sur la manipulation directe [205] d'éléments présents à l'écran (c'est-à-dire les fenêtres, menus et icônes) à l'aide d'un pointeur (c'est-à-dire la souris). Inventé à la fin des années 1970, ce type d'interface s'est largement répandu depuis leur apparition sur Macintosh, mis sur le marché en 1984. Aujourd'hui, il est le paradigme d'interaction le plus utilisé sur ordinateur, alors que les téléphones portables utilisent plutôt une interface dite post-WIMP, dont les bases reposent sur des paradigmes d'interactions différents.

Sous l'impulsion de l'informatique ubiquitaire [236], et grâce aux avancées technologiques, le nombre de modalités d'interaction ne cesse de croître. Les interfaces post-WIMP cherchent à tirer parti de ces nouvelles modalités pour aller au-delà du paradigme WIMP, moins adapté à certaines tâches ou certains contextes d'usage - un usage sans la vue par exemple.

1.2.5 Objets d'intérêt

Les objets sur lesquels l'utilisateur se concentre pour accomplir une tâche donnée, tels que le texte et les dessins d'un document, ou encore, les formules et les valeurs dans un tableur.

1.3 Justification de la grille d'analyse

Dans cette section, nous présentons les modèles conceptuels pertinents à l'établissement d'une grille d'analyse servant à étudier les modalités d'interaction en entrée destinées à un usage sans la vue. Cette section présente donc le cadre théorique de notre étude des modalités d'interaction en entrée pour les PDV.

Une première partie de ce cadre théorique porte sur les modèles caractérisant les dispositifs d'entrée au niveau physique et au niveau logique. La seconde partie de ce cadre théorique

porte sur les modèles décrivant l'interaction. Bien que nos travaux se concentrent sur l'interaction en entrée, les modèles présentés intègrent parfois les aspects d'entrée et de sortie. À la fin de chaque section, nous recentrons les contributions de ces modèles sur notre objet d'étude.

Nous concluons cette partie par une synthèse des éléments issus de ces modèles conceptuels que nous utilisons pour établir notre grille d'analyse.

1.3.1 Modèles conceptuels des dispositifs

Les modèles présentés dans cette section considèrent qu'une interaction se fait par le biais d'un dispositif captant les actions de l'utilisateur (niveau physique) dans le but de réaliser des tâches élémentaires d'interaction (niveau logique).

Niveau physique

Taxonomie de Buxton. Ce modèle conceptuel [40] caractérise les dispositifs physiques en entrée selon deux axes. Le premier regroupe les propriétés que le dispositif peut capturer : la pression, la position et le mouvement. Le deuxième définit le nombre de dimensions capturées pour chaque propriété (par exemple, le mouvement de la souris a deux dimensions).

Buxton précise que sa taxonomie ne permet de décrire que des dispositifs continus actionnables par le geste. De fait, les dispositifs discrets (c'est-à-dire les boutons) et les dispositifs non actionnables par le geste (par exemple, microphones, caméras, EEG) ne peuvent pas être traités à l'aide de cette taxonomie.

Nous retenons de ce modèle conceptuel que la richesse d'un dispositif peut dépendre du nombre de propriétés physiques capturées et du nombre de dimensions pour chacune de ces propriétés.

		Number of Dimensions			
		1	2	3	
Property Sensed	Position	Rotatory Pot	Tablet	3D Joystick	Mechanical
			Touch Tablet		Touch Sensitive
	Motion		Mouse	Trackball	Mechanical
			TASA X-Y Pad		Touch Sensitive
	Pressure	Torque Sending Pressure Pad	Isometric Joystick		

FIGURE 1.2 – Tableau récapitulant la taxonomie de Buxton. Tableau issu de Serrano [201], inspiré de Buxton [40]

Espace de conception de Card, Mackinlay et Robertson. Cet espace [44] traite également des dispositifs physiques d'entrée et étend le modèle de Buxton. Ce modèle permet de décrire un dispositif d'entrée à l'aide d'un vocabulaire de mouvements et d'opérateurs de composition.

Le vocabulaire de mouvements permet de décrire le fonctionnement d'un capteur physique. Il est défini à l'aide d'un sextuple $\langle M, In, Out, R, S, W \rangle$:

- M représente l'opérateur de manipulation appliqué par l'utilisateur ;
- In représente la plage des valeurs d'entrée possibles pour M ;
- Out représente le domaine des valeurs de sortie possibles du dispositif ;
- R définit la correspondance entre les domaines In et Out ;
- S représente l'état actuel du dispositif ;
- W désigne le fonctionnement du dispositif physique.

Les opérateurs de composition permettent de décrire la combinaison des capteurs entre eux au sein d'un même dispositif physique. Trois types de compositions sont identifiés : fusion, disposition et connexion. Pour illustrer ces opérateurs, prenons l'exemple d'une souris. La fusion combine les capteurs des deux axes de défilement X et Y pour former une seule sortie [X,Y]. La

disposition regroupe les boutons et les axes XY dans un même espace physique pour former un objet physique unique (c'est-à-dire la souris). La connexion relie la sortie [X,Y] de la souris à l'écran pour fournir les coordonnées dans son repère.

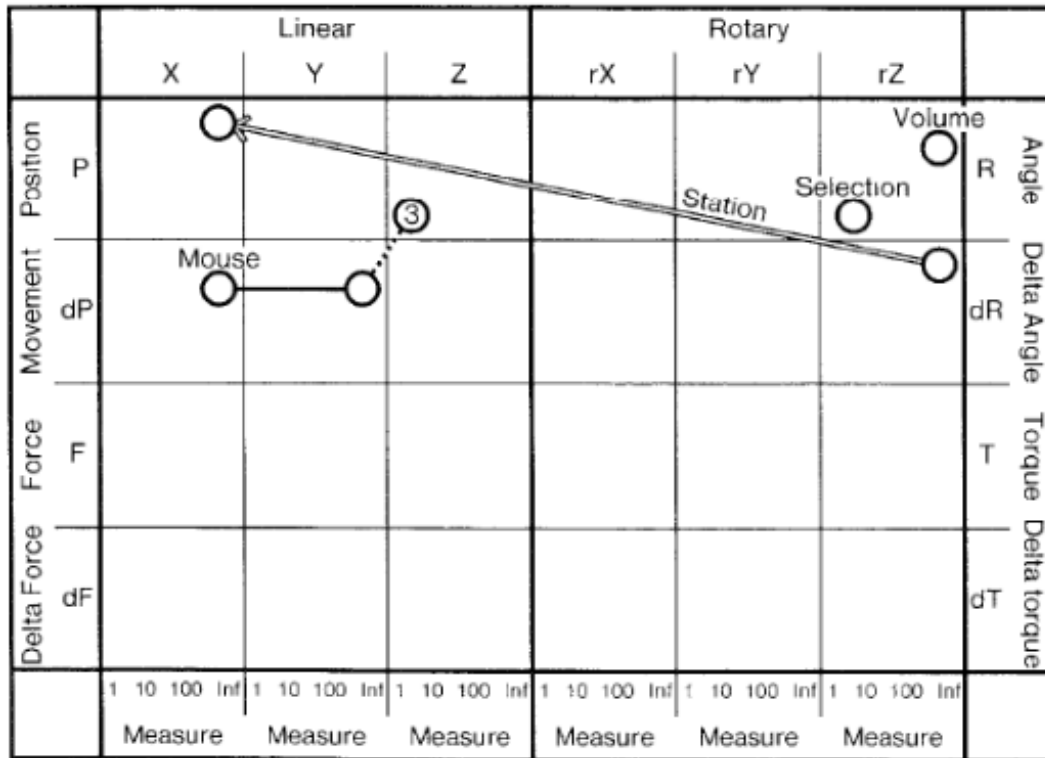


FIGURE 1.3 – Espace de conception des dispositifs en entrée de Card, Mackinlay et Robertson. Figure issue de [44].

La conjugaison du vocabulaire de mouvements avec les opérateurs de composition forme un espace de conception pour les dispositifs en entrée, illustré à la figure 1.3. Les cercles représentent les propriétés captées par les dispositifs. Les traits continus représentent la fusion, les traits discontinus la disposition et les traits fléchés la connexion. L'exemple de la souris est illustré à la figure 1.3 : elle est représentée par deux cercles reliés par un trait continu (c'est-à-dire les XY) et un troisième cercle avec le chiffre 3 relié par une ligne discontinue (c'est-à-dire les boutons). La connexion avec l'écran n'est pas représentée.

À l'instar du modèle de Buxton [40], le modèle de Card et al. ne permet de décrire que des dispositifs captant des informations spatiales (X, Y et Z pour la pression).

Nous retenons en plus de ce modèle conceptuel sa faculté à détailler les dispositifs d'interaction en entrée ainsi que sa représentation graphique facilitant l'évaluation de leur capacité d'expression.

Niveau logique

Taxonomie de Foley. Contrairement aux deux modèles précédents, ce modèle conceptuel [75] propose une taxonomie dans laquelle les dispositifs d'entrée d'un système numérique sont classés selon les tâches d'interaction qu'ils peuvent accomplir. Ces tâches d'interaction correspondent à une "unité d'information" élémentaire entrée dans le système par l'utilisateur et sont au nombre de six :

- Sélection (choisir un élément parmi un ensemble de choix),
- Position (spécifier une position en 1D, 2D ou 3D),
- Trajectoire (spécifier un certain nombre de positions sur une durée),
- Orientation (tourner un objet),
- Quantification (spécifier une valeur numérique),
- Saisie de texte (entrer une chaîne de caractères).

Comme les deux modèles précédents, ce modèle ne considère pas les dispositifs non actionnables, par exemple les commandes vocales qui permettraient d'accomplir toutes les tâches identifiées ne sont pas considérées.

Nous retenons de ce modèle conceptuel l'importance de la relation entre le dispositif et la tâche qu'il permet d'accomplir, ainsi que les 6 unités élémentaires de l'interaction qui définissent des tâches élémentaires d'interaction.

Modèle de Lachenal. Ce modèle conceptuel [119] propose de définir les dispositifs d'entrée (appelés instruments) à l'aide des attributs et des propriétés suivants :

- Actions possibles : pointer, sélectionner, saisir du texte, déplacer, etc.
- Précision : mesure de la précision avec laquelle l'utilisateur peut effectuer les actions possibles à l'aide de l'instrument.
- Résolution : la plus petite variation perceptible de la grandeur mesurée dans des conditions spécifiques.
- Stabilité : un instrument est considéré comme stable si la grandeur mesurée ne varie pas en l'absence d'action de l'utilisateur.
- Adéquation : capacité d'un instrument à faciliter les actions de l'utilisateur.
- Utilisabilité à distance : possibilité d'utiliser l'instrument pour agir sur une surface sans contact physique.
- Maniabilité : attributs de forme, taille, poids et texture de l'entité physique qui influencent sa maniabilité.
- Ré-assignabilité : capacité de l'instrument à être utilisé pour d'autres fonctions que celle pour laquelle il a été conçu (versatilité).

D'après son auteur, ces caractéristiques se restreignent à décrire un dispositif d'entrée haptique. Tout en étant en accord avec les modèles précédents, nous retenons de ce modèle conceptuel les caractéristiques suivantes pour décrire les dispositifs d'entrée : l'adéquation et la ré-assignabilité (versatilité).

1.3.2 Modèles conceptuels de l'interaction

Manipulation directe. Introduite initialement par Sutherland en 1963 avec son Sketchpad [149] et reprise par Shneiderman [205, 207], le modèle conceptuel de la manipulation directe désigne le fait de manipuler des objets graphiques sur l'écran à l'aide d'un dispositif de pointage et peut se résumer par trois principes :

- Représentation continue des objets d'intérêt ;
- Utilisation d'actions physiques pour manipuler les objets d'intérêt au lieu d'expressions selon une syntaxe (comme les interfaces reposant sur des lignes de commande) ;
- Rapidité et réversibilité des actions physiques sur les objets d'intérêt.

La manipulation directe offre plusieurs avantages tels qu'une réduction de la distance entre l'image mentale de la tâche et sa réalisation effective, un apprentissage rapide, une réduction des erreurs et la possibilité de défaire des actions. Cependant, dans la pratique, l'interaction n'est pas toujours directe. Les interfaces WIMP par exemple, omniprésentes sur ordinateur personnel, dérogent directement au principe #2 et utilisent souvent une manipulation indirecte des objets d'intérêt par le biais de la manipulation - directe cette fois - d'éléments d'interface tels que les menus et les boîtes de dialogue. Les interfaces sur téléphones portables, quant à elles, sont très limitées en terme d'espace d'affichage et dérogent souvent au principe #1 : elles ont souvent recours à des éléments d'interface (par exemple, bouton hamburger, onglets, barres de défilement) permettant d'afficher ou de cacher les objets d'intérêt selon les besoins de l'utilisateur afin d'économiser la surface d'affichage.

Nous retenons de ce modèle conceptuel les avantages qu'une interface gagnerait à proposer une interaction la plus directe possible.

Interaction instrumentale. Le modèle conceptuel de l'interaction instrumentale, introduit par Beaudouin-Lafon [24], est un modèle d'interaction qui étend les principes de la manipulation directe de Shneiderman. Il propose la notion d'instrument d'interaction, un instrument composé d'une partie physique et d'une partie logique, agissant comme médiateur entre les actions de l'utilisateur et les objets d'intérêt. L'interaction se produit donc à deux niveaux : entre l'utilisateur et l'instrument physique, et entre l'instrument logique et l'objet d'intérêt. Par exemple, une souris peut être l'instrument physique manipulant une barre de défilement, l'instrument logique, servant à faire défiler un texte.

Ce modèle met en évidence l'apport de représenter les conséquences des actions effectuées par l'utilisateur non seulement sur l'objet d'intérêt, mais aussi sur l'instrument, afin de faciliter la compréhension du système interactif. Le retour d'information vers l'utilisateur s'effectue à trois niveaux :

— retour de l'action sur l'instrument physique (par exemple, position de la souris) ;

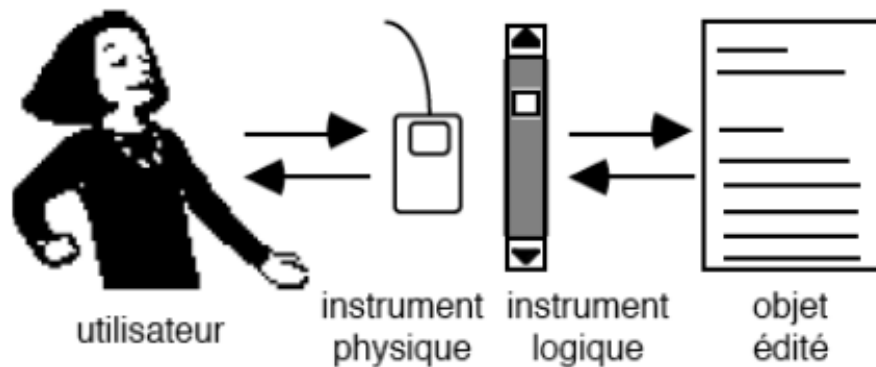


FIGURE 1.4 – Illustration du modèle de l'interaction instrumentale de Beaudouin-Lafon. Figure issue de [24].

- retour via la représentation de l'instrument logique (par exemple, mise en évidence de la barre de défilement lorsque le pointeur est dessus à l'aide d'un effet visuel) ;
- retour de l'effet de l'interaction sur l'objet d'intérêt (par exemple, défilement du document).

En outre, ce modèle définit trois propriétés essentielles des instruments d'interaction :

- Le degré d'indirection représente le décalage spatial et temporel entre l'instrument et l'objet d'intérêt. Par exemple sur ordinateur, fermer une fenêtre à l'aide de l'icône en forme de croix a un degré d'indirection plus faible que de fermer cette même fenêtre à l'aide du menu ;
- Le degré d'intégration est le rapport entre le nombre de dimensions physiques (voir Buxton 1.2) et le nombre de dimensions logiques manipulées par l'instrument. Par exemple, une barre de défilement est un dispositif 1D contrôlé par une souris 2D, ce qui signifie que son degré d'intégration est de $1/2$. Le degré d'intégration peut dépasser 1 : par exemple, le contrôle de 3 angles de rotation avec une souris 2D a un degré d'intégration de $3/2$;
- Le degré de compatibilité reflète la similarité entre les manipulations effectuées sur l'instrument et la réaction de l'objet manipulé. Par exemple, choisir une couleur directement sur un cercle chromatique dans une application de dessin a un degré de compatibilité plus élevé que d'entrer la valeur hexadécimale de cette couleur.

À l'instar de notre définition de la modalité d'interaction [156], ce modèle conceptuel présente la notion d'instrument d'interaction possédant deux niveaux d'abstraction : ici l'instrument physique incarne le dispositif et l'instrument logique matérialise le langage d'interaction des modalités d'interaction. Nous retenons de ce modèle l'intérêt d'un couplage fort entre l'action physique de l'utilisateur et l'effet opéré sur le système à plusieurs niveaux. Nous retenons en particulier la deuxième propriété pour évaluer le couplage <dispositif, langage d'interaction>d'une modalité d'interaction.

1.4 Synthèse

Nous avons adopté une approche ascendante pour présenter les modèles conceptuels de l'interaction. En effet, nous avons commencé par examiner les modèles traitant des dispositifs physiques et logiques, qui sont les éléments de base. Ensuite, nous avons étudié les modèles d'interaction. Afin de construire notre grille d'analyse des modalités d'interaction existantes pour les PDV, nous retenons des modèles présentés :

- Pour évaluer le dispositif de la modalité, nous considérons le nombre de propriétés physiques capturées et le nombre de dimensions pour chacune de ces propriétés [40, 44] ;
- Nous examinons également la maniabilité [119] de la modalité en considérant son dispositif ;
- Nous étudions l'adéquation du dispositif au langage d'interaction de la modalité en considérant le degré d'intégration [24] ;
- Nous étudions aussi l'adéquation de la modalité [119] aux tâches d'interaction [75].

1.5 Grille d'analyse des paradigmes et modalités d'interaction

Pour étudier les paradigmes et modalités d'interaction proposés pour l'interaction par des personnes ayant une déficience visuelle, nous considérons les quatre paradigmes d'interaction les plus couramment employés par les PDV : le Tactile, le Vocal, le Tangible, le Gestuel. Au sein de chaque paradigme, nous analysons les modalités d'interaction utilisables en situation de non-retour visuel afin d'en extraire des caractéristiques générales. Il serait illusoire de vouloir étudier quantitativement chaque modalité utilisée pour chaque tâche d'interaction, dans chaque contexte, avec chaque dispositif d'entrée possible. Nous nous limitons à identifier les avantages et les inconvénients de modalités d'interaction et d'en extrapoler les caractéristiques générales des paradigmes auxquels elles appartiennent.

Afin de pouvoir comparer ces modalités d'interaction nous créons une grille d'analyse dont certains critères sont issus des modèles conceptuels présentés précédemment.

Tandis que nous nous appuyons sur des critères issus des modèles conceptuels présentés dans la section précédente, la structure globale de la grille d'analyse repose sur la définition de l'utilisabilité déclinée en deux critères que sont l'efficacité et la satisfaction.

L'utilisabilité est définie par la norme ISO 9241-11 comme « le degré selon lequel un produit peut être utilisé, par des utilisateurs identifiés, pour atteindre des buts définis avec **efficacité**, **efficience** et **satisfaction**, dans un contexte d'utilisation spécifié ».

Nielsen [154] place la notion d'utilisabilité au sein d'un concept plus large d'acceptabilité d'un système. Il décrit ce modèle comme un arbre où l'acceptabilité d'un système est divisée en acceptabilité pratique et sociale. L'acceptabilité pratique englobe plusieurs caractéristiques, dont la fiabilité, les performances et l'utilité, que Nielsen subdivise ensuite en **utilité** (utility) et **utilisabilité** (usability). Il décompose à nouveau le concept d'utilisabilité en cinq critères princi-

poux d'un système utilisable :

- L'efficience
- La satisfaction (subjective)
- La facilité d'apprentissage
- La facilité de mémorisation
- La fiabilité

Ces critères sont à mettre en relation avec ceux proposés par Foley [75] :

- Le temps de complétion de la tâche
- La précision en entrée
- Le plaisir de l'utilisateur
- La facilité d'apprentissage et le temps de rappel du fonctionnement
- La charge cognitive des mémoires à court et à long terme
- La fatigue provoquée par l'utilisation et le taux d'erreur induite par la fatigue
- Le "naturel" du système : à quel point son utilisation est basée sur un concept d'une autre activité déjà connue de l'utilisateur
- Sa taille, l'espace de travail

Pour Nielsen [154], Foley [75] et autres auteurs [6, 41, 74, 24], l'adéquation entre la modalité d'interaction et la tâche d'interaction est primordiale afin de maximiser l'utilisabilité. En effet, moins la modalité d'interaction est compatible avec la tâche d'interaction, plus la modalité d'interaction devient complexe : par exemple, dessiner un point dans un plan en 2D avec une souris est facile, mais faire pivoter un objet 3D requiert davantage de manipulation. Dans ce cas précis, l'incompatibilité vient du faible degré d'intégration du dispositif (c'est-à-dire 3/2 = contrôle de 3 angles de rotation avec une souris 2D).

Or, tous les dispositifs d'entrée peuvent réaliser toutes les tâches qu'un système peut offrir. Il est en effet possible de tout faire sur un ordinateur avec seulement une souris - y compris entrer du texte, à l'aide d'un clavier virtuel. À l'extrême, un dispositif doté d'un seul bouton qui

émet des 0 et des 1 selon que l'on appuie dessus plus ou moins longtemps (à l'instar du Morse) permettrait de produire du code binaire interprétable par un ordinateur, et donc de le contrôler. Cette modalité d'interaction serait toutefois extrêmement lente et inefficace, et ce quelle que soit la tâche d'interaction à effectuer. Malgré tout, nous considérons que toutes les modalités d'interaction sont **efficaces**, dans le sens où elles permettent de réaliser toutes les tâches d'interaction. Nous ne tenons donc pas compte de cette caractéristique dans notre évaluation.

Nous évaluons l'efficacité des modalités d'interaction, au regard des critères suivants :

- L'adéquation entre la modalité d'interaction étudiée et les tâches d'interaction de Foley [75], que nous synthétisons, comme dans le modèle de Lachenal [119], à l'aide des degrés d'indirection, d'intégration et de compatibilité issus du modèle de Beaudouin-Lafon. [24]. Nous cherchons à tendre vers une interaction la plus directe possible, pour profiter des caractéristiques présentées par Shneiderman [205, 207].
- Les charges de travail perceptuelles, cognitives et motrices, taxonomie proposée par Foley [75] et basée sur des travaux en psychologie [46, 129, 120]. sont considérées à l'aide de concepts spécifiques au paradigme étudié comme le multiplexage spatial / temporel dans la section "Paradigme Tactile". Nous visons ainsi à estimer les forces et faiblesses de chacun des paradigmes étudiés selon des concepts liés à la charge de travail perceptuelle, cognitive et motrice.
- La pénibilité, que nous évaluons en fonction de la taille de l'effecteur prépondérant de la modalité d'interaction étudiée (par exemple, le doigt, le bras, le corps entier). Nous estimons simplement que plus l'effecteur est grand, plus le geste requiert d'énergie, plus il est pénible à effectuer. Nous pondérons notre évaluation selon que la modalité d'interaction requiert d'utiliser de la force pour manipuler (par exemple, pousser, porter) des objets ou non.

Nous évaluons aussi la **satisfaction** à l'usage, selon les critères suivants :

- La portabilité indique si le dispositif est raisonnablement transportable ou non. La taille

de référence étant celle du téléphone portable et de la tablette, nous considérons que tout dispositif sensiblement plus grand n'est pas désirable pour les PDV.

- La versatilité [119] mesure la capacité du dispositif à être utilisable pour plusieurs objectifs et dans différents contextes. Par exemple, une souris peut être utilisée pour une multitude de tâches (par exemple, mise en page, jeux, dessin), alors qu'un jeu d'échec connecté ne sert qu'à jouer aux échecs. L'intérêt des PDV pour le téléphone portable tient en sa versatilité, qui lui permet de remplacer une multitude d'autres TA. Nous considérons donc que tout dispositif supplémentaire n'est pas désirable pour les PDV s'il n'offre pas un haut niveau de versatilité.
- La discrétion reflète à quel point l'utilisateur peut se servir de la modalité d'interaction en public sans que cela ne soit ostensible. Elle dépend du mode d'interaction (c'est-à-dire vocal ou gestuel) [26], de la taille et l'emplacement du dispositif d'entrée (par exemple, casque de réalité augmentée, smartwatch), ainsi que de l'espace de travail nécessaire à l'utilisation de la modalité d'interaction (par exemple, amplitude d'un geste du bras).
- La confidentialité indique à quel point l'utilisateur peut se servir de la modalité d'interaction sans que le public ne puisse connaître le contenu de l'interaction. Elle dépend beaucoup du degré de discrétion et de la complexité du langage d'interaction.
- La discrimination, déterminante pour les PDV, marque à quel point l'utilisateur peut se servir de la modalité d'interaction sans que le public n'interprète cela comme un indicateur de déficience. Elle dépend du degré de discrétion et du niveau de banalité de la modalité d'interaction utilisée.

Nous synthétisons ces critères au sein du tableau suivant, où chaque colonne correspond à un critère d'analyse et de comparaison des modalités :

Paradigme									
Articles	Modalité(s)	Fait pour / testé par PDV	Adéquation	Charges de travail	Pénibilité	Portabilité	Versatilité	Discrétion et Confidentialité	Discrimination
-	-	-	-	-	-	-	-	-	-

FIGURE 1.5 – Tableau reprenant les éléments de notre grille d'analyse.

Cette grille d'analyse contient des critères génériques qui peuvent être utilisés pour évaluer tout type de modalités d'interaction, mais dans le cadre de nos travaux sur les PDV, les colonnes Adéquation, Versatilité et Discrimination nous paraissent particulièrement importantes.

Les modalités étudiées au sein des paradigmes d'interaction sont issues de la littérature et portent sur des interactions en situation de non-retour visuel. Toutes ne sont pas destinées aux PDV, et parmi celles qui le sont, toutes n'ont pas été testées auprès des PDV. Nous précisons à chaque fois que nous étudions une modalité de quel type il s'agit. Nous n'affirmons pas qu'avoir une cécité est la même chose qu'être placé dans une situation dans laquelle l'utilisation de la vue est impossible, mais nous estimons que les dispositifs conçus pour être utilisés sans la vue sont utilisables par les PDV.

Dans la section suivante et selon la grille d'analyse, nous présentons les 4 paradigmes d'interaction les plus utilisés par les PDV, en commençant par le tactile. Malgré ses désavantages en termes d'accessibilité pour les PDV, ce paradigme reste le plus utilisé car il s'agit du paradigme d'interaction par défaut des téléphones portables et notre objet d'étude.

1.6 Analyse des paradigmes et modalités d'interaction pour les Personnes ayant une Déficience Visuelle

1.6.1 Le paradigme Tactile

Le paradigme Tactile implique l'utilisation de gestes (par exemple, TAP, SWIPE) sur une surface tactile, de taille variable selon le dispositif (par exemple, téléphone portable, tablette, table tactile). Dans le cadre de nos travaux, nous nous focalisons sur la surface tactile la plus commune - celle des téléphones portables et tablettes - qui agit aussi comme un écran. Nous représentons ce paradigme comme utilisant exclusivement la modalité d'interaction suivante : <écran tactile, grammaire de gestes>.

A minima, l'écran tactile capte la Position en 2 dimensions. Voici sa représentation dans l'espace de Card et al. [44] :

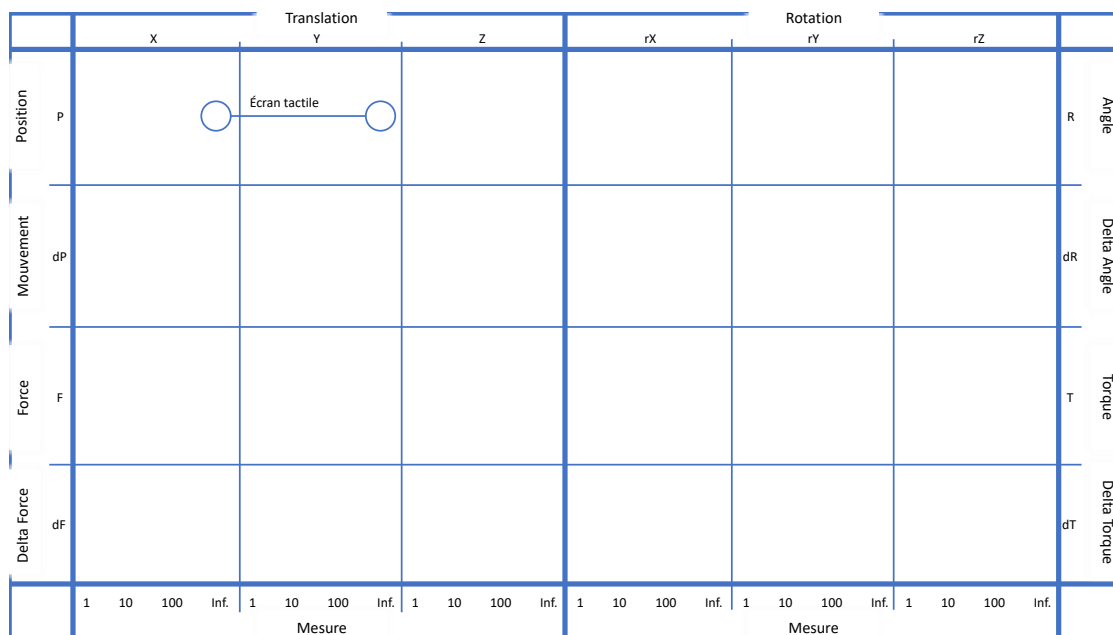


FIGURE 1.6 – Schéma représentant le dispositif "écran tactile" dans l'espace de conception des dispositifs en entrée de Card. Inspiré de [44].

Dans un contexte d'utilisation 1) sur écran et 2) par des personnes voyantes, ce paradigme

permet d'interagir directement (c'est-à-dire manipulation directe) avec les éléments graphiques présents à l'écran [206], généralement à l'aide des doigts. La vue permet de profiter de la versatilité d'affichage des écrans. En effet, une infinité d'éléments peut être affichée d'une infinité de façon afin de former une interface graphique spécifique à chaque application. La vue est donc essentielle dans ce cas pour pouvoir appréhender instantanément la disposition des éléments graphiques de l'interface et viser ces éléments.

Dans un contexte d'utilisation par des PDV néanmoins, l'écran tactile perd de son intérêt. Sans la vue, il devient difficile de viser les éléments composant une interface graphique [85, 86]. Sur téléphones portable et tablette les lecteurs d'écran VoiceOver (VO) et TalkBack (TB), présentés ci-dessous dans cette section, permettent d'interagir avec le système à l'aide de deux modalités d'interaction, résultats d'une recherche extensive sur le sujet [174, 223, 243, 106, 142, 8] : l'exploration et la navigation "pas à pas".

L'exploration consiste à parcourir l'écran avec le doigt jusqu'à trouver l'élément voulu. Lorsque le doigt passe sur un élément de l'interface l'élément est considéré comme sélectionné. Une commande supplémentaire (c'est-à-dire un DOUBLE TAP n'importe où sur l'écran) permet d'interagir avec l'élément (par exemple, cocher la case, cliquer sur le lien, ouvrir l'application). Si cette méthode permet à l'utilisateur d'utiliser l'interface dans sa disposition d'origine, dans ce cas de figure, la versatilité de l'écran devient un désavantage. Puisque chaque application a une interface graphique différente, cela nécessite d'explorer - ou d'apprendre par cœur - la disposition de chaque application, celle-ci pouvant aussi être variable au cours de l'interaction (par exemple, nouvelle icône d'un fichier, position de défilement variable). Par ailleurs, quand bien même la disposition des éléments serait connue, viser ces éléments et les atteindre du premier coup reste (selon la taille de la cible) un exercice difficile sans la vue, et requiert souvent d'ajuster la position du doigt [86, 102]. Cette tâche de pointage est d'autant plus difficile que la surface plane et lisse de l'écran n'offre aucun indice haptique (par exemple, relief²) qui aiderait l'utilisateur à se repérer sur l'écran autrement que visuellement [69, 38]. Le pointage direct reste donc utilisable, mais perd en efficacité car il faut chaque fois procéder à une étape préliminaire

2. Harmonizer for iPad : <http://www.musanim.com/harmonizer/>

d'exploration et d'ajustement.

La navigation pas-à-pas consiste à “transformer” l’interface en une liste de ses éléments. Cette liste peut être naviguée à l’aide de commandes discrètes afin de sélectionner l’un des éléments. Une commande supplémentaire (c’est-à-dire un DOUBLE TAP n’importe où sur l’écran) permet d’interagir avec l’élément de la même manière qu’avec la modalité d’interaction précédente.

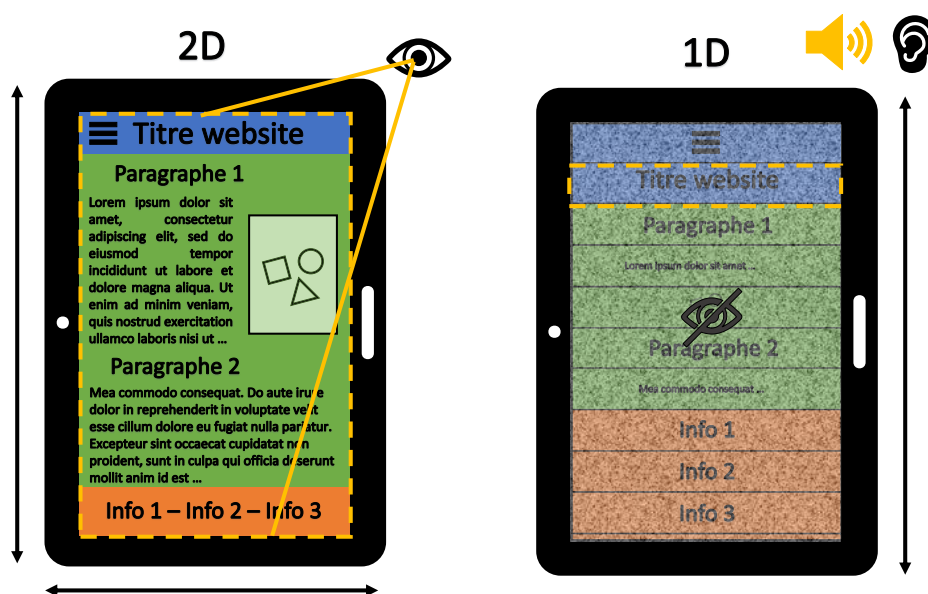


FIGURE 1.7 – Schéma illustrant la transformation d'une interface graphique (2D) en une liste d'éléments (1D).

Cette méthode permet à l'utilisateur de s'affranchir de la disposition de l'interface graphique. Ce faisant, cette modalité d'interaction s'affranchit aussi des avantages de la manipulation directe et du positionnement des éléments dans l'espace, peu accessibles sans la vue, mais qui offrent la possibilité à une interface de proposer plusieurs éléments en même temps à l'utilisateur en les répartissant dans l'espace de travail - dans ce cas, la surface 2D de l'écran. C'est ce qu'on appelle le **multiplexage spatial** [74] : dans un système, chaque fonction possède son élément associé occupant son propre espace à un moment donné. Avec cette propriété aug-

mente l'**expressivité** (c'est-à-dire le nombre d'options possibles à un instant donné) et la **bande passante** (c'est-à-dire la quantité d'information transmissible au système à un instant donné). Un exemple classique est celui du clavier : il possède un grand nombre de touches accessibles au même moment, et il est possible de les utiliser simultanément pour envoyer des informations plus ou moins complexes (par exemple, Maj + A pour entrer un A majuscule, Ctrl + Alt + Suppr pour ouvrir une fenêtre de contrôle de l'ordinateur sous Windows). Un autre exemple sur écran tactile : les claviers Braille sont dotés de 6 éléments répartis sur l'écran sur lesquels on place les doigts et qui agissent comme des touches (figure 1.8). Elles peuvent être utilisées simultanément, et les différentes combinaisons permettent d'écrire du texte [209].

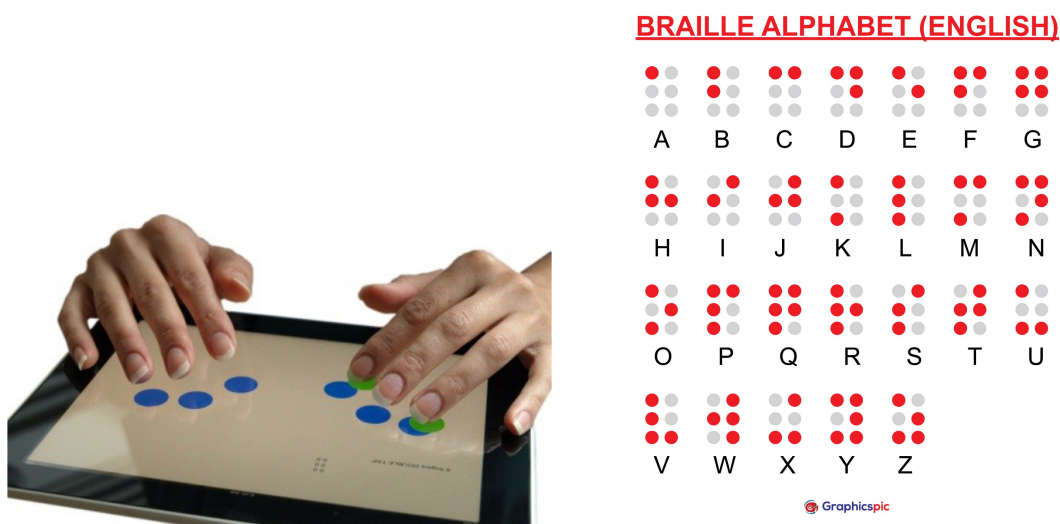


FIGURE 1.8 – **Gauche** : Photographie du dispositif BrailleTouch sur tablette. Illustration issue de [209]. **Droite** : Alphabet Braille.

Les modalités d'interactions accessibles aux PDV présentées précédemment (c'est-à-dire l'exploration et le pas-à-pas) font donc l'impasse sur ce multiplexage spatial, car il ne permettent fondamentalement d'interagir avec les éléments graphiques de l'interface que de façon séquentielle (c'est-à-dire un à un) et, dans le cas du pas-à-pas, linéaire (c'est-à-dire dans un ordre fixe), à l'aide d'un ensemble réduit de commandes d'entrée (c'est-à-dire précédent / suivant / actionner). C'est ce qu'on appelle le **multiplexage temporel** : un même dispositif ou une même commande peut servir à plusieurs fonctions, selon l'application ou le contexte - elle varie

au cours du temps. Par exemple, sur les claviers 3x4 de téléphones mobiles physiques, une même touche sert à écrire différentes lettres en fonction du nombre de fois où elle est pressée dans un intervalle de temps restreint. Sur ce type de clavier, écrire le mot "oui" nécessite d'appuyer sur les touches suivantes : trois fois sur le "6", puis deux fois sur le "8", puis trois fois sur le "4".



FIGURE 1.9 – Illustration d'un clavier 3*4 de téléphone mobile physique. Chaque touche correspond à plusieurs commandes possibles.

Dans une interface graphique, notamment sur écran, le multiplexage spatial est omniprésent. Toutefois, en raison de leur nature numérique, les interfaces graphiques sont reconfigurables et leur multiplexage spatial n'est donc pas fixe : sur téléphones portable par exemple, une même zone d'écran peut tour à tour héberger un élément textuel n'ayant aucune commande associée, une icône permettant d'ouvrir une application, un bouton, ou rien du tout. Dans ce cas, un TAP simple déclenche ce sur quoi le doigt est posé.

Dans notre contexte de navigation sur téléphones portable utilisant les modalités d'interaction accessibles aux PDV avec VO ou TB, c'est le DOUBLE TAP sur l'écran, indépendamment de sa position sur l'écran, qui devient l'unique commande déclencheuse de l'action préalablement sélectionnée. La différence entre ce DOUBLE TAP et le TAP simple d'une interface graphique classique tient à ce que la position de ce dernier sur l'écran a une signification, là où le premier n'en a pas. Pour discriminer la fonction de ces DOUBLE TAP, il faut donc passer

par une étape préalable de sélection à l'aide d'une des deux modalités d'interaction vues précédemment (l'exploration ou le pas à pas) dont l'expressivité et la bande passante est faible, bien moindre que celle basée sur les interfaces graphiques. À l'instar de notre exemple précédent sur les dispositifs d'interaction pouvant réaliser toutes les tâches, avec ces techniques d'interaction il est possible d'accéder à toutes les fonctionnalités du téléphones portable (sans forcément pouvoir les utiliser correctement, si leur usage est basé sur la vue, par exemple cadrer une photo). Atteindre ces fonctionnalités peut être long et fastidieux en raison de la taille et la complexité de l'arborescence du menu en permettant l'accès [145], que renforcent des problèmes d'utilisabilité plus classiques [57] tels qu'une difficulté à performer les DOUBLE TAP [187] ou une mauvaise architecture de l'information.

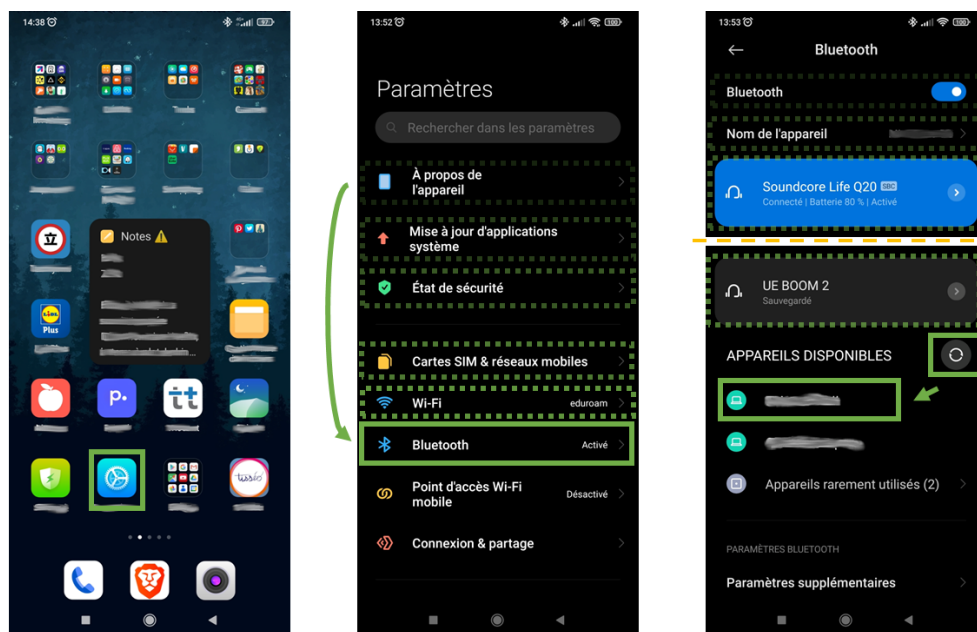


FIGURE 1.10 – Exemple de trajectoire d'interaction sur téléphone portable (Android) pour ajouter un périphérique bluetooth.

Cet ensemble d'actions à accomplir pour atteindre un objectif est appelé **trajectoire d'interaction**. La figure 1.10 illustre la trajectoire d'interaction nécessaire pour appairer un nouveau périphérique Bluetooth au téléphone portable en utilisant la technique d'interaction pas-à-pas. L'utilisateur doit tout d'abord localiser l'icône permettant d'accéder aux paramètres, puis il doit

naviguer dans les paramètres un par un jusqu'à trouver celui du Bluetooth, puis il doit descendre dans le nouvel écran jusqu'à trouver l'icône permettant de chercher un nouveau périphérique, puis sélectionner le nom du périphérique à ajouter. Alors qu'un utilisateur voyant nécessiterait 3 ou 4 actions, un PDV devra en effectuer une douzaine, sous réserve qu'il ait déjà trouvé l'icône "Paramètres" et qu'il connaisse exactement l'emplacement de la fonction qu'il recherche (c'est-à-dire chercher et ajouter un nouveau périphérique Bluetooth).

Pour raccourcir cette trajectoire d'interaction, les développeurs de VO et de TB ont ajouté d'autres techniques d'interaction utilisant les propriétés des écrans tactiles. Ces raccourcis utilisent le multitouch, la répétition et le dessin pour ajouter de l'expressivité et permettre un accès plus rapide à certaines fonctions du téléphone portable les plus communes. Ces systèmes permettent aussi d'enregistrer des raccourcis personnalisés fonctionnant comme des macros ou des marque-pages.

VO, le lecteur d'écran intégré aux appareils d'Apple (c'est-à-dire iPhones, iPad et Mac), utilise essentiellement le multitouch et la répétition pour augmenter le nombre de commandes possibles. Sur iOS 16 (septembre 2022), lorsque VO est activé, les gestes standard effectués sur l'écran fonctionnent différemment et il est alors possible d'utiliser l'exploration ou le pas-à-pas pour naviguer dans l'écran de haut en bas et de gauche à droite. Les gestes permettant de naviguer sont les suivants et proviennent du guide d'utilisation fourni par Apple³ :

3. Guide d'utilisation de l'iPhone - Apprendre les gestes VoiceOver sur l'iPhone : <https://support.apple.com/fr-fr/guide/iphone/iph3e2e2281/16.0/ios/16.0>

Action	Geste
Sélectionner et énoncer un élément	Touchez l'élément
Sélectionner l'élément suivant	Balayez vers la droite
Sélectionner l'élément précédent	Balayez vers la gauche
Sélectionner le premier élément à l'écran	Touchez le haut de l'écran avec quatre doigts
Sélectionner le dernier élément à l'écran	Touchez le bas de l'écran avec quatre doigts
Énoncer tout le contenu de l'écran en partant du haut	Balayez vers le haut avec deux doigts
Énoncer tout le contenu de l'écran en partant de l'élément sélectionné	Balayez vers le bas avec deux doigts
Interrompre ou poursuivre l'énonciation	Touchez avec deux doigts
Énoncer des informations supplémentaires, comme la position dans une liste ou l'état de sélection du texte	Touchez avec trois doigts

Les gestes permettant de faire défiler un document ou une page sont les suivants :

Action	Geste
Faire défiler d'une page vers le haut	Balayez vers le bas avec trois doigts
Faire défiler d'une page vers le bas	Balayez vers le haut avec trois doigts
Faire défiler d'une page vers la gauche	Balayez vers la droite avec trois doigts
Faire défiler d'une page vers la droite	Balayez vers la gauche avec trois doigts

Les gestes permettant d'interagir avec un élément de l'interface sont les suivants :

Action	Geste
Sélectionner un élément	Toucher
Activer l'élément sélectionné	Toucher deux fois
Toucher deux fois l'élément sélectionné	Toucher trois fois
Faire glisser un curseur	Touchez le curseur pour le sélectionner, puis balayez vers le haut ou le bas avec un doigt ; ou touchez deux fois le curseur et maintenez le doigt dessus jusqu'à entendre trois tonalités montantes, puis faites glisser le curseur
Lancer une action ou arrêter l'action en cours (par exemple, lancer ou mettre en pause la lecture d'un morceau ou d'une vidéo, prendre une photo dans Appareil photo, lancer ou arrêter un enregistrement, lancer ou arrêter le chronomètre)	Touchez deux fois avec deux doigts
Fermer une alerte ou revenir à l'écran précédent	Faites un zigzag avec deux doigts (déplacez deux doigts trois fois dans un mouvement de va-et-vient pour faire un « z »)
Modifier le libellé d'un élément afin de le retrouver plus facilement	Touchez deux fois avec deux doigts et maintenez la position

VO propose également d'utiliser le rotor, un menu circulaire proposant des commandes globales et contextuelles. Selon le guide d'utilisation d'Apple, le rotor fonctionne de la sorte :

1. Lorsque VO est activé, décrivez un cercle avec deux doigts sur l'écran comme si vous faisiez tourner un cadran. Si vous préférez utiliser un doigt de chaque main, faites glisser en même temps un doigt vers le haut et l'autre vers le bas. VO énonce le réglage du

rotor. Continuez de faire tourner vos doigts pour entendre d'autres réglages. Cessez de les faire tourner lorsque vous entendez le réglage voulu.

2. Faites glisser votre doigt vers le haut ou le bas sur l'écran pour utiliser le réglage. Les réglages disponibles dans le rotor et leurs effets varient selon la tâche que vous réalisez. Par exemple, si vous choisissez Entêtes pendant la consultation d'une page web, balayez vers le bas ou le haut pour déplacer le curseur VO vers l'en-tête suivant ou précédent.

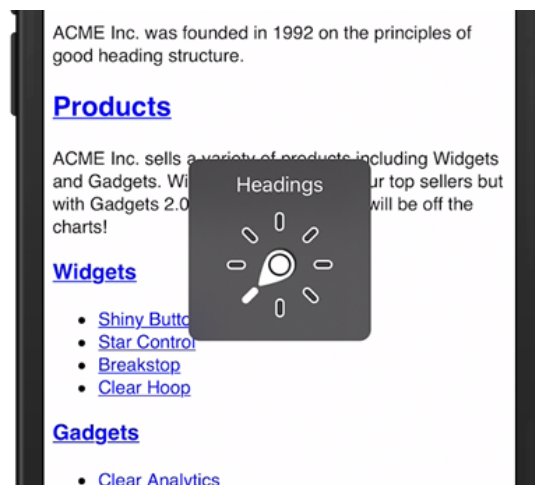


FIGURE 1.11 – Capture d'écran du rotor de VoiceOver, sous iOS.

Il permet de modifier les réglages, naviguer sur l'écran d'un élément à un autre, sélectionner des méthodes de saisie spéciales comme le Braille, et bien d'autres options. Le rotor sert notamment à sélectionner et modifier du texte. En voici la marche à suivre, issue du guide d'utilisation d'Apple :

- Pour déplacer le point d'insertion :
 - Balayez vers le haut ou le bas pour déplacer le point d'insertion vers l'avant ou l'arrière dans le corps du texte.
 - Utilisez le rotor pour choisir si vous voulez déplacer le point d'insertion d'un caractère, d'un mot ou d'une ligne.
 - Pour vous rendre au début ou à la fin, touchez deux fois le texte. VoiceOver émet un son lorsque le point d'insertion se déplace et énonce le caractère, le mot ou la

ligne sur lesquels il passe. Lorsque le déplacement est vers l'avant par mot, le point d'insertion est placé à la fin du mot énoncé, avant l'espace ou le signe de ponctuation qui suit. Lorsque le déplacement est vers l'arrière, le point d'insertion est placé à la fin du mot précédent, avant l'espace ou le signe de ponctuation qui suit.

- Pour placer le point d'insertion après le signe de ponctuation suivant un mot ou une phrase : Utilisez le rotor pour basculer en mode caractère. Lorsque le point d'insertion est déplacé ligne par ligne, VoiceOver énonce chaque ligne sur laquelle il passe. Lorsque le déplacement est vers l'avant, le point d'insertion est placé au début de la ligne suivante (sauf lorsque vous avez atteint la dernière ligne d'un paragraphe, auquel cas le point d'insertion est placé à la fin de la ligne énoncée). Lorsque le déplacement est vers l'arrière, le point d'insertion est placé au début de la ligne énoncée.
- Pour sélectionner du texte : Effectuez une ou plusieurs des opérations suivantes :
 - Réglez le rotor sur Sélection de texte, balayez vers le haut ou le bas pour choisir Caractère, Mot, Ligne ou Phrase, puis balayez vers la gauche ou la droite pour reculer ou avancer.
 - Réglez le rotor sur Modifier, balayez vers le haut ou le bas pour choisir Sélectionner ou Tout sélectionner, puis touchez deux fois l'écran. Si vous choisissez Sélectionner, le mot le plus proche du point d'insertion est sélectionné lorsque vous touchez deux fois. Pour augmenter ou réduire la sélection, faites un zigzag avec deux doigts (déplacez deux doigts trois fois dans un mouvement de va-et-vient pour faire un « z ») pour fermer le menu local, puis pincez.
- Pour couper, copier ou coller : Réglez le rotor sur Modifier, sélectionnez le texte, balayez vers le haut ou le bas pour choisir Couper, Copier ou Coller, puis touchez deux fois l'écran.
- Pour corriger les mots mal orthographiés : Réglez le rotor sur Mots mal orthographiés, puis balayez vers le haut ou le bas pour passer au mot mal orthographié précédent ou suivant. Balayez vers la gauche ou la droite pour choisir une suggestion de remplacement, puis touchez deux fois pour l'utiliser.
- Pour annuler : Secouez l'iPhone, balayez vers la gauche ou vers la droite pour choisir l'action à annuler, puis touchez deux fois.

TalkBack⁴ (TB), le lecteur d'écran intégré aux appareils Android, utilise des gestes similaires à VO pour explorer l'interface et naviguer dans le système : des SWIPE vers la gauche ou vers la droite, des TAP et double-TAP pour sélectionner et activer les éléments (c'est-à-dire pas-à-pas), et la possibilité de toucher ou faire glisser un doigt sur l'écran pour sélectionner les éléments présentés dans l'interface (c'est-à-dire exploration). Toutefois, contrairement à VO, il propose essentiellement des raccourcis sous forme de dessin : l'utilisateur doit effectuer différentes formes de "L" pour initier certaines commandes (Gauche puis Bas par exemple, pour montrer les notifications) ou accéder à différents endroits du système (par exemple, Haut puis Gauche pour l'écran d'accueil). Depuis la version 9.1 (c'est-à-dire mars 2021) toutefois, TB permet aussi de faire des gestes à plusieurs doigts à la place de certains dessins en "L".

D'après un sondage de 2021 effectué par l'Institute for Disability Research, Policy, and Practice⁵ (USA) auprès de 1568 PDV dans le monde, 90% utilisent un lecteur d'écran sur téléphone portable. Parmi eux, 71,5% utilisent VO et 29% utilisent TB. Ces deux lecteurs d'écran - et notamment VO - représentent aujourd'hui l'absolue majorité des usages sur téléphones portables et tablettes, et constituent la référence en matière de modalité d'interaction accessible sur ces supports.

Pourtant, malgré ces modalités d'interaction supplémentaires que sont les raccourcis, de nombreux problèmes persistent [57] : 1) les gestes servant à dessiner des angles ou des formes géométriques sont difficiles à faire sans la vue [109, 186, 137, 188]; 2) les gestes utilisant le multitouch et plusieurs traits sont perçus comme étant plus difficiles [182]; 3) les techniques d'interactions proposées par VO et TB comportent de nombreux gestes différents, difficiles à apprendre et à mémoriser [186, 131].

En effet, les PDV préfèrent les gestes simples à faire et à retenir [188, 42], utilisant un seul doigt, une seule direction, sans angle ni courbe. Les raccourcis proposés par VO et TB, s'ils permettent de tronquer les trajectoires d'interaction, ajoutent de la complexité à la technique d'in-

4. Aide Accessibilité sur Android : <https://support.google.com/accessibility/android/answer/6151827?sjid=531405605717568832-EU>

5. Institute for Disability Research, Policy, and Practice (USA) : <https://webaim.org/projects/screenreadersurvey9/#mobile>

teraction. Ce problème est en partie explicable en raison de l'incompatibilité entre la modalité d'interaction et les tâches d'interaction [24]. D'une part, les interfaces graphiques et les écrans tactiles sont conçus pour être utilisés en 2D mais les instruments proposés pour naviguer dans le système sont en 1D (c'est-à-dire SWIPE horizontaux et verticaux) : le degré d'intégration est donc de 1/2, et les tâches à accomplir requièrent plus de manipulations. D'autre part, si certains gestes effectués sur l'écran ont une certaine similarité avec leur effet sur le système (par exemple, SWIPE horizontaux pour sélectionner les éléments suivants / précédents), la majorité d'entre eux sont plus absconds (par exemple, VO : toucher le haut de l'écran avec quatre doigts pour lire le premier élément, ou dessiner un "Z" pour fermer une notification), notamment en ce qui concerne le nombre de doigts à utiliser (par exemple, VO : faire défiler d'une page avec un SWIPE à 3 doigts) ou le nombre de répétition du geste à faire (eg., VO : toucher deux fois avec 3 doigts pour activer / désactiver le son de VO). Cela traduit un faible degré de compatibilité, ce qui ne respecte pas le deuxième principe du modèle de la manipulation directe [205, 207], et ne favorise pas la facilité de prise en main et l'apprentissage du système [186, 131].

De plus, ces modalités d'interaction ne s'affranchissent pas de leur séquentialité et ne peuvent pas être utilisées de façon concomitante afin de profiter de combinaisons comme avec un système multiplexé dans l'espace.

D'autres modalités d'interaction ont été proposées par la communauté IHM afin de rendre les écrans tactiles plus accessibles et utilisables par les PDV pour des tâches de sélection. En 2011, Krajnc et al. [116] proposent une interface de téléphone portable simplifiée limitant le nombre d'éléments à 4, occupant chacun un quart de l'écran. En 2022, Pooripanyakun et al. [173] proposent trois exemples d'interfaces sur téléphone portable (c'est-à-dire un menu, une interface de contrôle avec des flèches, une calculatrice) utilisables d'une seule main et sans la vue. Les éléments sont disposés en "U" le long des bords de l'écran, qui sont souvent utilisés comme repères par les PDV pour s'orienter sur le téléphone [109, 131].

En positionnant les éléments sur l'écran de sorte à être accessibles même en situation de non-retour visuel, ces modalités d'interaction utilisent une version restreinte du multiplexage

spatial et du principe de manipulation directe pour améliorer l'interaction. Les possibilités de disposition et le nombre d'éléments affichés en même temps sont restreints, mais c'est un compromis qui permet de sortir des modalités d'interaction séquentielles et linéaires utilisées par défaut. Dans ces exemples, le degré d'intégration est de 1 : l'écran est en 2D et permet d'interagir avec des éléments positionnés en 2D. Le degré d'indirection (c'est-à-dire faible décalage spatio-temporel) et le degré de compatibilité sont bas (c'est-à-dire peu de similarité entre un TAP et son effet sur ces systèmes, comme "Aller à l'écran d'accueil" ou "Entrer l'opérateur Diviser").



FIGURE 1.12 – Captures d'écrans des dispositifs de Krajnc et al. (2011) (à gauche) et de Poo-ripanyakun et al. (2022) (au milieu et à droite). Illustrations issues respectivement de [116] et [173].

D'autres modalités d'interaction proposent aux PDV de pouvoir effectuer des tâches d'orientation, de position et de trajectoire : par exemple, AudioDraw, présenté par Grussenmeyer et al. [84] en 2016, permet de sélectionner des formes géométriques et de les placer à l'endroit voulu sur l'écran d'une tablette (figure 1.13). Il est possible à tout moment de "lire" la composition en passant le doigt sur les formes pour les retrouver et les éditer (c'est-à-dire pour changer leur emplacement ou leur angle). Si cette modalité d'interaction permet aux PDV de composer des graphiques, elle est toutefois très linéaire. À l'instar des modalités d'interaction "exploration" et "pas à pas" vues précédemment, AudioDraw requiert de sélectionner puis de valider chaque étape du processus : par exemple, pour ajouter un carré, l'utilisateur doit appuyer sur le bou-

ton correspondant en bas de l'écran, puis déplacer le doigt jusqu'à l'endroit voulu, relâcher, effectuer un DOUBLE TAP pour valider l'emplacement, reposer le doigt sur le carré, valider la sélection par un DOUBLE TAP, effectuer la rotation en tournant autour de la forme, relâcher puis valider de nouveau à l'aide d'un DOUBLE TAP.

Cette modalité d'interaction a un bon niveau d'indirection, d'intégration et de compatibilité puisqu'il favorise la manipulation directe. Comme dans les deux travaux présentés précédemment [116, 173], AudioDraw utilise le multiplexage spatial à l'aide de 4 larges boutons situés en bas de l'écran. Malgré cela, l'étude ne précise pas le temps nécessaire à la production d'un diagramme mais il est raisonnable de penser que les trajectoires d'interaction seraient particulièrement longues. Par ailleurs, afin de réduire la complexité de la tâche, cette technique d'interaction limite certaines commandes à des entrées discrètes : par exemple, l'orientation des formes géométriques se fait par incrément de 45° seulement, même si le geste de rotation, lui, peut être effectué de façon continue. L'article mentionne également la nécessité d'ajouter de l'automatisation pour faciliter le placement des formes, notamment une fonction "d'assemblage" des formes (c'est-à-dire snap). Cela montre les limitations du paradigme Tactile quant aux tâches d'orientation, de position et de trajectoire en situation de non-retour visuel, difficiles à effectuer de manière continue. En effet, le paradigme Tactile seul est particulièrement inefficace pour les tâches de position, d'orientation et de trajectoire en 2D et plus, surtout lorsque l'utilisateur ne connaît pas à l'avance les données exactes (de position, d'orientation etc.) et doit les ajuster de façon continue. Il est nécessaire d'avoir un retour d'information simultané en temps réel afin de corriger l'entrée de façon continue jusqu'à obtenir la donnée désirée : Foley parle de "circuit ouvert/fermé", en s'inspirant des travaux de Ramsey et Atwood sur le positionnement continu et discret [181, 75]. Or l'utilisation du dispositif de captation (c'est-à-dire écran tactile lisse) n'offre presque aucun retour d'information lors de son utilisation [24] - contrairement à un potentiomètre cranté ou un interrupteur par exemple, qui offrent davantage d'informations haptiques et proprioceptives pendant l'interaction. L'interaction en entrée sur écran tactile n'est donc pas auto-suffisante pour satisfaire les conditions nécessaires à une entrée continue en circuit fermé : c'est au système de produire ce retour d'information supplémentaire en temps réel, qui ne soit pas visuel, et qui permette de manipuler les données voulues de façon intelligible.

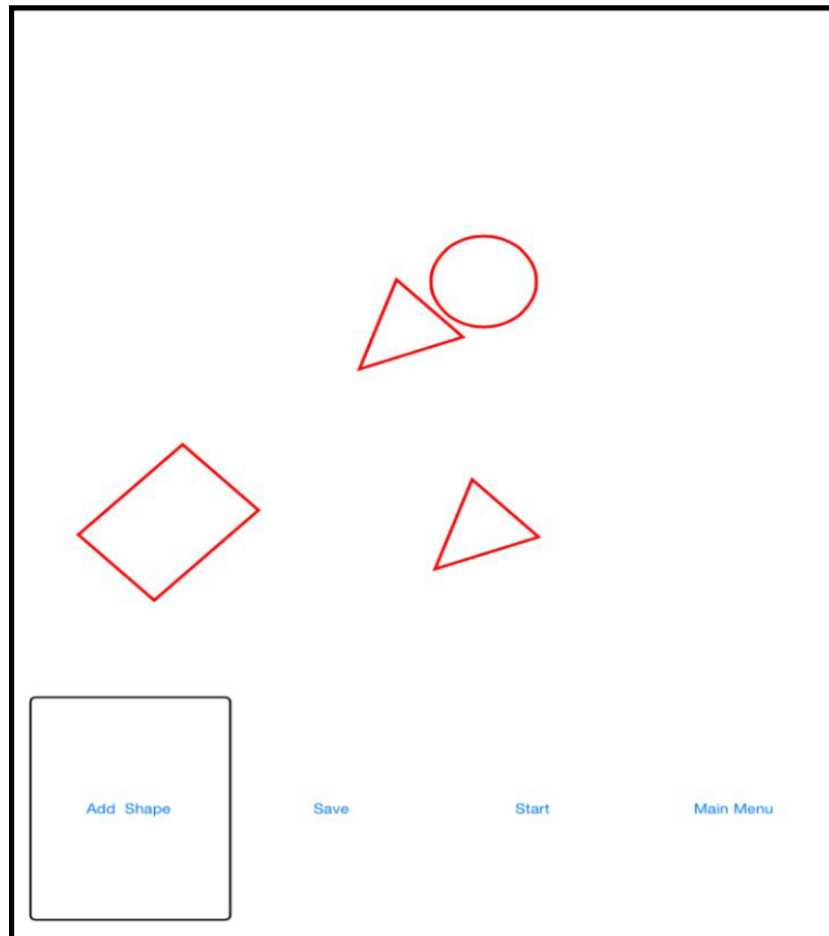


FIGURE 1.13 – Capture d'écran du dispositif AudioDraw de Grussenmeyer et al. (2016). Illustration issue de [84].

Concernant les tâches d'entrée de texte ou de nombres, elles sont fréquentes dans la littérature. Oliveira et al. [160] ont évalué la vitesse d'entrée de texte à l'aide du clavier QWERTY d'Apple sur tablette par 13 PDV et sont parvenus à une vitesse moyenne de 2,1 mots par minute (WPM, Word Per Minute). Les utilisateurs tenaient le téléphone portable avec leur main non directrice et utilisaient un doigt de l'autre main pour sélectionner la touche souhaitée en la touchant puis entrer la lettre en effectuant un double-TAP. Cette modalité d'interaction possède un excellent degré d'indirection, d'intégration et de compatibilité (puisque'il s'agit d'une métaphore directe d'un dispositif physique grand public), et si elle permet aux PDV de saisir du texte de ma-

nière similaire à une personne voyante grâce à une approche simple, elle requiert néanmoins un temps important d'acquisition de la cible car elle présente un grand nombre d'éléments de petite taille. À titre de comparaison, un utilisateur voyant sur téléphone portable peut parvenir à une vitesse moyenne de 50 WPM avec deux pouces, 36 WPM avec un seul doigt (par exemple, l'index) et 33 WPM avec un pouce [18].

Pour palier ce problème de ciblage, en 2010 Bonner et al. [33] présentent No-Look, une modalité d'interaction qui propose de découper l'écran en 8 zones (figure 1.14) - plus larges qu'une touche de clavier QWERTY. Pour entrer une touche, l'utilisateur sélectionne la zone contenant la lettre souhaitée en la touchant, puis valide la sélection en effectuant un TAP à l'aide d'un second doigt (c'est-à-dire split-TAP). Puisque les zones contiennent plusieurs lettres, une seconde étape de sélection permet de les discriminer : l'écran est alors verticalement divisé en 3 et l'utilisateur doit procéder de la même manière que précédemment pour entrer la lettre voulue. Ils parviennent à une vitesse moyenne de 1,3 WPM en une session d'une heure. Cette modalité a un bon degré d'indirection et d'intégration, mais possède un faible degré de compatibilité : il y a peu de similarité entre le geste effectué (et sa localisation sur l'écran, par exemple, un split-TAP sur la partie Nord) avec la réponse du système. En outre, ce type de modalité d'interaction fonctionne de la même manière que les claviers physiques 3*4 vu précédemment et souffre de problèmes similaires : basés sur un multiplexage temporel fort, ils requièrent plusieurs gestes pour entrer un seul caractère et allongent la trajectoire d'interaction.

Tinwala et al. [216] présentent en 2010 une modalité d'interaction pour entrer du texte basée sur le dessin. Pour saisir du texte, les utilisateurs tracent des traits sur la surface de l'écran à l'aide d'un doigt. À la fin d'un trait, l'application en analyse la forme pour identifier le caractère souhaité. Ils parviennent à une vitesse moyenne de 8 WPM dans leur mode sans assistance, mais il faut aux utilisateurs en moyenne 1,5 tentative pour entrer un caractère qui soit reconnu par le système. Finalement, il s'agit essentiellement une tâche de type Trajectoire, et cette modalité d'interaction, bien qu'ayant un haut degré d'intégration, d'indirection et de compatibilité, souffre des mêmes problèmes que le prototype AudioDraw [84]. Problèmes qui ont été exposés dans des travaux antérieurs [109, 186, 137, 188] : sur une surface tactile lisse, sans retour de la

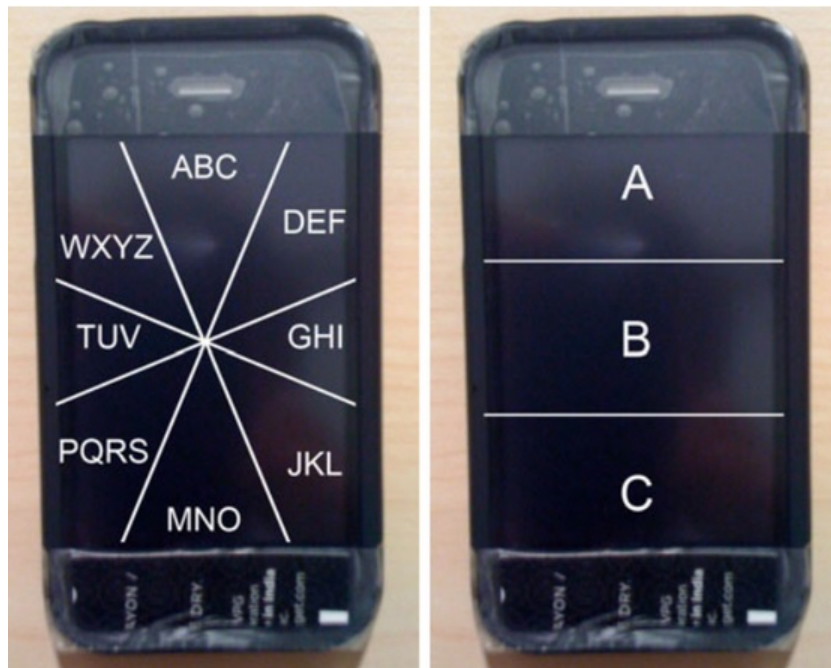


FIGURE 1.14 – Photographie du dispositif No-Look de Bonner et al. (2010). Photographie issue de [33].

part du système, orienter un geste et tracer un dessin de façon uniforme est difficile sans la vue.

D'autres auteurs ont basé leur modalités d'interaction sur des gestes simples, unidirectionnels et relatifs. Samanta et al. présentent en 2020 VectorEntry [193], une technique de saisie de texte utilisant des SWIPE dans les 8 directions cardinales pour sélectionner des caractères sur un clavier conçu selon le schéma traditionnel du clavier téléphonique 3*4. À l'instar de ce clavier et de No-Look, les touches contiennent plusieurs lettres et doivent être discriminées par une deuxième commande (figure 1.15). Ils parviennent à une vitesse moyenne de 3,3 WPM. En 2020, Ye et al. [247] présentent QB-Gest, une technique de saisie de texte bimanuelle utilisable sans retour visuel, où les utilisateurs déplacent leurs pouces dans une des 8 directions cardinales par rapport au point de contact (figure 1.16). Ils parviennent à une vitesse moyenne de 11,1 WPM après 4 sessions. Cependant, les touches du clavier contiennent plusieurs lettres qui sont discriminées automatiquement par le système. Ces modalités d'interaction ont un bon niveau d'indirection, d'intégration et de compatibilité. Toutefois, les orientations possibles sont

limitées afin de préserver la faisabilité, limitant aussi l'expressivité (c'est-à-dire le nombre de gestes utilisables). Il est alors nécessaire d'effectuer plusieurs commandes pour accéder à la lettre voulue ou d'utiliser une forme d'automatisation pour aider l'utilisateur.

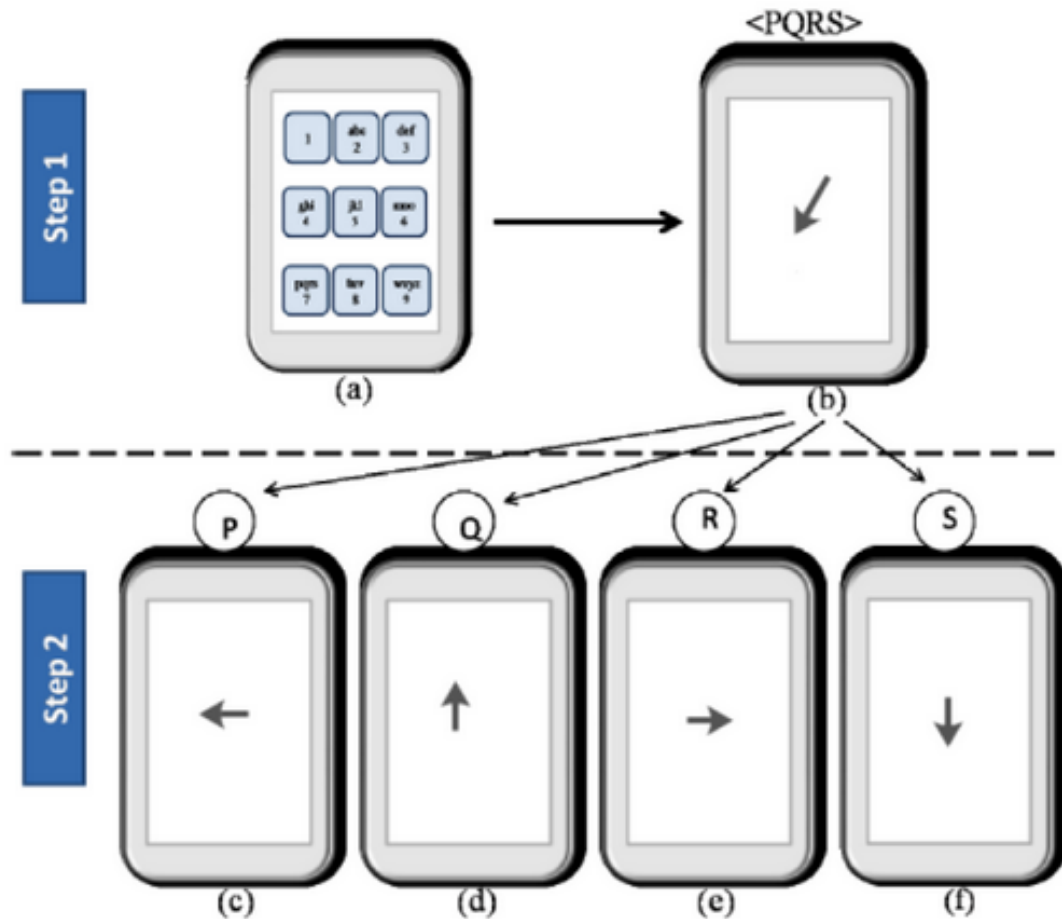


FIGURE 1.15 – Schéma du fonctionnement du dispositif VectorEntry de Samanta et al. (2020). Illustration issue de [193].

En résumé, les modalités présentées ont un niveau intermédiaire d'adéquation [24, 119] avec les tâches d'interaction, néanmoins le paradigme Tactile tel qu'utilisé sur téléphone portable est peu efficace et pose encore de nombreux défis aux PDV en raison d'une expressivité et bande-passante faible :



FIGURE 1.16 – Photographies du dispositif QB-Gest de Ye et al. (2020). Photographies issue de [247].

- le multiplexage spatial est possible mais simplifié, les études portant sur ce type d'interface basent leur conception sur la simplification (une dizaine d'éléments à l'écran au maximum, 5 étant optimal [115]) et la répartition des éléments à des endroits clés de l'écran (c'est-à-dire coins, bords) ;
- le vocabulaire d'entrée est limité à des gestes simples, peu nombreux, et possédant un faible degré de compatibilité avec les tâches d'interaction (par exemple, VectorEntry [193]).

Cette faible expressivité et bande-passante font que les trajectoires d'interaction sont allongées, ce qui en retour impacte négativement les charges de travail :

- motrice, car le faible nombre de gestes utilisables fait qu'il est nécessaire de les répéter plusieurs fois pour effectuer une tâche d'interaction ;
- cognitive, car la faible compatibilité entre les gestes et leur effets sur le système rend le vocabulaire d'entrée difficile à apprendre et à retenir (par exemple, VoiceOver et TalkBack [186, 131]).

La charge perceptive, quant à elle, est lésée par l'incapacité des modalités d'interaction étudiées à permettre un retour d'information instantané, continu (c'est-à-dire circuit fermé [75]) et auto-suffisant. Cela implique que le système doit générer un retour d'information supplémentaire pour que l'utilisateur sache ce avec quoi il interagit.

Cependant, le Tactile est un paradigme discret, socialement acceptable et ancré dans les usages. Les PDV souhaitent conserver ces caractéristiques et le téléphone portable comme

outil principal, donc plutôt que de le remplacer, nous cherchons un autre paradigme qui pourrait compléter le Tactile tout en conservant des caractéristiques désirables pour un PDV.

Tactile									
Articles	Modalité(s)	Fait pour / testé par PDV	Adéquation	Charges de travail	Pénibilité	Portabilité	Versatilité	Discretion et Confidentialité	Discrimination
8	< écran tactile , gestes >	-	Indirection ✓ Intégration ✓ Compatibilité ✗	Perception ✗ Cognition ✗ Motricité ✓	--	++	++	++ ++	--

FIGURE 1.17

Tactile									
Article	Modalité(s)	Fait pour / testé par PDV	Adéquation	Charges de travail	Pénibilité	Portabilité	Versatilité	Discretion et Confidentialité	Discrimination
Krajnc et al. (2011)	< écran tactile, tap>	Oui / Oui (n=7)	Indirection ✓ Intégration ✓ Compatibilité ✗	Perception ✓ Cognition ✓ Motricité ✓	--	++	++	++ ++	--
Pooripanyakun et al. (2022)	< écran tactile, tap>	Non / Non (n=11)	Indirection ✓ Intégration ✓ Compatibilité ✗	Perception ✗ Cognition ✓ Motricité ✓	--	++	++	++ ++	--
Grussenmeyer et al. (2016)	< écran tactile, grammaire de gestes >	Oui / Oui (n=8)	Indirection ✓ Intégration ✓ Compatibilité ✓	Perception ✗ Cognition ✗ Motricité ✓	--	++	+	++ ++	--
Oliveira et al. (2011)	< écran tactile, grammaire de gestes >	Non / Oui (n=13)	Indirection ✓ Intégration ✓ Compatibilité ✓	Perception ✗ Cognition ✓ Motricité ✗	--	++	+	++ ++	--
Bonner et al. (2010)	< écran tactile, grammaire de gestes >	Oui / Oui (n=10)	Indirection ✓ Intégration ✓ Compatibilité ✗	Perception ✓ Cognition ✓ Motricité ✓	--	++	+	++ ++	--
Tinwala et al. (2010)	< écran tactile, alphabet gestuel >	Non / Non (n=12)	Indirection ✓ Intégration ✓ Compatibilité ✓	Perception ✗ Cognition ✓ Motricité ✗	--	++	+	++ ++	--
Samanta et al. (2020)	< écran tactile, grammaire de gestes >	Oui / Oui (n=8)	Indirection ✗ Intégration ✓ Compatibilité ✗	Perception ✗ Cognition ✓ Motricité ✓	--	++	++	++ ++	--
Ye et al. (2020)	< écran tactile, grammaire de gestes >	Non / Non (n=10)	Indirection ✓ Intégration ✓ Compatibilité ✗	Perception ✗ Cognition ✓ Motricité ✓	--	++	++	++ ++	--

FIGURE 1.18

1.6.2 Le paradigme Vocal

Le paradigme Vocal implique l'utilisation du mode vocal [26] - qu'il s'agisse de langage naturel, pseudo-naturel ou de sons dépourvus de signification linguistique, dont la captation se fait à l'aide d'un seul microphone. Nous représentons ce paradigme comme utilisant les modalités d'interaction suivantes : <microphone, langage naturel>, <microphone, langage pseudo-naturel>, <microphone, grammaire de sons>.

L'espace de Card ne permet normalement pas de représenter les microphones. Mais les sons possèdent deux caractéristiques physiques : leur fréquence et leur amplitude, qui déterminent respectivement leur sonorité (c'est-à-dire grave / aiguë) et leur volume. Nous arguons que l'amplitude peut être représentée à l'aide d'une dimension dans la catégorie "Force" et que cette amplitude peut être virtuellement infinie (c'est-à-dire nombre réel). De la même façon, la fréquence peut être représentée à l'aide d'une dimension dans la catégorie "Delta Force", puisqu'il s'agit d'une mesure de variation de la Force sur une période déterminée (par exemple, une seconde). La fréquence aussi peut être virtuellement infinie, d'où la représentation proposée ci-après (figure 1.19).

Les interfaces vocales offrent un paradigme d'interaction basé sur l'audio et la voix. Au lieu de faire des gestes sur un écran tactile, les utilisateurs peuvent énoncer des commandes ou poser des questions à haute voix au système, qui interprète et répond en conséquence. Ce mode d'interaction non visuel offre aux PDV l'opportunité d'avoir un accès à des services ou interfaces peu accessibles - voire totalement inaccessibles - autrement.

En effet, alors que les utilisateurs voyants considèrent ces interfaces comme "gadget" [136, 177], ces technologies sont considérées comme des outils sérieux pour les PDV [17], sinon comme des outils de productivité essentiels [14]. C'est en particulier le cas des assistants vocaux (par exemple, Siri d'Apple, Alexa d'Amazon, Cortana de Microsoft) : des systèmes vocaux interactifs et multifonctions, généralement intégrés aux téléphones portables ou sur des enceintes connectées. Selon une étude de 2018 explorant les critiques de produits rédigées par

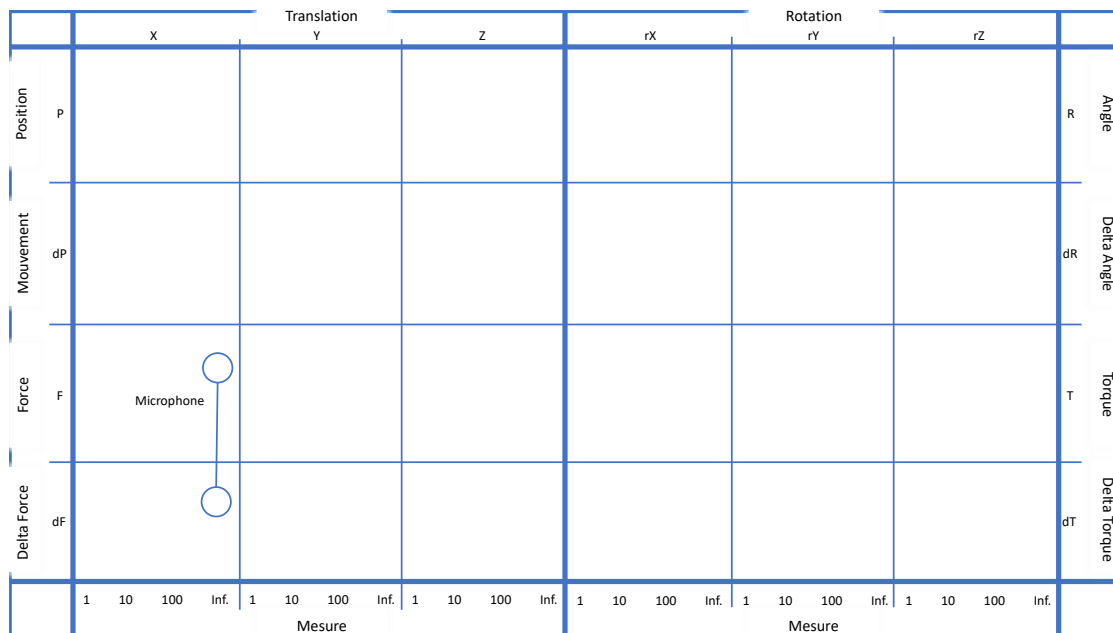


FIGURE 1.19 – Schéma représentant le dispositif "microphone" dans l'espace de conception des dispositifs en entrée de Card. Inspiré de [44].

des propriétaires d'Amazon Echo, Pradhan et al. ont montré qu'environ 38% de celles-ci mentionnent l'utilité de ces assistants vocaux pour les PDV [177]. Si les assistants vocaux, sont principalement utilisés pour des tâches telles que le contrôle d'un lecteur de musique, l'accès à la météo, le contrôle d'objets connectés, la rédaction de mails et la configuration de minuteries, alarmes et rappels [221, 134], ils rendent toutefois plus aisées ces tâches du quotidien et jouent un rôle important en renforçant le sentiment d'indépendance et d'autonomie chez les PDV qui les utilisent [4].

Cependant, ce paradigme d'interaction est intrinsèquement multiplexé dans le temps. Avec la voix il n'est pas possible de produire plusieurs unités d'information en même temps (c'est-à-dire il est impossible de parler plusieurs fois au même instant). La modalité vocale ne permet d'interagir avec les éléments du système que de façon séquentielle, limitant ainsi la bande-passante de la même manière que pour le paradigme Tactile (en situation de non-retour visuel). Certains systèmes néanmoins, notamment les assistants vocaux, peuvent être utilisés à l'aide du langage naturel. Or, utilisé comme langage d'interaction, celui-ci offre un niveau d'expressi-

tivité potentiellement illimité.

Ces particularités, propres à la modalité <microphone, langage naturel>, font que les tâches de Sélection sont grandement simplifiées. Dans leur étude de 2021, Abdolrahmani et al. [4] ont organisé des groupes de discussion composés de voyageurs aveugles et voyants pour explorer la possibilité d'étendre l'utilisation des assistants vocaux grand public à des interactions plus utilitaires dans un contexte de navigation en intérieur - notamment, dans les aéroports. Les résultats indiquent que ces assistants vocaux seraient essentiellement employés à obtenir des informations contextuelles en temps réel qu'il serait difficile d'obtenir en cherchant soi-même du fait de la complexité de la recherche (c'est-à-dire accessibilité des sources, masse de l'information etc.). Par exemple, demander oralement au système afin de connaître précisément les modalités de transport des bagages en soute (c'est-à-dire taille, poids etc.) plutôt que d'aller chercher dans la FAQ de la compagnie. Cette caractéristique permettant de spécifier directement l'effet voulu est au cœur de l'interaction vocale en langage naturel. En 2014, Zhong et al. ont présenté JustSpeak [253], un outil permettant un contrôle vocal du système d'exploitation Android, éliminant les tâches de localisation et de pointage de la cible (par exemple, localiser l'icône sur l'écran d'accueil pour lancer une application). Cette fonctionnalité a depuis été intégrée aux assistants vocaux nativement présents dans les appareils grand public (c'est-à-dire Siri, Google Now ...). Plus récemment, dans un article de 2021, Zhang et al. ont présenté VoiceEmoji [249], un système qui produit des emoji pertinents en réponse à des commandes vocales et qui permet aux PDV de les utiliser sans avoir à naviguer séquentiellement dans des listes d'emoji - réduisant ainsi le temps de saisie de 91% par rapport à la saisie utilisant le clavier tactile du téléphone.

Ces résultats montrent que le paradigme Vocal, surtout la modalité d'interaction utilisant le langage naturel, est extrêmement efficace pour les tâches de Sélection, puisqu'elle permet d'explicitement ce que l'utilisateur souhaite sans passer par une étape de recherche et de navigation dans le système pour trouver la commande. Le degré d'indirection spatiale est donc élevé (c'est-à-dire l'interface sur laquelle est faite la demande et celle où se trouve l'objet d'intérêt sont souvent éloignés) mais c'est tout l'intérêt de cette modalité d'interaction. En effet,

ces interactions vocales s'avèrent être plus rapides pour effectuer certaines tâches que d'opérer une succession de gestes et sont souvent préférées aux interactions sur écrans tactiles [4, 249]. Cela traduit un fort degré de compatibilité entre la modalité d'interaction et son effet sur le système, du fait que l'interaction s'effectue à l'aide d'un langage d'interaction considéré comme naturel pour l'utilisateur et sur un modèle transactionnel lui aussi très courant (c'est-à-dire demander pour obtenir quelque chose).

À l'inverse, tout comme pour le paradigme Tactile, le paradigme Vocal est particulièrement inefficace pour les tâches de position, d'orientation et de trajectoire, notamment lorsque l'utilisateur ne connaît pas à l'avance les données exactes (de position, d'orientation, etc.) et doit les ajuster de façon continue (c'est-à-dire circuit fermé [75]). Or l'utilisation du dispositif de captation (c'est-à-dire microphones) des modalités d'interaction vocale n'offre aucun retour d'information lors de son utilisation [24]. L'interaction en entrée du paradigme Vocal n'est donc pas auto-suffisante pour satisfaire les conditions nécessaires à une entrée continue en circuit fermé : c'est au système de produire ce retour d'information supplémentaire en temps réel, qui ne soit pas visuel, ni auditif car il ne serait pas confortable d'écouter et de parler en même temps, et qui permette de manipuler les données voulues de façon intelligible.

D'autre part, le langage naturel n'est pas pratique pour une utilisation continue : soit l'utilisateur doit parler tant qu'il souhaite que l'action dure, soit il doit donner l'ordre au système de poursuivre la commande jusqu'à ce qu'un nouvel ordre soit donné. La première solution n'est ni engageante ni efficace [96], et la seconde manque de précision et de maîtrise.

Le paradigme vocal peut toutefois être utilisé pour accomplir des tâches de position, d'orientation et de trajectoire de façon discrète. En 2016, Brewer et al. [35] ont réalisé une étude avec cinq PDV âgés de plus de 60 ans afin de comprendre comment concevoir un système qui leur permettrait d'éditer du contenu audio en utilisant une interface uniquement audio elle aussi. Leurs résultats montrent que les participants souhaitent utiliser des boutons physiques pour les tâches de navigation "génériques" (c'est-à-dire Lecture, Stop, Avance / Retour rapide) et des commandes vocales pour les tâches spécifiques (par exemple, "Avance d'une minute", "Va à

la section 4”). En 2019, Vtyurina et al. [228, 229], ont exploré les différences entre les lecteurs d'écran et les assistants vocaux dans le contexte de la navigation web. Ils ont élaboré “VERSE”, un prototype conçu pour améliorer les assistants vocaux en intégrant des fonctionnalités de navigation inspirées des lecteurs d'écran et permettant une exploration plus approfondie du web. Des tests d'utilisabilité auprès de 12 PDV ont révélé que l'utilisation de gestes serait plus désirables pour les commandes utilisées fréquemment, par exemple, pour la navigation dans les menus, ce qui permettrait d'éviter l'enchaînement de commandes vocales comme "suivant - suivant - suivant - sélectionner".

Ces travaux indiquent que, dans un contexte exploratoire (c'est-à-dire sans connaître exactement la cible), effectuer une multitude de commandes “logistiques” discrètes, simples, courtes et répétitives à haute voix pour naviguer dans un système (par exemple, un menu - tâche de position en 1D) n'est ni “naturel” ni efficace. Cela traduit un faible degré de compatibilité entre la modalité d'interaction et son effet sur le système, du fait que l'interaction s'effectue à l'aide d'un langage d'interaction très expressif pour opérer des commandes, elles, très simples. Le paradigme vocal est toutefois utilisable pour effectuer des tâches de position, d'orientation et de trajectoire de manière discrète lorsque les données de position sont connues de l'utilisateur. Dans un article de 2021, Gadiraju et al. [78] ont interrogé 59 PDV via un questionnaire en ligne sur la façon dont ils souhaiteraient piloter un drone. 71% des personnes interrogées ont indiqué qu'elles seraient intéressées par la possibilité de piloter (c'est-à-dire positionner, orienter, définir une trajectoire) un drone à l'aide de commandes vocales telles que : “Monte de 10 mètres” et “Avance de 3 mètres et dis-moi ce que tu vois”. Dans ce cas aussi la modalité d'interaction et son effet sur le système possèdent un fort degré de compatibilité. Par ailleurs, lorsque les données sont connues, elles se font souvent sous forme numérique - ce qui implique une tâche de saisie.

Or, de par sa nature, le paradigme Vocal est très pertinent pour les tâches de saisie. Si les précédentes études ont montré que le paradigme Vocal permet d'éviter les étapes de recherche et de navigation dans un système lorsque la cible est connue, cette caractéristique est encore plus prégnante pour le texte puisque taper du texte sur un clavier consiste, en somme, à naviguer dans un “menu” (c'est-à-dire le clavier) de façon séquentielle et répétée pour sélectionner

une multitude d'objets (c'est-à-dire les touches). Plusieurs recherches ont montré l'intérêt d'utiliser le vocal pour des tâches de création et d'édition de texte. Dans une tâche de saisie de texte sur un iPhone 6 Plus effectuée par des personnes voyantes, Ruan et al. [189] ont comparé en 2018 les performances entre un système de Speech-to-Text (STT) et une saisie au clavier tactile et ont constaté que le taux de saisie était approximativement 3 fois plus rapide avec un système STT qu'avec un clavier dans leur conditions expérimentales (153 WPM contre 52, en anglais). Dans une tâche similaire effectuée par des PDV, en 2013, Azenkot et al. [17] ont constaté que le taux de saisie était approximativement 5 fois plus rapide avec un système STT qu'avec un clavier (19,5 WPM contre 4,3). Utilisé de la sorte, le degré de compatibilité est très fort entre la modalité d'interaction et le résultat opéré sur la système, l'entrée et la sortie étant toutes deux en langage naturel.

Si la voix reste une modalité pertinente pour dicter du texte et envoyer des emails [217, 157, 138, 169, 92], elle semble insuffisante pour faire de l'édition plus complexe [80, 58]. Dans un article de 2018, Ghosh et al. [80] ont réalisé une étude auprès de 12 utilisateurs voyants pour étudier la viabilité d'une tâche d'édition de texte sans regarder et ont constaté que la tâche requérait une charge cognitive élevée et que l'usage d'une modalité supplémentaire utilisable sans la vue serait utile dans ce contexte, en particulier pour naviguer dans le système. En 2023, Darvishy et al. [58] ont proposé et évalué un système vocal permettant aux PDV de créer et éditer du texte sur téléphones portables. Leurs résultats indiquent que les commandes vocales ne sont pas adaptées pour se déplacer dans le texte avec un curseur (c'est-à-dire comme dans un éditeur de texte classique), mais plus appropriées pour de la sélection directe (par exemple, "Sélectionne le mot 'Lorem' dans la deuxième phrase"). Ils ont constaté que se fier uniquement aux commandes vocales n'était pas suffisant pour les PDV, et qu'une combinaison de gestes et d'interaction vocale serait plus robuste.

En synthèse, les modalités présentées ont un niveau élevé d'adéquation [24, 119] avec les tâches de sélection et de saisie, mais faible avec les autres, surtout lorsqu'il s'agit d'entrée continue et/ou lorsque l'objet d'intérêt n'est pas bien identifié ou identifiable. Cela est notamment dû au degré de compatibilité, qui change grandement selon la tâche d'interaction à effectuer. Les

deux autres caractéristiques sont peu influencées par la modalité d'interaction : le degré d'indirection temporel est faible, les retours d'information étant toujours concomitants à l'entrée et le degré d'indirection spatial est élevé, ce qui est désirable dans un mode vocal. Le degré d'intégration est quant à lui toujours de 1, les modalités présentées utilisent toutes le langage naturel, qui nécessite des variations de sonorité et de volume dans le temps. Nous n'avons pas trouvé d'exemple de travaux utilisant seulement une seule dimension (par exemple, le volume) en entrée et qui aurait donc un degré d'interaction de $\frac{1}{2}$. Ce type d'interaction aurait toutefois une granularité très forte pour être utilisable par le grand public et souffrirait donc d'un manque d'expressivité, en plus du manque inhérent de bande-passante.

Le paradigme Vocal, même s'il tire parti du fort pouvoir d'expression du langage naturel, est peu adapté à des tâches complexes et pose encore de nombreux défis en situation de non-retour visuel qui impactent négativement la charge cognitive :

- son fonctionnement exclusivement multiplexé dans le temps et son mode de communication mono-canal font qu'il s'agit d'un paradigme intrinsèquement séquentiel, perçu comme insuffisant, répétitif et pénible pour les tâches qui requièrent peu d'expressivité mais une forte bande-passante [80, 229, 58] ;
- par ailleurs, bien qu'elle permette dans certains cas d'éviter la navigation dans un système et d'aller à l'essentiel, il place sur l'utilisateur la charge d'identifier l'objet d'intérêt (dans le cas d'une sélection directe) et de fournir un effort cognitif important pour réfléchir à la formulation des commandes ou des requêtes [54, 50] - en particulier lorsqu'elles sont utilisées dans une langue autre que l'anglais pour des personnes non natives en anglais [30].
- ceci est d'autant plus vrai que l'interaction vocale est intangible et éphémère, contrairement à d'autres types d'interfaces (par exemple, graphiques, tangibles) possédant des repères (par exemple, visuels, spatiaux) permettant aux utilisateurs de se baser sur la reconnaissance (par exemple, icônes) plutôt que sur la mémorisation [153].

Dans ce paradigme, la charge motrice est faible, l'effecteur étant le système phonatoire.

La charge perceptive est fortement dépendante de la modalité d'interaction en sortie toutefois. Dans les travaux présentés, le retour d'information est essentiellement vocal et en langage naturel - ce qui implique une faible charge perceptive, dans les conditions acceptables d'utilisation.

De plus, le paradigme vocal peut être parfois à éviter selon l'emplacement ou la présence de tiers : son usage dans les lieux publics peut-être associé à un sentiment d'embarras [54, 204, 15] et requiert une certaine prudence en raison du risque de divulgation d'informations privées [66, 68]. Les environnements bruyants peuvent en outre fortement réduire les capacités du système à réagir correctement aux commandes vocales [60].

Vocal									
Articles	Modalité(s)	Fait pour / testé par PDV	Adéquation	Charges de travail	Pénibilité	Portabilité	Versatilité	Discrétion et Confidentialité	Discrimination
9	< microphone, langage naturel > < microphone, langage pseudo-naturel > < microphone, grammaire de sons >	-	Indirection ✗ Intégration ✓ Compatibilité ✓/✗	Perception ✓ Cognition ✗ Motricité ✓	--	++	++	--	+/-

FIGURE 1.20

Vocal									
Abdolrahmani et al. (2021)	< microphone, langage naturel >	Focus groupe (n=9)	Indirection ✗ Intégration ✓ Compatibilité ✓	Perception ✓ Cognition ✗ Motricité ✓	--	++	++	--	-
Zhong et al. (2014)	< microphone, langage naturel >	Oui / Non	Indirection ✗ Intégration ✓ Compatibilité ✓	Perception ✓ Cognition ✗ Motricité ✓	--	++	++	--	+/-
Zhang et al. (2021)	< microphone, langage naturel >	Oui / Oui (n=12)	Indirection ✗ Intégration ✓ Compatibilité ✓	Perception ✓ Cognition ✗ Motricité ✓	--	++	++	--	-
Brewer et al. (2016)	< microphone, langage naturel >	Oui / Oui (n=6)	Indirection ✗ Intégration ✓ Compatibilité ✗	Perception ✓ Cognition ✗ Motricité ✓	--	++	+/-	--	+/-
Vtyurina et al. (2019)	< microphone, langage naturel >	Oui / Oui (n=12)	Indirection ✗ Intégration ✓ Compatibilité ✗	Perception ✓ Cognition ✗ Motricité ✓	--	++	++	--	+/-
Gadiraju et al. (2021)	< microphone, langage naturel >	Questionnaire (n=59)	Indirection ✗ Intégration ✓ Compatibilité ✓	Perception ✓ Cognition ✗ Motricité ✓	--	++	++	--	-
Azenkot et al. (2013)	< microphone, langage naturel >	Oui / Oui (n=8)	Indirection ✗ Intégration ✓ Compatibilité ✓	Perception ✓ Cognition ✗ Motricité ✓	--	++	++	--	-
Ghosh et al. (2018)	< microphone, langage naturel >	Non / Non (n=12)	Indirection ✗ Intégration ✓ Compatibilité ✓	Perception ✓ Cognition ✗ Motricité ✓	--	++	+/-	--	+/-
Darvishy et al. (2023)	< microphone, langage naturel >	? / ? (n=?)	Indirection ✗ Intégration ✓ Compatibilité ✓	Perception ✓ Cognition ✗ Motricité ✓	--	++	+/-	--	+/-

FIGURE 1.21

1.6.3 Le paradigme Tangible

Le paradigme Tangible implique l'utilisation d'objets physiques que les utilisateurs peuvent manipuler pour interagir avec des systèmes interactifs. Selon cette définition, les dispositifs traditionnels tels que le clavier et la souris peuvent être considérés comme des interfaces tangibles (TUI, pour Tangible User Interface) mais la notion de TUI a évolué pour inclure un niveau d'interaction supplémentaire, allant au-delà de l'échange standard d'entrée et de sortie que proposent ces dispositifs. En effet, la tangibilité doit servir à la fois de support d'entrée et de sortie : la manipulation physique du système étant directement associée à son état [20] (par exemple, la position d'un interrupteur ou d'un potentiomètre).

La définition de ce paradigme est donc large et ne s'arrête pas à quelques modalités d'interaction. Celles-ci sont au contraire multiples, et il serait impossible de toutes les modéliser dans l'espace de Card, tant les dispositifs sont différents. Nous modélisons toutefois quelques-uns des travaux issus de la littérature (figure 1.22) que nous analysons dans cette section.

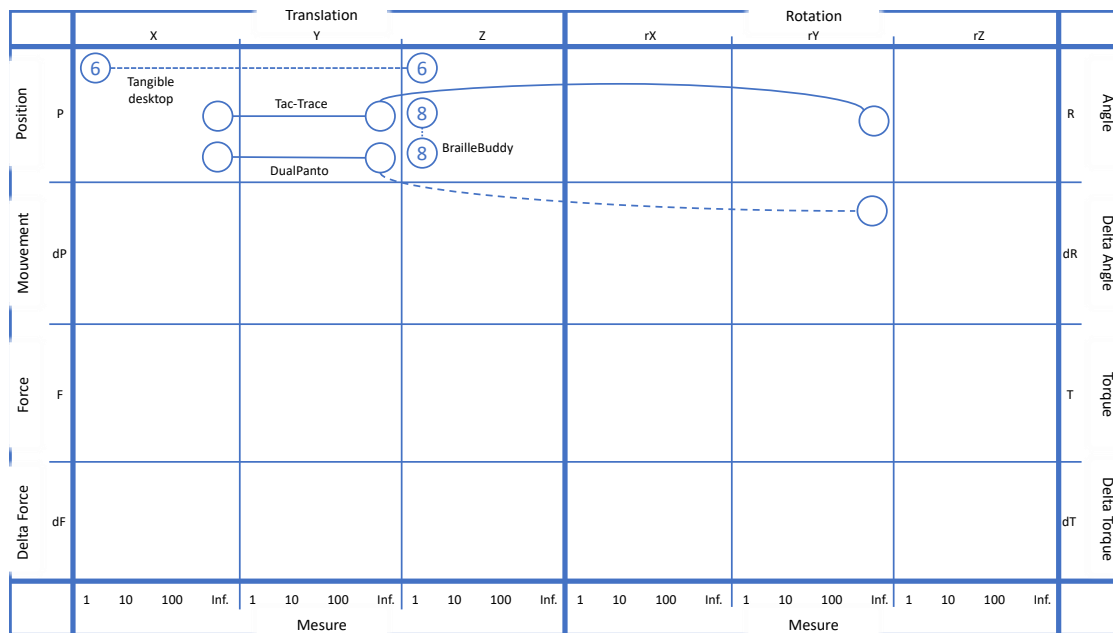


FIGURE 1.22 – Schéma représentant différents dispositifs relevant du paradigme Tangible dans l'espace de conception des dispositifs en entrée de Card. Inspiré de [44].

En introduisant de la tangibilité dans l'interface, ce paradigme offre aux PDV l'opportunité d'utiliser leur perception tactile et leur proprioception pour s'orienter et interagir avec le système [202, 36, 122]. Le toucher peut en effet - dans une certaine mesure - se substituer à la vision pour percevoir des informations graphiques telles que des cartes, des diagrammes ou des dessins [47].

"The Tangible Desktop" en est un bon exemple [19]. Présenté par Baldwin et al. en 2017 dans un article éponyme, il s'agit d'un dispositif à brancher sur ordinateur, composé d'icônes physiques ("phycones") à placer sur un plateau doté d'un potentiomètre rectiligne (figure 1.23). Les phycones représentent une entité numérique unique (par exemple, page web, fichier, dossier, logiciel) et sont différenciables à l'aide d'une zone sur le dessus possédant différentes formes et textures. Lorsqu'elles sont placées dans le plateau, le système ouvre l'entité en question. Le potentiomètre permet de passer le focus de l'une à l'autre des entités lorsque plusieurs phycones sont sur le plateau. Le système est doté d'un second potentiomètre rectiligne permettant de faire défiler les documents avec un positionnement absolu. Lors de tests utilisateurs avec 6 PDV, Baldwin et al. [19] ont obtenu une amélioration moyenne de 39% du temps d'exécution d'une tâche d'achat en ligne par rapport à l'utilisation d'un lecteur d'écran traditionnel. Ce système a un degré d'indirection spatial élevé, puisque la manipulation de l'interface se fait à une certaine distance de l'objet d'intérêt sans impact sur l'interaction. Le degré d'intégration est de 1, chaque icône étant reliée à son entité, et le potentiomètre (1D) permet de naviguer de l'une à l'autre de ces entités à la manière d'une liste (1D). Le degré de compatibilité est, quant à lui, élevé puisqu'il s'agit d'une représentation physique de la barre des tâches. Cette modalité d'interaction est très adaptée aux tâches de sélection et de position.

En 2018, Schneider et al. [197] présentent DualPanto, un dispositif à retour de force composé de deux pantographes dont les positions sont liées l'une à l'autre, permettant à des PDV de jouer à un jeu vidéo contenant des objets en mouvement (figure 1.24). Les utilisateurs interagissent avec DualPanto en manipulant la poignée "ME" avec une main et en tenant la poignée "IT" avec l'autre main. Ils guident leur avatar (c'est-à-dire mouvement et orientation) à travers un monde virtuel en utilisant la poignée "ME" tandis que la poignée "IT" se déplace d'elle-même,



FIGURE 1.23 – Photographies du dispositif Tangible Desktop de Baldwin et al. (2017), avec les "phycons" à gauche et le dispositif complet à droite. Illustration issue de [19].

permettant aux utilisateurs de sentir où se trouve la cible (par exemple, le ballon dans un jeu de football). Un système à retour de force bloque la poignée "ME" lorsque le personnage rencontre un obstacle en jeu (par exemple, mur, adversaire). Tout comme le système précédent, ce système a un degré d'indirection spatial élevé, sans que cela ne soit un problème. Le degré d'intégration est de 1, la poignée "ME" étant associée de manière bijective à l'avatar. De fait, le degré de compatibilité est très élevé puisque l'avatar agit de la même manière que la poignée "ME". Cette modalité d'interaction fonctionne bien pour les tâches de position, d'orientation et de trajectoire.

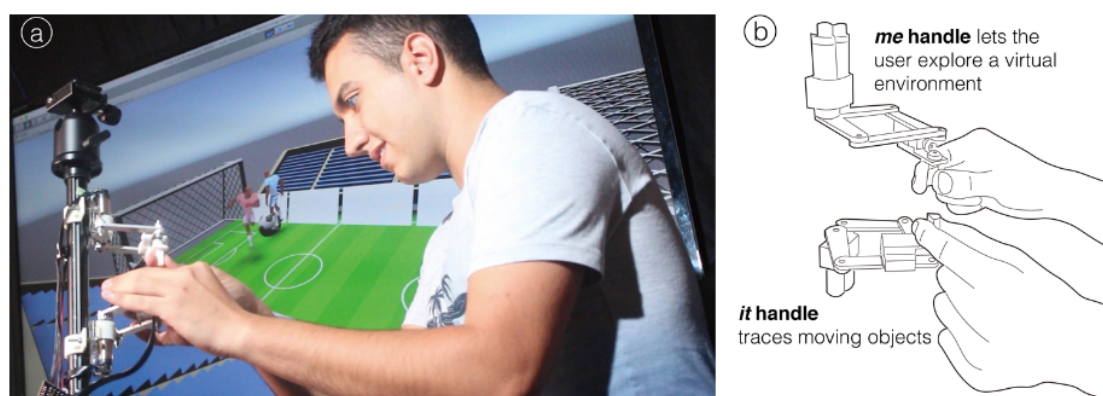


FIGURE 1.24 – Photographie et schéma illustrant le dispositif DualPanto de Schneider et al. (2018). Photographie issue de [197].

Bornschein et al. présentent un poste de travail pour PDV [34], utilisant une matrice Braille

rafraîchissable⁶ (figure 1.25) permettant, entre autres, de dessiner et d'éditer directement sur la zone braille à l'aide des doigts ou d'un stylet. Elle fonctionne comme une tablette tactile - la différence étant que l'écran affiche les informations en relief afin d'être perçue de façon haptique. Le degré d'indirection spatiale est plus faible, puisque l'interaction se fait au niveau de l'objet d'intérêt. Du reste, le degré d'intégration est de 1 et le degré de compatibilité avec la tâche de dessin (c'est-à-dire trajectoire) est élevé.

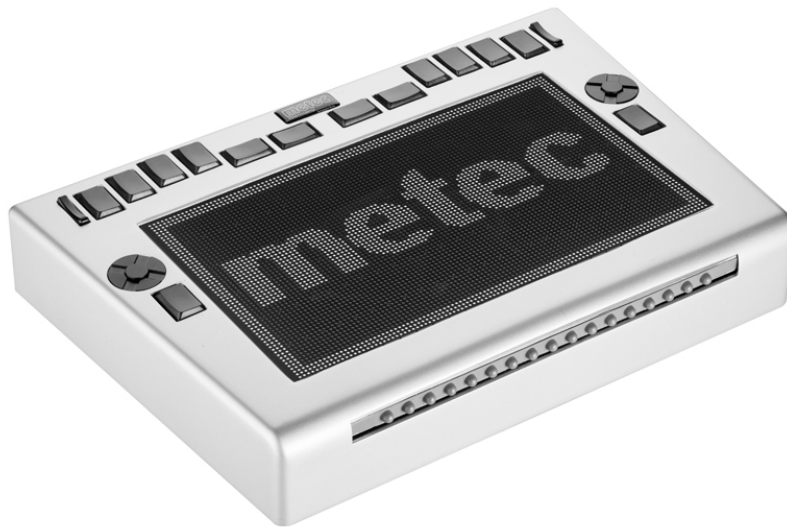


FIGURE 1.25 – Matrice braille rafraîchissable et tactile : l'HyperBraille F de Metec. Photographie issue du Site web de Metec.

Les TUI sont des interfaces qui illustrent bien le principe de la manipulation directe des interfaces graphiques, et ces trois exemples mettent particulièrement en lumière ces caractéristiques : leur manipulation est intrinsèquement physique ; de plus en terme de modalités de sortie, leur tangibilité leur donne une permanence similaire à celle d'une interface graphique, les objets d'intérêts étant continuellement représentés - contrairement aux interfaces vocales en sortie dont le support est éphémère et qui requiert soit de garder en mémoire l'état du système, soit d'en demander un rappel.

6. Metec AG. 2016. Hyperbraille S Display 6240. techreport. Metec AG, Stuttgart, Germany.

En outre, leur tangibilité permet aussi un multiplexage spatial, leur offrant une plus grande bande passante : les TUI peuvent paralléliser les entrées (par exemple, appuyer sur plusieurs boutons à la fois), et peuvent bénéficier d'un effet combinatoire qui démultiplie le nombre d'actions possibles en une seule commande. Par ailleurs, selon la modalité d'interaction, les TUI peuvent être utilisées pour des tâches d'orientation, position et trajectoire de façon continue (c'est-à-dire circuit fermé [75]) car les dispositifs physiques utilisés sont capables de fournir un retour d'information haptique et proprioceptif suffisant et de façon simultanée à l'entrée.

Ces caractéristiques font que les TUI peuvent en quelque sorte "imiter" les interfaces graphiques, ce qui les rend très attrayantes pour les PDV. Cependant ces dispositifs sont peu pratiques à utiliser au quotidien : ils sont encombrants, complexes à installer et à utiliser, et peuvent être chers (par exemple, une matrice Braille rafraîchissable comme l'HyperBraille F de Metec coûte environs 55 000\$ [161]). Ils sont en outre limités à des usages très spécifiques de par leur format, les dispositifs étant ad hoc, et ne peuvent souvent pas être modifiés aisément.

Toutefois les TUI ne s'arrêtent pas à ces dispositifs : la démocratisation des technologies telles que la micro-électronique et l'impression 3D ouvre la porte à la création d'objets physiques pour les TUI de façon rapide et à faible coût [95]. Au-delà des dispositifs déjà présentés, il existe des systèmes plus simples, basant leur interaction sur la position et la combinaison de petits éléments tangibles. Cette modularité permet d'élargir les possibilités de l'interface d'une façon qui n'est pas sans rappeler la liberté offerte par les jeux de construction (par exemple, Lego™).

- En 2010, McGookin et al. [143] présentent Tangible Graph Builder, une solution conçue pour permettre aux PDV de produire et lire des graphes à l'aide de marqueurs 3D à placer sur une grille tangible.
- En 2016, Brulé et al. [38] présentent MapSense (figure 1.26), une carte interactive multi-sensorielle pour les enfants PDV utilisant des objets 3D et une surface texturée.
- En 2022, Jafri et al. [99] présentent Tac-Trace (figure 1.26), une solution basée sur des objets imprimés en 3D dotés de marqueurs fiduciaux et suivis par un système de vision

par ordinateur pour aider les enfants PDV à apprendre les formes géométriques.



FIGURE 1.26 – **Gauche** : Photographies du dispositif MapSense de Brulé et al. (2016). **Droite** : Photographie du dispositif Tac-Trace de Jafri et al. (2022). Illustrations issues respectivement de [38] et [99].

Dans ces travaux, la manipulation de ces éléments tangibles est très adaptée aux tâches de sélection, position et orientation. Le degré d'indirection spatiale est négligeable, puisque l'objet d'intérêt est manipulé directement. Le degré d'intégration est de 1 et le degré de compatibilité avec les tâches respectives est très élevé puisque les dispositifs sont conçus de façon ad hoc.

Les TUI disposent d'une autre caractéristique intéressante : leur utilisation peut être grandement facilitée par leur affordance (et les contraintes associées). Une définition de l'affordance par Norman [158] est : "la caractéristique d'un objet ou d'un environnement qui suggère à son utilisateur son mode d'usage ou autre pratique"⁷. Concrètement, il s'agit de concevoir les TUI de sorte que leurs caractéristiques (par exemple, forme, poids, matière, texture, etc.) guident l'utilisateur dans la manière d'interagir (par exemple, une poignée peut être tirée, tenue ou tournée). À l'inverse, les contraintes empêchent, ou entravent physiquement, certaines actions d'être effectuées (par exemple, insérer un objet cubique dans un orifice rond de moindre taille). Les contraintes et les affordances peuvent ainsi fonctionner en tandem de sorte que l'interface

7. Affordance - Wikipedia : <https://fr.wikipedia.org/wiki/Affordance>

soit utilisable intuitivement - rendant les utilisateurs plus autonomes.

Par exemple, en 2018, Marichal et al. présentent LETSmath [139], un système tangible d'apprentissage des mathématiques pour enfants PDV. Il se compose de blocs tangibles, dont la taille est proportionnelle au chiffre qu'ils représentent, qui peuvent être additionnés : lorsqu'ils sont posés ensemble sur une surface de travail spécifique, cela déclenche un retour d'information auditif et vibrotactile indiquant le nombre total de cubes présents. Ici l'une des affordance est basée sur la taille des blocs, qui représente leur valeur numérique dans le système.



FIGURE 1.27 – Schéma illustrant le dispositif LETSmath de Marichal et al. (2018). Illustration issue de [139].

Lang et al., en 2023, présentent BrailleBuddy [122], un système d'aide à l'apprentissage du Braille en autonomie par des enfants DV. Le système est composé d'une boîte dotée d'un emplacement de grande taille dans lequel peut être inséré une plaque surmontée d'une figurine représentant un animal ou un objet, et de sept emplacements plus petits destinés à recevoir des plaques surmontées d'une lettre en braille. Le but est d'utiliser les petites plaques pour écrire le nom de l'objet ou de l'animal présent sur la grande plaque. Ici l'une des affordance est basée sur une contrainte : la taille des plaques, qui ne peuvent être insérées que dans les emplacements appropriés.

Ces deux modalités d'interaction respectent aussi les critères de la manipulation directe de

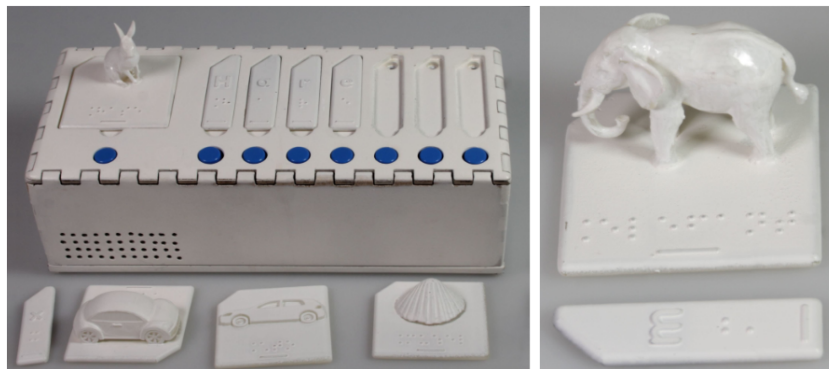


FIGURE 1.28 – Photographies du dispositif BrailleBuddy de Lang et al. (2023). Illustration issue de [122].

l'objet et permettent aisément d'effectuer des tâches de sélection et de position. Elles sont aussi utilisées pour des tâches de saisie : de nombre pour LetsMath, de texte pour BrailleBuddy. Ces dispositifs sont toutefois grandement limités : le premier ne possède que des blocs allant du chiffre 1 à 5, le second ne possède que 7 emplacements pour y placer des lettres. Par ailleurs, trouver le bon élément parmi tous et le placer au bon endroit est une opération longue et coûteuse en charge de travail perceptive et motrice. Ce n'est pas un souci cependant, car l'objectif de ces systèmes est pédagogique et non performatif (c'est-à-dire saisir du texte le plus rapidement possible). Ces exemples illustrent toutefois les limites des TUI basés sur la manipulation directe de petits éléments dans les tâches de saisie.

Une fois encore, le degré d'indirection spatiale est négligeable, puisque l'objet d'intérêt est manipulé directement, le degré d'intégration est de 1 et le degré de compatibilité avec leurs tâches respectives est très élevé puisque les dispositifs sont conçus de façon ad hoc.

En synthèse, les modalités présentées ont toutes un niveau élevé d'adéquation [24, 119] avec les tâches de sélection, de position, d'orientation et de trajectoire, du fait de leur haut degré de manipulabilité directe. L'affordance que leur confère leur tangibilité permet un usage plus instinctif et donc impacte positivement sur la charge perceptive et cognitive. Cependant, l'exploration tactile des TUI affecte davantage la charge cognitive par rapport à l'exploration visuelle [88, 192, 224, 133] : étant donné que le champ perceptif tactile est très limité [132], il

est nécessaire de se déplacer pour explorer l'ensemble de l'objet, et l'intégration doit se faire en agrégeant une succession de détails incomplets jusqu'à ce que suffisamment d'informations aient été acquises pour synthétiser et comprendre l'ensemble. La manipulation d'objets physiques peut aussi poser davantage de problèmes de fatigue à l'utilisateur [202] que l'utilisation d'une souris ou d'un écran tactile, puisque l'effecteur est plus grand.

De plus, malgré leurs caractéristiques intéressantes, les TUI font aussi face à des défis inhérents à leur physicalité. Si certains dispositifs sont plus petits que d'autres, il s'agit pour la plupart de systèmes stationnaires et encombrants [183, 108] et ils sont principalement utilisés dans des environnements où les utilisateurs n'ont pas besoin de les transporter, tels que les salles de classe. Même pour les dispositifs plus compacts, ils sont souvent dotés de petits éléments susceptibles d'être facilement perdus, notamment par des PDV. Edge et Black décrivent cela comme un problème de voluminosité ("bulkiness") [67]. De plus, cette voluminosité physique peut rendre la manipulation de plusieurs représentations (par exemple, une carte) en même temps difficile, nécessitant un espace suffisant pour les disposer. En comparaison, les représentations numériques peuvent être enregistrées et rouvertes sans occuper d'espace physique supplémentaire.

Par ailleurs, une fois produits, ces éléments physiques sont statiques. Ils ne peuvent pas s'adapter facilement à la diversité des applications, rendant ainsi complexes les mises à jour dynamiques du contenu. Par exemple, avec une carte numérique, il est possible de zoomer et de dézoomer pour ajuster l'échelle. En revanche, avec un modèle physique, la taille de chaque élément est fixe : ajuster l'échelle nécessiterait de modifier l'ensemble du modèle en le reconstruisant. Poupyrev et al. [176] soulignent que les TUI sont généralement conçues pour faciliter un ensemble limité de tâches spécifiques : contrairement aux interfaces graphiques, plus polyvalentes, les objets physiques ne peuvent pas être transformés en différents objets ou modifier leurs propriétés physiques, même si de tels changements seraient appropriés.

1.6. ANALYSE DES PARADIGMES ET MODALITÉS D'INTERACTION POUR LES PERSONNES AYANT UNE DÉFICIENCE VISUELLE

Tangible									
Articles	Modalité(s)	Fait pour / testé par PDV	Adéquation	Charges de travail	Pénibilité	Portabilité	Versatilité	Discretion et Confidentialité	Discrimination
8	Ad hoc	-	Indirection ✓ Intégration ✓ Compatibilité ✓	Perception ✓ Cognition ✓ Motricité ✓	-	--	-	+ +	Selon le contexte et le dispositif

FIGURE 1.29

Tangible									
Article	Modalité(s)	Fait pour / testé par PDV	Adéquation	Charges de travail	Pénibilité	Portabilité	Versatilité	Discretion et Confidentialité	Discrimination
Baldwin et al. (2017)	< interface tangible, manipulation d'objets >	Oui / Oui (n=6)	Indirection ✓ Intégration ✓ Compatibilité ✓	Perception ✓ Cognition ✓ Motricité ✓	+/-	-	+/-	+ +	+/-
Schneider et al. (2018)	< pantographes, mouvements 2D >	Oui / Oui (n=6)	Indirection ✓ Intégration ✓ Compatibilité ✓	Perception ✓ Cognition ✓ Motricité ✓	+/-	--	+/-	+ +	+/-
Bornschein et al. (2018)	< plage braille, dessin >	Oui / Oui (n=12)	Indirection ✓ Intégration ✓ Compatibilité ✓	Perception ✓ Cognition ✓ Motricité ✓	-	--	+	+ +	+/-
McGookin et al. (2010)	< marqueurs fiduciaires, manipulation d'objets >	Oui / Oui (n=12)	Indirection ✓ Intégration ✓ Compatibilité ✓	Perception ✓ Cognition ✓ Motricité ✓	-	--	-	+ +	+/-
Brulé et al. (2016)	< écran tactile, gestes et manipulation d'objets >	Oui / Oui (n=10)	Indirection ✓ Intégration ✓ Compatibilité ✓	Perception ✓ Cognition ✓ Motricité ✓	-	-	-	+ +	+/-
Jafri et al. (2022)	< marqueurs fiduciaires, manipulation d'objets >	Oui / Oui (n=12)	Indirection ✓ Intégration ✓ Compatibilité ✓	Perception ✓ Cognition ✓ Motricité ✓	-	--	--	+ +	+/-
Marichal et al. (2018)	< blocs connectés, manipulation d'objets >	Oui / ? (n=?)	Indirection ✓ Intégration ✓ Compatibilité ✓	Perception ✓ Cognition ✓ Motricité ✓	-	-	-	+ +	+/-
Lang et al. (2023)	< interface tangible, manipulation d'objets >	Oui / Oui (n=?)	Indirection ✓ Intégration ✓ Compatibilité ✓	Perception ✓ Cognition ✓ Motricité ✓	-	-	-	+ +	+/-

FIGURE 1.30

1.6.4 Le paradigme Gestuel

Le paradigme Gestuel implique l'utilisation du mode Gestuel [26]. Il s'agit des modalités d'interaction utilisant comme langage d'interaction les mouvements et postures du corps, indépendamment du dispositif d'entrée. Plus précisément, et ce pour mieux distinguer ce paradigme des paradigmes Tactile et Tangible, nous nous focalisons sur les mouvements du corps 1) dont l'effecteur est plus grand que les doigts et la main seuls, et 2) effectués en l'air et/ou sur le corps de l'utilisateur. Nous représentons ce paradigme comme utilisant les modalités d'interaction suivantes : <[dispositif], grammaire de gestes en l'air>, <[dispositif], grammaire de gestes sur le corps>.

La définition de ce paradigme est donc large, et comprend une grande variété de modalités d'interaction. La captation des gestes est variée et nous modélisons quelques dispositifs de captation dans l'espace de Card (figure 1.31).

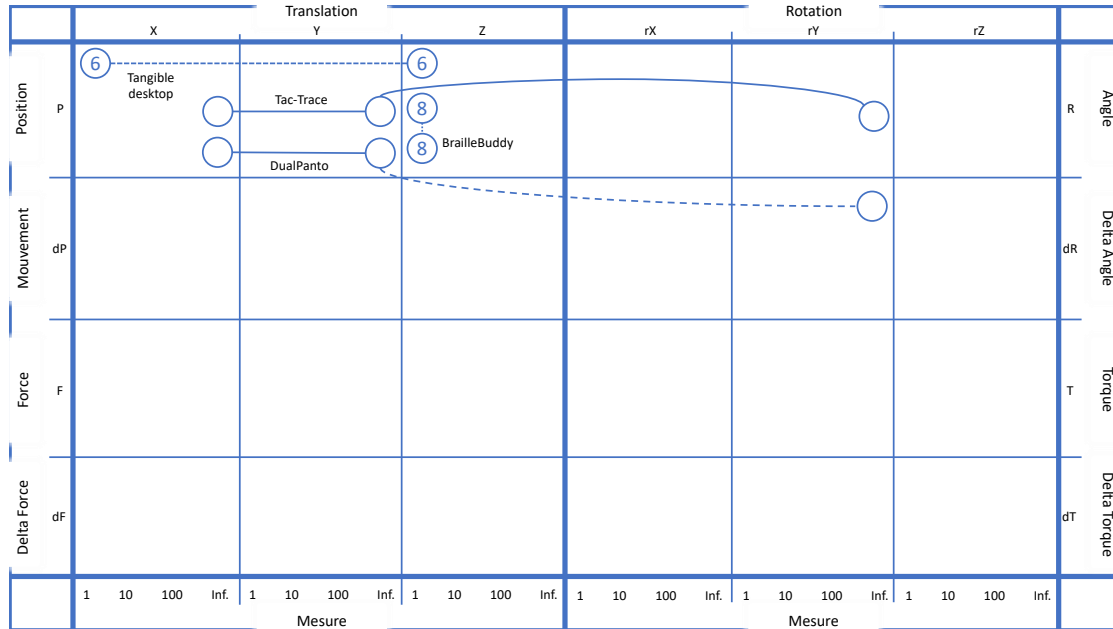


FIGURE 1.31 – Schéma représentant différents dispositifs relevant du paradigme Gestuel dans l'espace de conception des dispositifs en entrée de Card. Inspiré de [44].

À l'instar du paradigme Tangible, le paradigme Gestuel offre aux PDV la possibilité d'utiliser leur proprioception pour interagir avec le système [91]. En effet, dans un article de 1993, Hocherman [91] a montré qu'après entraînement, il était possible d'effectuer des mouvements de précision de manière proprioceptive vers des emplacements cibles dans un plan horizontal. Si la perception proprioceptive n'est pas suffisante pour se substituer à la vision - comme pour les interfaces Tangibles - elle est suffisante toutefois pour s'orienter et permettre un usage d'interfaces multiplexées dans l'espace.

En 2009, Li et al. [128] présentent VirtualShelves, une modalité d'interaction permettant d'activer des objets d'intérêts en orientant le bras dans un espace hémisphérique (c'est-à-dire 180°/180°) segmenté en régions situées devant l'utilisateur. Les auteurs montrent que les utili-

sateurs peuvent pointer avec précision 7 régions sur le plan horizontal et 4 régions sur le plan vertical en utilisant uniquement leur mémoire kinesthésique. Costa et al. [53] ont proposé en 2020 une modalité d'interaction permettant de contrôler une TV à l'aide de commandes de navigation horizontales et verticales, en déplaçant l'avant-bras dans la direction souhaitée. Un geste de pincement permet de valider la sélection. Dans ces travaux, le degré d'indirection spatiale est élevé, car l'objet d'intérêt - voire le dispositif - est situé à distance de l'utilisateur. Le degré d'intégration dépend du dispositif et n'est pas pertinent pour ces travaux : les gestes effectués en l'air sont intrinsèquement en 3D mais le système ne capture et filtre que ce dont il a besoin pour permettre l'interaction, par exemple, la position de la main sur le plan horizontal et vertical dans les travaux de Costa et al. [53]. En revanche, le degré de compatibilité est élevé : les gestes effectués sont très similaires avec leur effets sur le système car il s'agit de gestes de pointage (Déictiques), naturels et particulièrement appropriés lorsqu'il s'agit de sélectionner quelque chose parmi un ensemble (c'est-à-dire tâches de sélection).



FIGURE 1.32 – **Gauche** : Photographies illustrant le fonctionnement de VirtualShelves de Li et al. (2009). **Droite** : Schéma illustrant le fonctionnement du dispositif de Costa et al. (2020). Illustrations issues respectivement de [128] et [53].

Toutefois, comme le montrent ces travaux, utiliser la proprioception seule dans une tâche de sélection par pointage direct limite la précision et donc la granularité de l'interaction ainsi que la densité d'information : Hocherman [91] n'a utilisé que 5 cibles dans son étude tandis que Li et al. [128] ont placé la limite à 7 régions sur le plan horizontal et 4 sur le plan vertical. Par ailleurs, les gestes de pointage souffrent des tremblements de main, ce qui fait qu'il est difficile de sélectionner de petites cibles, même avec un retour visuel [227].

Selon la modalité d'interaction, le paradigme Gestuel peut être utilisé pour des tâches d'orientation, position et trajectoire de façon continue (c'est-à-dire circuit fermé [75]). Dicke et al. [61] ont proposé en 2015 une modalité d'interaction permettant de choisir une musique parmi 20 à l'aide d'un geste de pointage effectué sur un arc de cercle de 110° orienté sur le plan horizontal. En 2016, Dim et al. [114] proposent deux études d'élicitation par des PDV de gestes en l'air destinés à contrôler une TV. Ils ont trouvé plusieurs gestes permettant de changer le volume ou les chaînes de façon continue : pour les commandes de volume, les gestes consistent à déplacer la main dominante vers le haut/le bas ; pour les commandes de chaînes, les gestes consistent à tourner la main de gauche à droite, imitant ainsi la rotation de la molette des canaux sur les anciennes télévisions.

Cependant, les modalités d'interactions utilisant des gestes en l'air sont peu adaptées pour les tâches de position, d'orientation et de trajectoire lorsque l'utilisateur ne connaît pas à l'avance les données exactes et doit les ajuster de façon continue car effectuer un geste en l'air n'offre presque aucun retour d'information sinon celui de la proprioception. Or nous avons vu pour les tâches de sélection que cette proprioception seule n'offre pas une granularité assez fine pour permettre un contrôle précis. Poupyrev et al. [175] ont proposé en 2002 une méthode d'interaction permettant de faire défiler une liste (1D) plus ou moins rapidement en fonction du degré d'inclinaison d'un dispositif portable. Cette étude a été réalisée en situation de retour visuel, et pourtant, ils ont constaté que le positionnement précis, par exemple pour faire défiler d'une seule ligne, était difficile. Une raison possible est que l'angle seuil était trop petit et la vitesse de défilement initiale était trop rapide, ce qui faisait que les sujets ne pouvaient pas s'arrêter suffisamment rapidement lorsque la liste commençait à défiler. Ni et al. [152] ont proposé rapMenu en 2011, une modalité d'interaction permettant de naviguer dans un menu radial contenant 12 objets d'intérêts à l'aide de la rotation de la main (180°) en situation de non-retour visuel. Si cette rotation est continue, les seuils sont toutefois très larges, car seules trois régions sont accessibles de la sorte dans le menu radial : le sélecteur positionné au nord, à l'est ou au sud (figure 1.33). Un TAP effectué avec le pouce sur l'un des quatre autres doigts permet de sélectionner l'objet d'intérêt souhaité dans cette région. Il s'agit d'une adaptation de la modalité d'interaction de Rahman et al. [180] en 2009, qui proposent une navigation (en situation de re-

tour visuel) dans un menu radial à 12 ou 16 objets en utilisant uniquement la rotation de la main - et ce sur un angle de 125° seulement. Les seuils sont donc beaucoup plus petits, et si les utilisateurs parviennent à utiliser cette modalité en situation de retour visuel, les auteurs Ni et al. [152] affirment qu'il est peu probable qu'ils puissent faire de même sans retour d'information visuelle. Dans leurs travaux, Dicke et al. [61] ont considéré jusqu'à 20 objets mais l'interface utilise un retour d'information auditif immédiat.



FIGURE 1.33 – Illustrations représentant le fonctionnement du dispositif rapMenu de Ni et al. (2011). Illustrations issues de [152].

Ces études [180, 152, 61] montrent que les modalités d'interaction utilisant les gestes en l'air ne sont pas auto-suffisantes pour satisfaire les conditions nécessaires à une entrée continue en circuit fermé [75] : c'est au système de produire ce retour d'information supplémentaire en temps réel, qui ne soit pas visuel, et qui permette de manipuler les données voulues de façon intelligible. Dans ces travaux, le degré d'indirection spatiale est élevé comme pour toutes les modalités d'interaction gestuelle en l'air. Là encore, le degré d'intégration dépend du dispositif et n'est pas pertinent pour ces travaux. En revanche, le degré de compatibilité est élevé car les gestes effectués sont très similaires avec leur effets sur le système : les langages des modalités gestuelles sont toutes basées sur une métaphore ayant une affordance simple. La modalité gestuelle dans le système de Poupyrev et al. [175] utilise une affordance basée sur la gravité : pencher le dispositif d'un côté ou de l'autre fait "tomber" l'objet d'intérêt plus ou moins vite dans la direction idoine. La modalité gestuelle de Rahman et al. et de Ni et al. [180, 152] utilise la métaphore du potentiomètre rotatif. La modalité du système de Dicke et al. [61] implémente des gestes de pointage.

Le paradigme Gestuel peut aussi être utilisé pour des tâches de saisie de texte. Kajastila et al. [105] proposent en 2009 une modalité d'interaction en situation de non-retour visuel utilisant l'orientation en 3D d'un dispositif pour naviguer dans un menu sphérique (figure 1.34). Les auteurs ne précisent pas la performance moyenne en WPM mais seulement le temps moyen de sélection d'une lettre, qui est de 3.82 secondes. Avec cette donnée, nous calculons naïvement un score de 15,7 WPM, ce qui est plus de trois fois supérieur aux résultats obtenus par Azenkot et al. [17] avec des PDV utilisant le clavier tactile d'un téléphone portable (4,3 WPM).

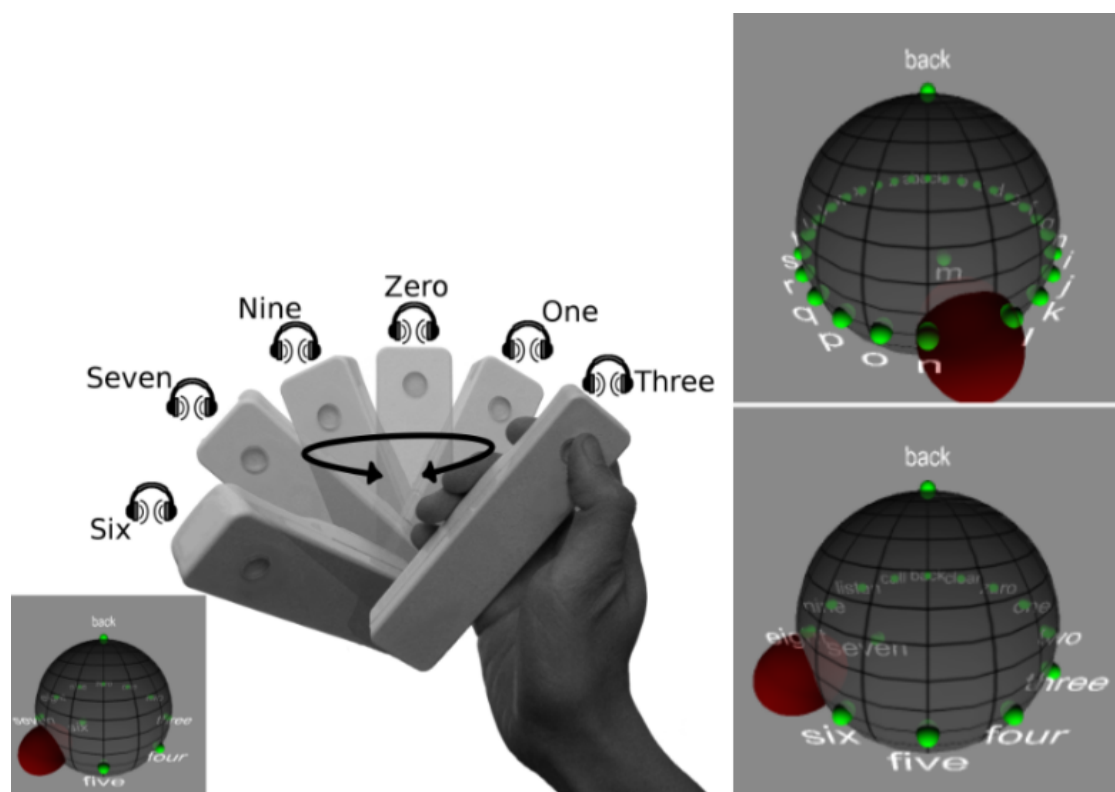


FIGURE 1.34 – Illustrations représentant le fonctionnement du dispositif de Kajastila et al. (2009). Illustrations issues de [105].

Ici, le degré de compatibilité est faible. Les gestes effectués sont très similaires avec leur effets sur le système si l'on considère leur représentation visuelle : la rotation du dispositif fait écho au pointeur autour de la sphère (figure 1.34). Mais dans un contexte sans retour visuel,

cette similarité est moins prégnante, le seul retour d'information obtenu est celui de l'objet d'intérêt actuellement sélectionné : par exemple, placer le dispositif dans une position particulière permet d'avoir un "A", un "R" dans une autre position - en somme, la translation entre l'orientation du dispositif et son effet sur le système n'est pas aussi évident.

En définitive, les modalités d'interaction utilisant les gestes en l'air respectent en majorité les principes de la manipulation directe des interfaces graphiques : les objets d'intérêt sont toujours accessibles, les interactions sont réversibles et se font essentiellement à l'aide d'actions physiques. Elles ont un niveau d'adéquation raisonnable, quoique le degré de compatibilité diverge grandement d'une modalité à l'autre. Bien que limitées en termes de résolution dans l'espace [128], le multiplexage spatial est suffisant pour offrir une expressivité correcte : la plupart des modalités d'interaction présentées permettent de choisir parmi plus d'une dizaine d'objets d'intérêt à un instant donné. Cette expressivité peut être augmentée en utilisant d'autres types de gestes, comme les gestes symboliques (par exemple, dessiner un "M" en l'air) ou métaphoriques (par exemple, tourner le poignet avec les doigts tendus comme pour tourner un bouton) [111, 9, 53]. Cependant, les travaux présentés dans la section sur le paradigme Tactile montrent que les PDV ont des difficultés à dessiner des formes régulières [109, 186, 137, 188] : les gestes symboliques nécessitant de dessiner sont donc peu accessibles aux PDV. De fait, il serait raisonnable de penser que les modalités d'interaction telles que l'AirStroke de Ni et al. [151], présenté en 2011, qui permet d'entrer du texte en dessinant les lettres en l'air (à raison de 6,5 WMP en situation de retour visuel), ne serait pas non plus accessible aux PDV. Dim et al. [114] ont par ailleurs confirmé dans leur étude d'élicitation que les PDV ne souhaitaient pas utiliser ces gestes symboliques en l'air comme langage d'interaction pour contrôler un système à distance.

A contrario, les gestes métaphoriques, souvent basés sur des habitudes d'usage de dispositifs physiques plus traditionnels (c'est-à-dire selon le principe du skeuomorphisme⁸), sont favorisés par les PDV [114]. Basés sur des habitudes d'usage antérieures, il semble raisonnable de penser que ces gestes seraient plus faciles à apprendre et à mémoriser.

8. Skeuomorphisme - Wikipedia : <https://fr.wikipedia.org/wiki/Skeuomorphisme>

Les modalités d'interaction basées sur les gestes effectués en l'air souffrent toutefois d'un défaut majeur : elles occasionnent de la fatigue pour l'utilisateur - aussi appelé syndrome du bras de gorille ("gorilla arm effect") [23, 100, 90]. En effet, cette fatigue se manifeste lorsqu'un membre est maintenu en position contre la gravité et dépend de plusieurs facteurs : l'amplitude et la durée de maintien de la contraction musculaire. Par conséquent une interaction gestuelle doit être concise et rapide afin de minimiser l'effort. En particulier, l'implémentation de contrôle gestuel doit éviter des gestes requérant une grande précision pendant une longue période. Cette caractéristique fait que les modalités d'interaction basées sur les gestes en l'air sont appropriées pour des commandes ponctuelles et rapides (par exemple, sélectionner un objet d'intérêt spécifique, à l'instar d'un marque-page) mais le sont beaucoup moins pour des tâches longues et/ou répétitives (par exemple, rédiger un long texte).

Nonobstant, il est possible de réduire l'amplitude des mouvements pour aussi réduire la fatigue engendrée par les gestes, notamment en effectuant les gestes plus près du corps, voire sur le corps, plutôt qu'en l'air.

En 2013, Perrault et al. [171] ont proposé WatchIt, une modalité d'interaction permettant d'utiliser des gestes (c'est-à-dire TAP et SWIPE) sur le poignet pour naviguer en situation de non-retour visuel dans 3 menus de 5 objets chacun. Les auteurs utilisent pour cela un bracelet de montre tactile mais on peut imaginer d'autres dispositifs de captation permettant d'interagir directement sur la peau - par exemple basés sur la vision par ordinateur [231]. En 2014, Lissermann et al. [130] ont proposé EarPut, une modalité d'interaction permettant d'utiliser le contour de l'oreille comme surface tactile et de contrôler un lecteur de musique. Xu et al. [245] ont proposé un système similaire en 2020 : earBuddy, qui permet d'effectuer des gestes directement sur la tête (par exemple, derrière l'oreille, sur la tempe, sur la joue), là aussi pour contrôler un lecteur de musique.

Comme pour les gestes en l'air, le degré d'intégration n'est pas pertinent. Le degré d'indirection spatiale est élevé si l'on considère que le geste est effectué à distance du système

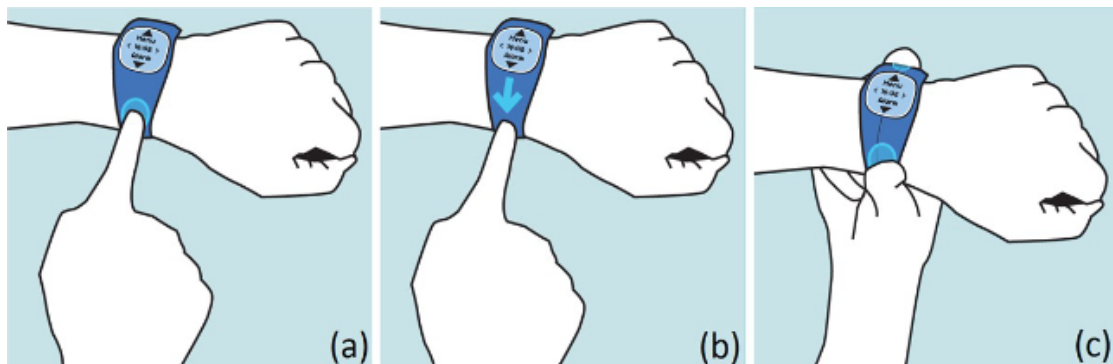


FIGURE 1.35 – Illustrations représentant le fonctionnement du dispositif WatchIt de Perrault et al. (2013). Illustrations issues de [171].

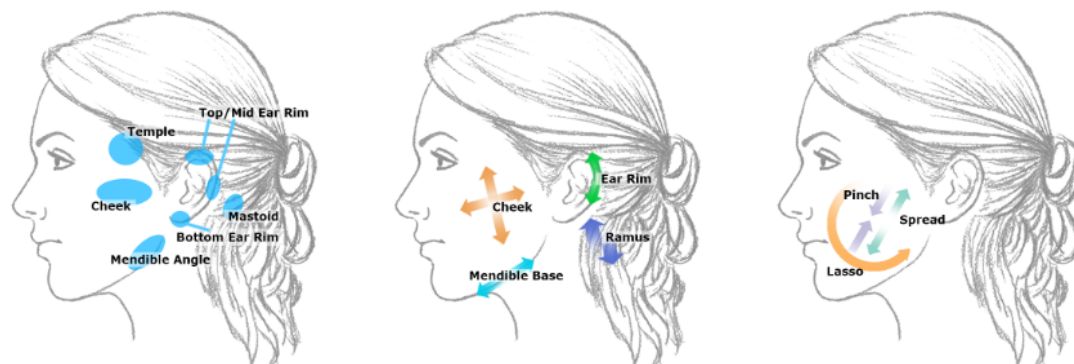


FIGURE 1.36 – Illustrations représentant les interactions possibles du dispositif EarBuddy de Xu et al. (2020). Illustrations issues de [245].

interactif. Il est faible si l'on considère que le geste est effectué à l'emplacement où le retour d'information est fait (par exemple, les oreilles pour un retour audio). Le degré de compatibilité n'est pas très élevé, les gestes et leur localisation étant arbitraires et rarement corrélés - que ce soit directement ou de façon métaphorique. Seul le geste permettant d'augmenter ou de baisser le volume à l'aide d'un SWIPE vers le haut ou vers le bas le long du rebord de l'oreille (EarPut [130]) - et d'autres, similaires - aurait pour sa part un degré de compatibilité élevé.

Par ailleurs, dans ces trois modalités, la localisation sur laquelle est effectué le geste a son importance : WatchIt [171] distingue deux zones sur le poignet (c'est-à-dire avant / arrière) elles même subdivisées en 5 (une par objet d'intérêt) et pouvant être accédées de façon directe via un TAP ; EarPut [130] divise le contour de l'oreille en quatre zones, elles aussi atteignables di-

rectement via un TAP ; tandis qu'earBuddy [245] permet d'interagir sur 7 zones différentes de la tête. Ces travaux montrent que la proprioception peut être très précise lorsqu'il s'agit de cibler une partie spécifique du corps en situation de non-retour visuel, davantage que lorsqu'il s'agit de cibler un point arbitraire dans l'espace. Cela permet un multiplexage spatial plus dense, augmente donc l'expressivité, et favorise les tâches de sélection.

Les modalités d'interaction utilisant les gestes effectués sur le corps offrent en sus un support physique, une tangibilité, qui permet d'utiliser la perception tactile en plus de la proprioception. Le retour d'information haptique qui en découle se fait à deux niveaux : celui de l'effecteur et celui du récepteur. Ce double retour d'information fait qu'il est possible d'agir et de percevoir simultanément dans l'espace l'endroit où s'effectue l'interaction, permettant donc d'être utilisé pour des tâches d'orientation, de position et de trajectoire en circuit fermé [75] même sans retour de la part du système. Le système earBuddy de Xu et al. [245], présenté précédemment, permet d'effectuer des commandes continues - bien qu'aucune n'ait été testée dans leur étude. Le système EarPut de Lissermann et al. [130] permet de changer le volume du lecteur de musique à l'aide d'une commande continue (c'est-à-dire un SWIPE le long de l'oreille). ThroughHand, présenté en 2021 par Jung et al. [103], transforme la main en "écran tactile" (figure 1.37), permettant de jouer en temps réel à des jeux 2D d'ordinaire réservés aux personnes voyantes en exploitant le potentiel du sens tactile à pouvoir relier spatialement les deux côtés de la main l'un par rapport à l'autre. Le dispositif transmet un stimuli haptique spatialisé et continu sur la paume de la main tandis que l'utilisateur utilise le dos de la main comme dispositif d'entrée, à la manière d'une surface tactile, lui permettant d'interagir de façon continue et en 2D - pour déplacer notamment un avatar dans un jeu.

Les modalités utilisant les gestes sur le corps peuvent être utilisées pour des tâches de saisie de texte. PalmGesture, présenté par Wang et al. [231] en 2015, est une modalité d'interaction permettant d'interagir avec des systèmes en dessinant des gestes sur la paume de la main sans retour visuel. Les gestes reconnus dans ce système sont des formes géométriques et des lettres. Les auteurs n'utilisent ces gestes dans leur système que comme des raccourcis vers des applications, toutefois, il semble tout à fait possible de faire plusieurs gestes à la suite

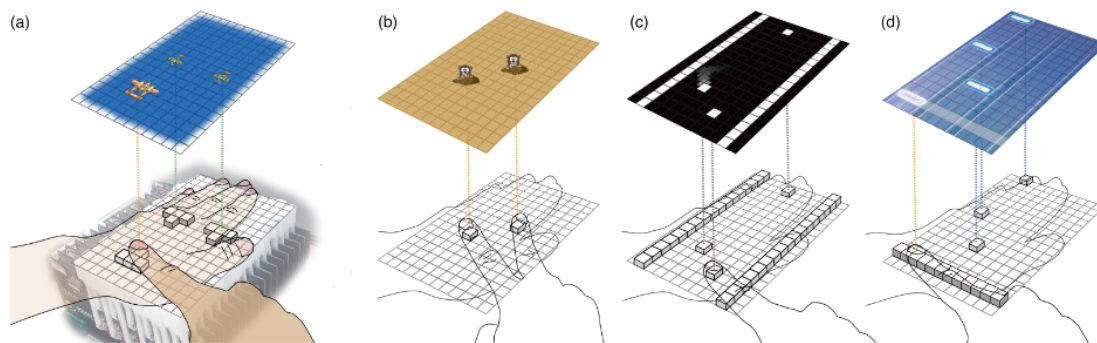


FIGURE 1.37 – Illustrations du fonctionnement du dispositif ThroughHand de Jung et al. (2021). Illustrations issues de [103].

pour écrire des mots et des phrases, à l'instar de Tinwala et al. [216], et tout comme ce dernier, PalmGesture possède un bon niveau de compatibilité.

Dans ces conditions, les gestes effectués sur le corps, en particulier sur la main, fonctionnent de manière très similaire à ceux du paradigme Tactile. Toutefois, la paume de la main renvoie davantage d'information qu'une surface tactile : le toucher se produit à un endroit facilement identifiable de la paume, tandis que les indices tactiles perçus par le doigt effecteur sont intrinsèquement ambigus [87]. Cette perception tactile spatiale fait que les interfaces du paradigme Gestuel utilisant les gestes effectués sur le corps peuvent être utilisées de façon spatiale, comme les interfaces graphiques pour les personnes voyantes. Notamment, et contrairement aux résultats de Kane et al. [109] utilisant une surface tactile, les utilisateurs de PalmGesture (des personnes voyantes avec les yeux bandés) sont parvenus à dessiner des formes symboliques, certaines fermées et/ou composées de plusieurs traits.

En définitive, les modalités d'interaction utilisant les gestes sur le corps sont similaires à celles utilisant les gestes en l'air : elles respectent en majorité les principes de la manipulation directe des objets de la tâche ; ont un niveau d'adéquation intermédiaire, quoique le degré de compatibilité diverge d'une modalité à l'autre ; et sont basées sur un multiplexage spatial suffisant pour offrir une expressivité correcte. Mieux encore, la proprioception et la perception tactile font que les gestes sur le corps peuvent être utilisés dans l'espace de façon plus précise que

les gestes en l'air, ce qui permet de proposer une densité d'affichage de l'interface plus grande, augmentant ainsi l'expressivité et réduisant la charge perceptive. Par ailleurs, contrairement au paradigme Tangible, qui requiert souvent un équipement ad hoc et encombrant, les modalités gestuelles en l'air et sur le corps ne requièrent pas de dispositif supplémentaire - sinon celui de captation, qui peut être intégré dans un dispositif grand public tel qu'un téléphone portable [105] ou une smartwatch [171]. Les langages d'interaction basés sur des gestes sont en outre considérés dans la littérature comme naturelles [126, 238] et intuitives [155, 239, 148], ce qui réduirait la charge cognitive.

Toutefois, les gestes en l'air comme les gestes sur le corps ont généralement une amplitude plus grande que les gestes sur téléphone portable et souffrent de leur manque de discrétion [184, 185, 211, 179, 93, 213]. Plusieurs travaux indiquent que les utilisateurs n'apprécient pas faire de grands gestes visibles et inhabituels, notamment en public. C'est particulièrement vrai lorsque l'interaction se fait avec un dispositif non visible par les tiers (par exemple, un téléphone dans la poche) [211]. Les résultats montrent également que les utilisateurs préfèrent les mouvements subtils et agréables, ceux similaires à ce qui existe déjà dans les dispositifs grand public et similaires à ceux utilisés dans la vie quotidienne.

Gestuel									
Articles	Modalité(s)	Fait pour / testé par des PDV	Adéquation	Charges de travail	Pénibilité	Portabilité	Versatilité	Discrétion et Confidentialité	Discrimination
7	< [dispositif], grammaire de gestes en l'air >	-	Indirection ✗ Intégration - Compatibilité ✓	Perception ✓ Cognition ✓ Motricité ✗	++	++	++	-- +	+/-
5	< [dispositif], grammaire de gestes sur le corps >	-	Indirection ✓ Intégration + Compatibilité ✓/✗	Perception ✓ Cognition ✓ Motricité ✓	+/-	++	++	+ +	-

FIGURE 1.38

1.6. ANALYSE DES PARADIGMES ET MODALITÉS D'INTERACTION POUR LES PERSONNES AYANT UNE DÉFICIENCE VISUELLE

Gestuel									
Article	Modalité(s)	Fait pour / testé par PDV	Adéquation	Charges de travail	Pénibilité	Portabilité	Versatilité	Discrétion et Confidentialité	Discrimination
Li et al. (2009)	< accéléromètres, orientation dans l'espace >	Non / Non (n=7)	Indirection ✗ Intégration - Compatibilité ✓	Perception ✓ Cognition ✓ Motricité ✗	++	++	++	-- +	+/-
Costa et al. (2020)	< caméra, grammaire de gestes >	Oui / Oui (n=16)	Indirection ✗ Intégration - Compatibilité ✓	Perception ✓ Cognition ✓ Motricité ✗	++	++	++	-- +	+/-
Dicke et al. (2015)	< optitrack, orientation du bras et gestes >	Non / Non (n=6)	Indirection ✗ Intégration - Compatibilité ✓	Perception ✓ Cognition ✓ Motricité ✗	++	++	++	-- +	+/-
Dim et al. (2016)	< indéfini, grammaire de gestes >	Étude d'élicitation (n=12)	Indirection ✗ Intégration - Compatibilité ✓	Perception ✓ Cognition ✓ Motricité ✗	+/-	++	++	-- +	+/-
Poupyrev et al. (2002)	< accéléromètre, degré d'inclinaison >	Non / Non (n=10)	Indirection ✓ Intégration - Compatibilité ✓	Perception ✗ Cognition ✓ Motricité ✗	+/-	++	++	-- +	+/-
Ni et al. (2011)	< gant électronique, orientation et gestes >	Oui / Oui (n=12)	Indirection ✗ Intégration - Compatibilité ✓	Perception ✓ Cognition ✓ Motricité ✗	++	++	++	-- +	+/-
Kajastila et al. (2009)	< accéléromètres, orientation dans l'espace >	Non / Non (n=11)	Indirection ✗ Intégration - Compatibilité ✗	Perception ✓ Cognition ✓ Motricité ✗	++	++	++	-- +	+/-
Perrault et al. (2013)	< surfaces tactiles, pointage direct et gestes >	Non / Non (n=8)	Indirection ✓ Intégration - Compatibilité ✗	Perception ✓ Cognition ✓ Motricité ✓	-	++	++	++ ++	-
Lissermann et al. (2014)	< surface tactile, pointage direct et gestes >	Non / Non (n=7)	Indirection ✓ Intégration - Compatibilité ✗	Perception ✓ Cognition ✓ Motricité ✓	+	++	++	- +	-
Xu et al. (2020)	< microphone, pointage direct et gestes >	Non / Non (n=12)	Indirection ✓ Intégration - Compatibilité ✗	Perception ✓ Cognition ✓ Motricité ✓	+	++	++	- +	-
Jung et al. (2021)	< caméra, pointage direct >	Oui / Oui (n=10)	Indirection ✓ Intégration - Compatibilité ✓	Perception ✓ Cognition ✓ Motricité ✓	-	-	++	+	+/-
Wang et al. (2015)	< caméras, dessin >	Non / Non (n=24)	Indirection ✗ Intégration - Compatibilité ✓	Perception ✓ Cognition ✓ Motricité ✓	+/-	++	++	+	-

FIGURE 1.39

1.6.5 Synthèse

Nous avons présenté les 4 paradigmes principaux utilisés par les PDV. Ils ont tous des avantages et des inconvénients :

- le Tactile est peu efficient mais c'est le paradigme par défaut des téléphones portables, il est intégré dans les usages et discret (tableau 1.17) ;
- le Vocal, si limité aux tâches de sélection ponctuelles et de saisie de texte, est très efficient, versatile et transportable, mais requiert une certaine charge cognitive et souffre d'un manque de discrétion et de confidentialité (tableau 1.20) ;
- le Tangible est très efficient mais requiert l'utilisation de dispositifs ad hoc généralement peu transportables et peu versatiles (tableau 1.29) ;

— le Gestuel est plutôt efficient - les gestes effectués sur le corps, plus particulièrement, bénéficient de la tangibilité sans souffrir de l'utilisation de dispositifs supplémentaires encombrants, offrent une bonne versatilité, ainsi qu'une bonne expressivité et bande-passante en raison de leur multiplexage dans l'espace et leur précision (tableau 1.38).

Dans la section suivante, nous étudions la combinaison du paradigme Tactile avec chacun des trois autres pour trouver une combinaison efficiente et accessible aux PDV.

1.7 Les combinaisons de paradigmes

Dans cette section, nous présentons des systèmes interactifs combinant plusieurs modalités d'interaction appartenant à des paradigmes différents. Plus spécifiquement, nous présentons des systèmes basés sur le paradigme Tactile et utilisant en plus une ou plusieurs modalités d'interaction appartenant à l'un des trois autres paradigmes présentés dans la section précédente (c'est-à-dire Vocal, Tangible, Gestuel) et les évaluons au regard de la même grille d'analyse utilisée précédemment. Dans le cadre de nos travaux, nous nous focalisons sur les écrans tactiles, notamment de téléphones portables.

1.7.1 Tactile + Vocal

Le paradigme Tactile combiné au Vocal implique l'utilisation de gestes (par exemple, TAP, SWIPE) sur une surface tactile et l'utilisation du mode vocal [26]. Nous représentons ce paradigme comme utilisant les modalités d'interaction suivantes : <écran tactile, grammaire de gestes>+ <microphone, langage naturel>ou <microphone, langage pseudo-naturel>ou <microphone, grammaire de sons>.

Voici sa représentation dans l'espace de Card et al. [44] :

Nous avons vu dans la section dédiée au paradigme Tactile que les modalités utilisant l'écran tactile est peu efficient et pose encore de nombreux défis aux PDV en raison d'une expressi-

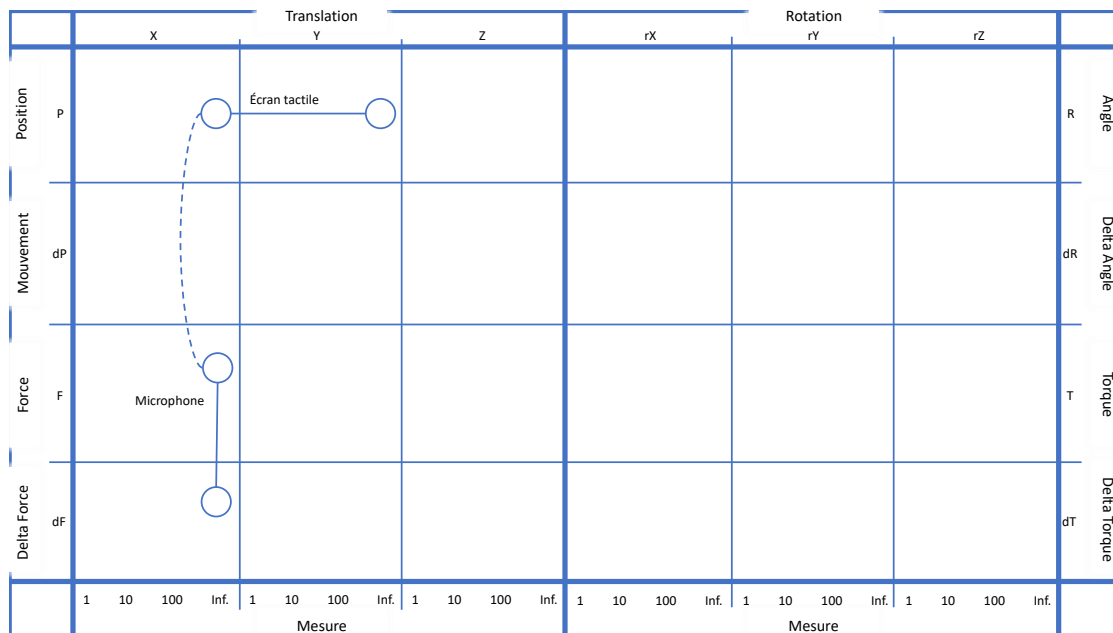


FIGURE 1.40 – Schéma représentant les dispositifs "écran tactile" et "microphone" dans l'espace de conception des dispositifs en entrée de Card. Inspiré de [44].

tivité et bande-passante faible, puisque le multiplexage spatial est très limité et le vocabulaire d'entrée est réduit à des gestes simples, peu nombreux, et possédant un faible degré de compatibilité avec les tâches d'interaction.

Le paradigme Vocal, quant à lui, en particulier les modalités utilisant le langage naturel, permet d'augmenter grandement cette expressivité, en particulier pour les tâches de sélection et de saisie. Cette combinaison est déjà présente dans les appareils grand public avec les assistants vocaux (par exemple, Siri sur iOS, Google Assistant sur Android), et beaucoup d'utilisateurs l'emploient pour des tâches de sélection ponctuelles (par exemple, ouvrir une application, appeler un numéro, faire une recherche sur internet) et saisir du texte. Nous avons vu [35, 228, 229] que dans ce cas, les modalités tactiles servent alors aux tâches de sélection répétitives et simples (par exemple, naviguer dans un menu élément par élément), nécessitant peu d'expressivité mais une forte bande-passante.

Azenkot [16] propose en 2014 SpeakNSwipe, un système permettant aux utilisateurs de

saisir, consulter et modifier du texte à l'aide de 5 gestes prédéfinis et du langage naturel en contexte de non retour visuel. En 2017, Duarte et al. [64] proposent un système similaire en Magicien d'Oz, avec davantage de gestes, qui sont définis par les utilisateurs. Dans ces deux systèmes, le texte est dicté de vive voix, et les gestes servent à naviguer dans le système (par exemple, débiter la diction, démarrer la relecture, sélectionner un segment précis à modifier ou à relire . . .). Ces systèmes ont un bon niveau d'indirection, d'intégration et de compatibilité, chaque modalité étant utilisée là où elle est la plus efficiente.

Cependant, ces deux paradigmes, même combinés, souffrent des mêmes problèmes qu'individuellement : notamment, leur manque de bande-passante et l'absence de retour d'information suffisant et en temps réel ne permettent pas d'effectuer des tâches de position, d'orientation, et de trajectoire en continu (c'est-à-dire circuit fermé [75]) de façon efficiente.

Malgré leur combinaison, les paradigmes Tactile + Vocal , sont peu adaptés à des tâches complexes et pose encore de nombreux défis en situation de non-retour visuel pour les mêmes raisons que lorsqu'ils sont utilisés de façon indépendante :

- dans le cas d'une sélection directe, ils placent sur l'utilisateur la charge d'identifier l'objet d'intérêt et de fournir un effort cognitif important pour réfléchir à la formulation des commandes ou des requêtes [54, 50] - en particulier lorsque les commandes sont énoncées dans une langue autre que l'anglais [30] pour des utilisateurs non-natifs en anglais ;
- dans le cas d'une navigation exploratoire, leur fonctionnement très largement multiplexé dans le temps, très séquentiel, fait que l'interaction par le biais de ces modalités est perçue comme insuffisante, répétitive et pénible pour les tâches qui requièrent peu d'expressivité mais une forte bande-passante [80, 229, 58] ;
- ceci est d'autant plus vrai qu'en terme de modalité de sortie l'interaction vocale est intangible et éphémère, contrairement à d'autres types d'interfaces possédant des repères (par exemple, visuels, spatiaux) permettant aux utilisateurs de se baser sur la reconnaissance (par exemple, icônes) plutôt que sur la mémorisation [153] ;
- de plus, le paradigme vocal peut être à éviter selon l'environnement, pour des raisons

d'acceptation sociale [54, 204, 15], de confidentialité [66, 68] ou à cause de bruit pouvant impacter les capacités du système à réagir correctement aux commandes vocales [60].

Tactile + Vocal									
Article	Modalité(s)	Fait pour / testé par PDV	Adéquation	Charges de travail	Pénibilité	Portabilité	Versatilité	Discretion et Confidentialité	Discrimination
2	< microphone, langage naturel > < microphone, langage pseudo-naturel > < microphone, grammaire de sons >	-	Indirection ✓/✗ Intégration ✓/✗ Compatibilité ✓/✗	Perception ✗ Cognition ✗ Motricité ✓	--	++	++	--	+/-
Azenkot et al. (2014)	< écran tactile, grammaire de gestes > < microphone, langage naturel >	Non / Non (n=9)	Indirection ✓/✗ Intégration ✓/✗ Compatibilité ✓/✗	Perception ✗ Cognition ✗ Motricité ✓	--	++	++	--	+/-
Duarte et al. (2017)	< écran tactile, grammaire de gestes > < microphone, langage naturel >	Oui / Oui (n=9)	Indirection ✓/✗ Intégration ✓/✗ Compatibilité ✓/✗	Perception ✗ Cognition ✗ Motricité ✓	--	++	++	--	+/-

FIGURE 1.41

1.7.2 Tactile + Tangible

Le paradigme Tactile combiné au Tangible implique l'utilisation de gestes (par exemple, TAP, SWIPE) sur une surface tactile et l'utilisation d'objets physiques que les utilisateurs peuvent manipuler pour interagir, la manipulation physique du dispositif étant directement associée à son état [20] (par exemple, la position d'un interrupteur ou d'un potentiomètre).

Du fait de la grande variabilité des objets physiques utilisables, la définition de cette combinaison de paradigmes est large. Les modalités sont multiples, et il serait impossible de toutes les modéliser dans l'espace de Card [44], tant les dispositifs (objets physiques manipulés) sont différents. Nous modélisons toutefois les exemples que nous traitons dans cette section à la figure 1.42.

En introduisant de la tangibilité sur un écran tactile, cette combinaison de paradigmes offre aux PDV l'opportunité d'utiliser leur perception tactile et leur proprioception pour s'orienter et interagir avec le système [202, 36, 122] tout en bénéficiant - dans une certaine mesure - de la modularité et la versatilité des écrans.

	Translation												Rotation											
	X			Y			Z			rX			rY			rZ								
Position P				Écran tactile Caméra Écran tactile				• TouchPlates • ClipWidgets													Angle R			
Mouvement dP				• El-Glaly et al. • Slate Master • Interactiles																		Delta Angle dR		
Force F																						Torque T		
Delta Force dF																						Delta Torque dT		
	1	10	100	Inf.	1	10	100	Inf.	1	10	100	Inf.	1	10	100	Inf.	1	10	100	Inf.	1	10	100	Inf.
	Mesure												Mesure											

FIGURE 1.42 – Schéma représentant les dispositifs "écran tactile" et tangibles dans l'espace de conception des dispositifs en entrée de Card. Inspiré de [44].

Les "TouchPlates" sont un bon exemple [108] (figure 1.43). Présentées par Kane et al. en 2013, il s'agit d'un ensemble d'accessoires conçus pour être superposés à l'écran tactile d'une tablette ou d'une table afin de fournir un guide perceptible tactilement pour une variété d'usages (par exemple, clavier, souris, menus, boutons...).

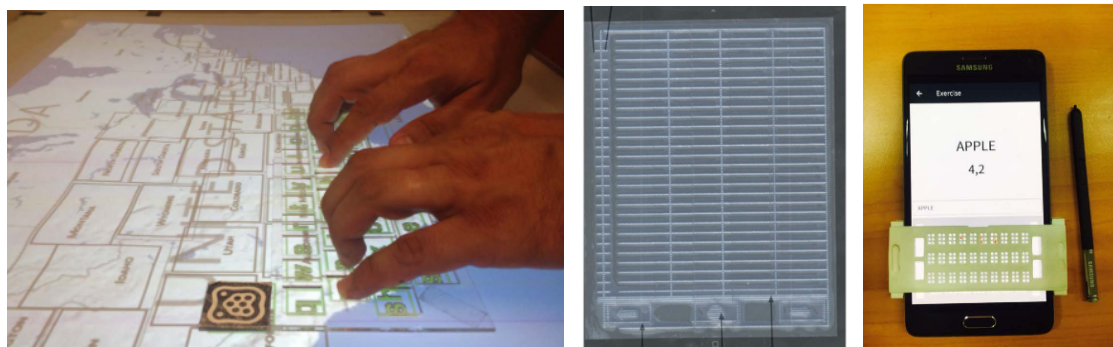


FIGURE 1.43 – Photographies de divers dispositifs Tactile + Tangible. **Gauche** : une TouchPlate (clavier) de Kane et al. (2013). **Centre** : surcouche pour liseuse de El-Glaly et al. (2013). **Droite** : le Slate Master de Lee et al. (2017). Photographies issues respectivement de [108], [69] et [125].

D'autres travaux ont une approche similaire, équipant divers dispositifs possédant un écran tactile lisse de surcouches tangibles : en 2013, El-Glaly et al. [69] ajoute une feuille de plastique texturée et ajourée permettant de suivre les lignes d'un texte et de localiser les boutons dans une application de lecture sur tablette (figure 1.43). En 2017, Lee et al. [125] présentent Slate Master, un outil didactique conçu pour faciliter l'apprentissage de l'utilisation de la tablette Braille (un outil utilisé par les PDV pour écrire manuellement du texte en Braille sur du papier) (figure 1.43). Le système se compose d'une réglette tangible, d'un stylet et d'une application qui utilise une interface de saisie personnalisée imitant une ardoise Braille. En 2018, Zhang et al. [250] proposent Interactiles, une coque imprimée en 3D venant ajouter des volets tangibles rabattables sur l'écran du téléphone portable (figure 1.44).



FIGURE 1.44 – Photographies du dispositif Interactiles de Zhang et al. (2018). Photographies issu de [250].

Enfin, en 2022, Visschedijk et al. [226] proposent ClipWidgets, une coque modulaire imprimée en 3D pour téléphones portables (figure 1.45). Ce système utilise un miroir conique et un étui de téléphone personnalisé pour rediriger le champ de vision de la caméra arrière

d'un téléphone portable vers la périphérie du téléphone pour détecter les entrées provenant de contrôleurs tangibles imprimés en 3D fixés à l'étui du téléphone.



FIGURE 1.45 – Photographies du dispositif ClipWidgets de Visschedijk et al. (2022). Photographies issu de [226].

Dans ces travaux, le degré d'indirection spatiale est négligeable, puisque l'objet d'intérêt est manipulé directement. En outre, leur tangibilité leur confère une permanence qui, lorsqu'ils se superposent à l'écran tactile, permet de rendre les éléments de l'interface graphique perceptibles et utilisables pour les PDV pour toutes les tâches d'interaction, y compris les tâches d'orientation, position et trajectoire de façon continue (c'est-à-dire circuit fermé [75]). Le degré d'intégration est de 1 et le degré de compatibilité avec leurs tâches respectives est très élevé puisque les dispositifs sont conçus de façon ad hoc.

De plus, les problèmes de charge cognitive dus à l'exploration tactile de l'interface, ainsi que ceux de la charge physique, sont grandement diminués du fait de la taille réduite des dispositifs par rapport aux dispositifs uniquement tangibles présentés précédemment.

Cependant ces dispositifs sont peu pratiques à utiliser au quotidien : s'ils sont moins encombrants que les dispositifs tangibles stationnaires décrits précédemment, ils n'en restent pas moins volumineux pour certains (par exemple, TouchPlates, Interactiles, ClipWidgets [108, 250, 226]1) - à minima, ils nécessitent d'utiliser des éléments supplémentaires (susceptibles d'être facilement perdus ou cassés) qui réduisent la praticité et la mobilité du téléphone portable utilisé seul. Ils sont en outre limités à des usages très spécifiques de par leur format, les dispositifs étant ad hoc, et ne peuvent souvent pas être modifiés aisément. Étant donné que chaque application a une disposition et une utilité différente, cette solution nécessiterait l'utilisation de nombreux accessoires supplémentaires.

Tactile + Tangible									
Article	Modalité(s)	Fait pour / testé par PDV	Adéquation	Charges de travail	Pénibilité	Portabilité	Versatilité	Discretion et Confidentialité	Discrimination
5		-	Indirection ✓/X Intégration ✓/X Compatibilité ✓/X	Perception X Cognition X Motricité ✓	--	+	-	+	+/-
El-Ghaly et al. (2013)	< écran tactile, gestes et manipulation d'objets >	Oui / Oui (n=16)	Indirection ✓ Intégration ✓ Compatibilité ✓	Perception ✓ Cognition ✓ Motricité ✓	--	+	-	+	+/-
Kane et al. (2014)	< écran tactile, boutons >	Oui / Oui (n=9)	Indirection ✓ Intégration ✓ Compatibilité ✓	Perception ✓ Cognition ✓ Motricité ✓	--	+	+/-	+	+/-
Lee et al. (2017)	< écran tactile, dessin >	Oui / Oui (n=3)	Indirection ✓ Intégration ✓ Compatibilité ✓	Perception ✓ Cognition ✓ Motricité ✓/X	--	+	--	+	+/-
Zhang et al. (2018)	< écran tactile, gestes et manipulation d'objets >	Oui / Oui (n=5)	Indirection ✓ Intégration ✓ Compatibilité ✓	Perception ✓ Cognition ✓ Motricité ✓	--	-	+/-	+	+/-
Visschedijk et al. (2022)	< écran tactile, gestes et manipulation d'objets >	Non / Non (n=)	Indirection ✓ Intégration ✓ Compatibilité ✓	Perception ✓ Cognition ✓ Motricité ✓	--	--	+/-	+	+/-

FIGURE 1.46

1.7.3 Tactile + Gestuel

Le paradigme Tactile combiné au Gestuel implique l'utilisation de gestes (par exemple, TAP, SWIPE) sur une surface tactile et ceux effectués en l'air et/ou sur le corps de l'utilisateur. Nous représentons ce paradigme comme utilisant les modalités d'interaction suivantes : <écran tactile, grammaire de gestes>+ <[dispositif], grammaire de gestes en l'air>, <[dispositif], grammaire de gestes sur le corps>.

La définition de ce paradigme est donc large elle aussi, et ne s'arrête pas à quelques modalités d'interactions. Celles-ci sont au contraire multiples, et il serait impossible de toutes les modéliser dans l'espace de Card, tant les dispositifs sont différents. Nous modélisons toutefois les exemples que nous traitons dans cette section à la figure 1.47.

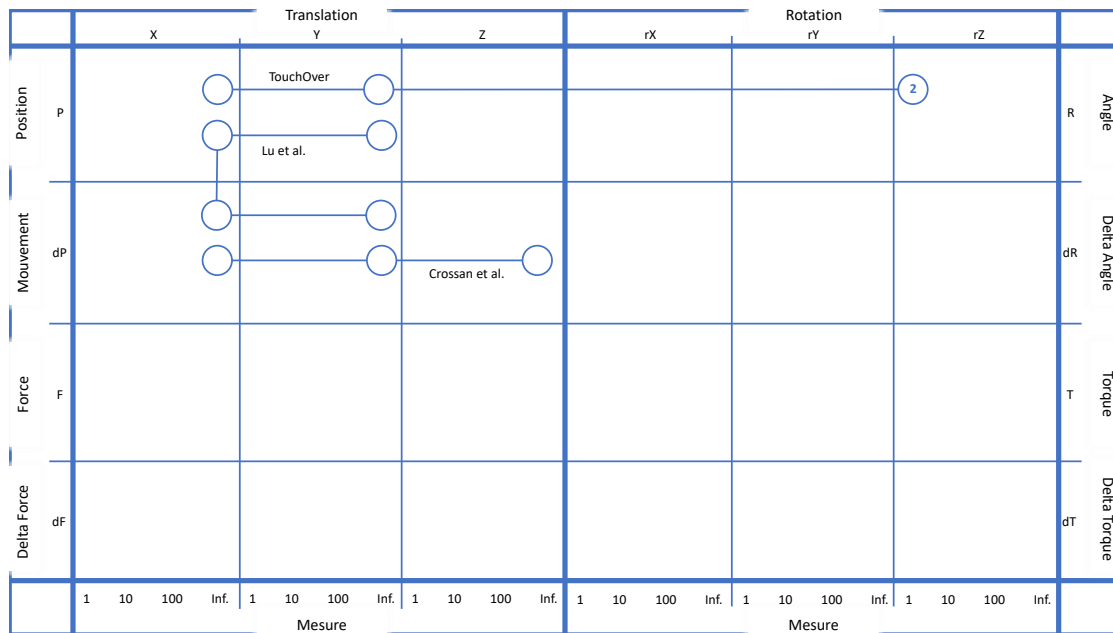


FIGURE 1.47 – Schéma représentant les dispositifs "écran tactile" et gestuel dans l'espace de conception des dispositifs en entrée de Card. Inspiré de [44].

À l'instar du paradigme Tangible, le paradigme Gestuel offre aux PDV la possibilité d'utiliser leur proprioception pour interagir avec le système [91]. Cependant, les gestes utilisés dans le paradigme Gestuel pourraient gêner l'utilisation de l'écran tactile : par exemple, il serait certainement difficile d'effectuer une tâche quelconque avec le téléphone portable tandis que celui-ci est déplacé dans l'espace à bout de bras comme avec VirtualShelves [128], mentionné précédemment (figure 1.32).

Pour permettre une utilisation simultanée des gestes et de l'écran tactile, il est nécessaire que les modalités d'interaction ne se parasitent pas l'une l'autre, par exemple, en ayant une amplitude de mouvement trop large. Un système qui serait basé sur l'orientation du téléphone

portable à la manière de Poupyrev et al. [175], voire Ni et al. [152], et permettrait d'accéder à des commandes différentes selon l'angle de l'appareil pourrait convenir. Scoditti et al. [198] notamment présentent TouchOver en 2011, une technique d'interaction utilisant l'inclinaison du téléphone portable sur le plan sagittal pour basculer entre deux modes (c'est-à-dire "aperçu" et "sélection") lorsque le doigt est en contact avec l'écran tactile. Ce geste est utilisé comme interrupteur et ajoute de l'expressivité à l'écran tactile de manière multiplicative : c'est à dire qu'un geste permettant de basculer entre 2 modes multiplie le nombre de commandes par 2.

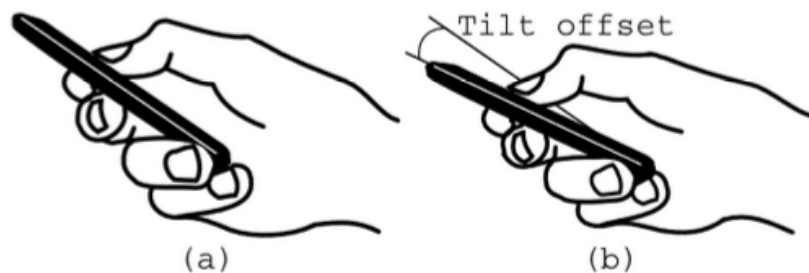


FIGURE 1.48 – Schéma représentant le fonctionnement du dispositif TouchOver de Scoditti et al. (2011). Illustration issue de [198].

Les gestes et l'écran tactile peuvent aussi être utilisés de façon séquentielle : en 2011, Ruiz et al. [190] ont présenté les résultats d'une étude d'élicitation de gestes effectués en l'air pour invoquer des commandes sur un téléphone portable en situation de retour visuel. Ces gestes servent 1) soit à des actions ponctuelles (par exemple, lever le téléphone portable pour répondre à un appel, le secouer pour retourner à l'écran d'accueil ...) et sont utilisables indépendamment de l'écran tactile ; 2) soit à des actions répétitives de navigation/manipulation (par exemple, bouger rapidement le téléphone portable vers la gauche/droite pour aller à l'élément précédent/suivant, l'avancer/reculer devant soi pour contrôler le zoom sur un élément ...). Ces gestes peuvent donc être utilisés comme raccourcis vers des fonctions, ou pour naviguer dans une interface et manipuler ses éléments. Ils ajoutent ainsi de l'expressivité à l'écran tactile de manière additive : c'est à dire que 5 gestes vont ajouter 5 nouvelles commandes au système.

Peu de travaux existent sur le sujet pour PDV, mais nous avons vu que dans un contexte

sans retour visuel, l'utilisation de la proprioception en l'air (et donc à fortiori, l'orientation d'un dispositif en l'air) seule ne serait ni suffisamment précise ni confortable à utiliser (pour rappel, à cause du syndrome du bras de gorille [23, 100, 90]) et manquerait par ailleurs singulièrement de discrétion.

Une autre possibilité serait d'utiliser des gestes utilisant une autre partie du corps - en particulier les pieds. L'orientation et la position de ceux-ci pourraient modifier en temps réel le comportement des gestes effectués sur l'écran tactile. Plus simplement, les gestes effectués avec le pied ajouteraient autant de commandes au système : Lu et al. [135] ont présenté en 2012 un jeu de football jouable en situation de retour visuel sur téléphone portable, utilisant le pointage du doigt sur l'écran tactile pour contrôler le gardien et des gestes du pied pour contrôler le ballon. Crossan et al. [55] ont présenté en 2010 une technique d'interaction utilisant des TAP effectués avec le pied pour naviguer dans un menu hiérarchique en situation de non-retour visuel. Ils montrent que cette technique d'interaction est utilisable avec un taux d'erreur comparable à la condition de référence (c'est-à-dire usage avec la main en situation de retour visuel) mais est sensiblement plus longue, en particulier si la trajectoire d'interaction est supérieure à 4 commandes. Par ailleurs, Velloso et al. [225] rapportent que les gestes nécessitant de soulever ou de glisser le pied sur le sol sont fatigants pour l'utilisateur, tandis que Scott et al. [199] montrent que les gestes de rotation du pied sont peu précis et recommande de discrétiser les commandes en entrée utilisant ces gestes. De plus, les gestes effectués avec le pied peuvent manquer de discrétion - en particulier s'ils sont répétés - car la situation pourrait paraître inhabituelle.

La dernière option serait d'utiliser les gestes sur le corps à l'aide d'une main tandis que l'autre main utilise le téléphone portable. Toutefois, à l'instar des gestes effectués en l'air ou avec les pieds par exemple, ces gestes peuvent manquer de discrétion surtout s'ils sont utilisés de façon récurrente.

En effet, bien qu'ils puissent potentiellement démultiplier l'expressivité des écrans tactiles, les gestes sur le corps souffrent de leur manque de discrétion. Plusieurs travaux [184, 185,

211, 179, 93, 213] indiquent que les utilisateurs n'apprécient pas faire de grands gestes visibles et inhabituels en public. Les utilisateurs préfèrent les mouvements subtils et agréables, ceux similaires à ce qui existe déjà dans les dispositifs grand public et similaires à ceux utilisés dans la vie quotidienne.

Tactile + Gestuel									
Article	Modalité(s)	Fait pour / testé par PDV	Adéquation	Charges de travail	Pénibilité	Portabilité	Versatilité	Discretion et Confidentialité	Discrimination
4	-	-	Indirection ✗ Intégration - Compatibilité ✓	Perception ✓ Cognition ✓ Motricité ✗	+/-	++	++	-- +	+/-
Scoditti et al. (2011)	< smartphone, gestes et orientation du dispositif >	Non / Non (n=18)	Indirection ✓ Intégration - Compatibilité ✗	Perception ✓ Cognition ✓ Motricité ✗	+/-	++	++	- ++	-
Ruiz et al. (2011)	< smartphone, mouvements et orientation du dispositif >	Étude d'élicitation (n=20)	Indirection ✓ Intégration - Compatibilité ✓	Perception ✓ Cognition ✓ Motricité ✗	+/-	++	++	-- +	+/-
Lu et al. (2012)	< smartphone, gestes sur l'écran et gestes du pied en l'air >	Non / Non (n=12)	Indirection ✓ Intégration - Compatibilité ✓	Perception ✓ Cognition ✓ Motricité ✓/✗	+/-	++	++	- ++	-
Crossan et al. (2010)	< accéléromètre, grammaire de gestes du pied >	Non / Non (n=12)	Indirection ✗ Intégration - Compatibilité ✓	Perception ✓ Cognition ✓ Motricité ✓/✗	+	++	++	- ++	-

FIGURE 1.49

Pour répondre à ce défi, il existe un sous-ensemble du paradigme Gestuel, plus discret et moins fatigant car utilisant des gestes avec de petites amplitudes : les micro-gestes (μ G). Objet de nos travaux, nous définissons ce paradigme et nous étudions sa combinaison avec le paradigme tactile.

1.8 Le paradigme Micro-Gestuel

Le paradigme Micro-Gestuel (voire Microgestuel, ou encore μ Gestuel) est contenu dans la paradigme Gestuel. Indépendamment du dispositif d'entrée, il s'agit des modalités d'interaction utilisant comme langage d'interaction des mouvements du corps effectués avec de petites amplitudes. Il n'existe pas de définition consensuelle de la taille de cette amplitude, ni même des microgestes (μ G) en général :

- Ashbrook [12] décrit les μ Gestes comme de petits mouvements délibérés, effectués en tant qu'entrée, et qui durent moins de quatre secondes ;
- Wolf et al. [241, 240]. décrivent les μ Gestes comme de petits mouvements exécutés avec les mains et les doigts ;
- Chan et al. [49] définissent les μ Gestes à une main comme des gestes effectués par un doigt d'une main sur lui-même ;
- ... mais d'autres parties du corps peuvent également être utilisées, par exemple, les pieds [10, 55], la tête [59], les yeux [98, 215], voire la mâchoire [13, 79, 51].

Chan et al. [49] ont mené une étude d'élicitation de 1632 μ Gestes à une seule main, et ont constaté que le pouce était utilisé 88% du temps : il s'agit en particulier de ce que nous appelons les μ Gestes "thumb-to-finger" (TTF) (c'est-à-dire le pouce touchant un autre doigt).

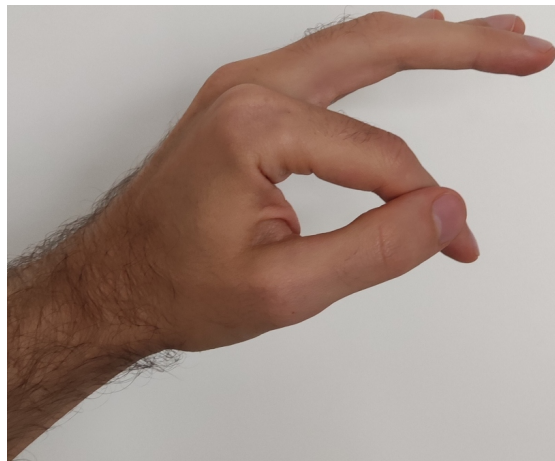


FIGURE 1.50 – Photographie illustrant un μ Geste TTF, plus spécifiquement un TAP du pouce sur l'index.

Dans cette section, ainsi que dans le reste du manuscrit, nous nous focalisons sur les μ Gestes TTF, que nous représentons comme étant la modalité d'interaction suivante : <[dispositif], grammaire de gestes du pouce sur un autre doigt>.

La définition de ce paradigme ne permet pas de préciser le dispositif de captation, les études menées portant sur les μ Gestes TTF utilisant au choix : un système stationnaires et/ou des ap-

pareillages portatifs (par exemple, gants, anneaux, montres . . .) utilisant la vision par ordinateur [233], des radars [232], des champs magnétiques [70], des capteurs analogiques (par exemple, de flexion, de pression, de contact . . .) [245] voire des microphones [117]. Il serait impossible de tous les modéliser dans l'espace de Card, tant les dispositifs sont différents. Nous modélisons toutefois les exemples issus de la littérature que nous étudions dans cette section à la figure 1.47.

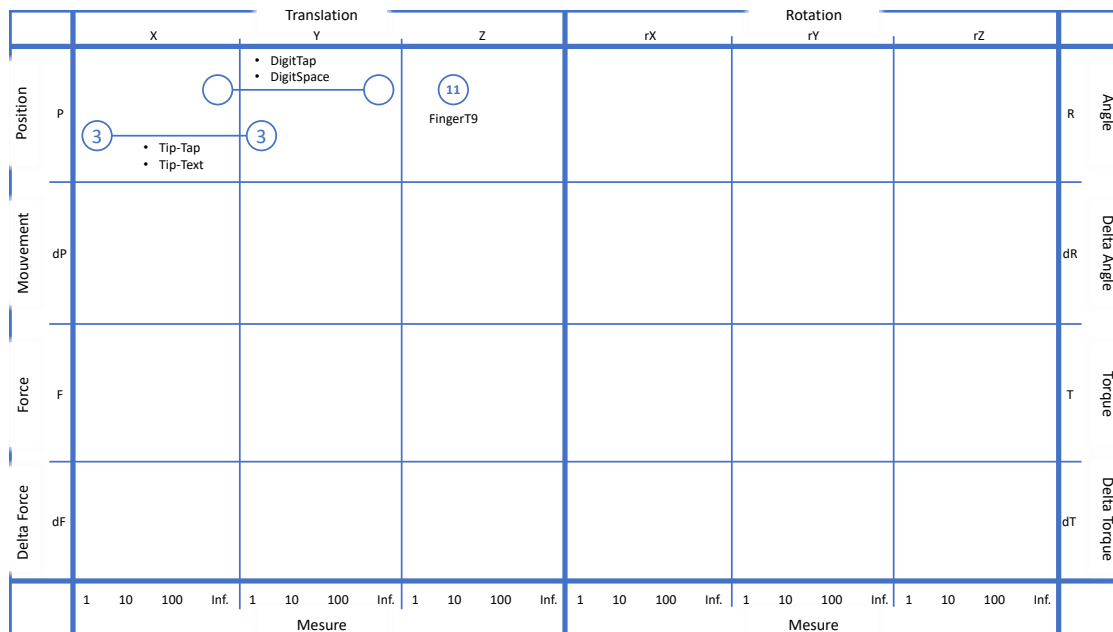


FIGURE 1.51 – Schéma représentant les dispositifs Micro-Gestuels dans l'espace de conception des dispositifs en entrée de Card. Inspiré de [44].

Kuo et al. ont étudié l'espace de travail fonctionnel du pouce et des doigts (c'est-à-dire la plage de mouvements) et ont constaté que les paires Pouce-Index et Pouce-Majeur avaient l'espace de travail fonctionnel maximal [118]. Huang et al. ont constaté que les μ Gestes de TAP et de Swipe effectués par le pouce sur l'index ou le majeur de la main dominante étaient confortables [94]. Chaque doigt peut être divisé en deux ou trois segments, se superposant principalement aux phalanges, qui sont clairement délimitées par les articulations et les plis. Plusieurs études ont utilisés les phalanges comme cibles tactiles [101, 178, 219, 237], tandis que d'autres les ont subdivisé davantage [220, 94].

La préférence et le confort perçu des zones des doigts pour les μ Gestes TTF ont été explorés par plusieurs études [94, 101, 110, 219, 230]. Elles montrent que l'index et le majeur sont les plus confortables à utiliser, suivis par l'annulaire puis le petit doigt dans un ordre décroissant. Le classement du confort suit une tendance similaire le long du doigt, avec l'extrémité du doigt qui est préférée, suivie de la section médiane, puis de la section inférieure. La préférence entre le côté volaire (côté paume) et le côté radial (côté faisant face au pouce) du doigt a été étudiée et a montré une tendance vers le côté radial pour l'index et le majeur, mais le côté volaire pour l'annulaire et le petit doigt [219, 101]. La zone de l'ongle a également été étudiée : plusieurs études concluent qu'il s'agit d'un emplacement viable, toujours disponible et apprécié [110, 124].

En exploitant la dextérité de la main, ainsi que la proprioception et le retour haptique propres aux gestes effectués sur le corps, les μ Gestes TTF offrent une grande expressivité car ils permettent une variété de gestes possibles (par exemple, TAP, Swipes) qui peuvent être réalisés avec précision sur des parties distinctes des doigts [208], et ce même sans retour visuel [178, 94, 246, 112]. En 2014, Prätorius et al. ont proposé DigitTap [178], une modalité d'interaction permettant d'effectuer des Taps du pouce sur douze emplacements situés sur les autres doigts (c'est-à-dire les articulations et le bout des doigts) (figure 1.52). Dans une étude utilisateur avec 12 participants, ils obtiennent un taux de réussite de 92%.

Huang et al. [94] présentent en 2016 une étude dans laquelle 14 participants doivent positionner leur pouce sur une zone précise de l'un des quatre autres doigts en situation de non-retour visuel. Ils testent plusieurs agencements avec un nombre de zones à atteindre variable et des zones réparties plus ou moins densément sur les doigts (figure 1.53). Leurs résultats expérimentaux suggèrent que les participants peuvent discriminer jusqu'à 5 zones sur l'index et le majeur, 4 sur l'annulaire et 2 sur l'auriculaire - pour un total de 16 zones.

Dans une autre étude, 12 participants doivent utiliser leur pouce pour effectuer des gestes en formes de lettres sur leur index seul, ou sur leur index et leur majeur joints, à la manière

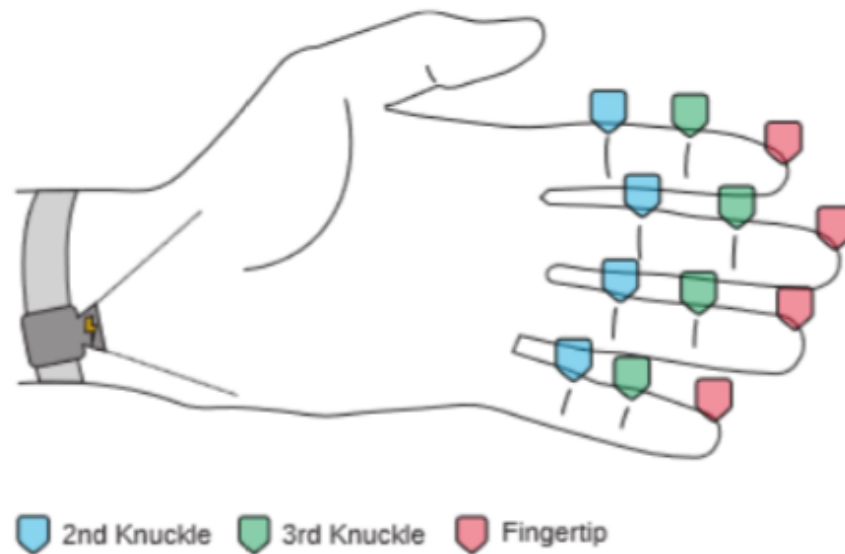


FIGURE 1.52 – Schéma représentant le dispositif DigitTap de Prätorius et al. (2014). Illustration issue de [178].

d'une surface tactile. Leurs résultats indiquent que la tâche est réalisable et obtiennent un taux de précision de 92,7% à l'aide d'un reconnaiseur personnalisé à chaque participant. Basé sur ces résultats, ils proposent DigitSpace, un prototype permettant de contrôler une smartwatch à l'aide de μ Gestes TTF. L'utilisateur peut facilement saisir des chiffres ou déclencher une fonction souhaitée en appuyant à l'aide du pouce sur une certaine zone de l'un des doigts. L'utilisateur change de mode en joignant l'index et le majeur, puis en effectuant rapidement un double TAP sur l'une des phalanges. Il peut ensuite soit dessiner sur les doigts pour saisir du texte, soit faire glisser son pouce le long d'un doigt pour contrôler une barre de défilement. Les auteurs avancent qu'un utilisateur familiarisé avec l'interface peut manipuler la smartwatch sans retour visuel.

Katsuragawa et al. [112] ont proposé Tip-Tap en 2019, une modalité d'interaction permettant d'entrer des commandes à l'aide de gestes du pouce sur l'index (figure 1.54). Leur système reconnaît des commandes tactiles discrètes en deux dimensions en détectant l'intersection entre deux séries de points de contact : une série le long de l'extrémité du doigt index et l'autre le long de l'extrémité du pouce. Ils montrent que la précision sans regarder est de 86% après une

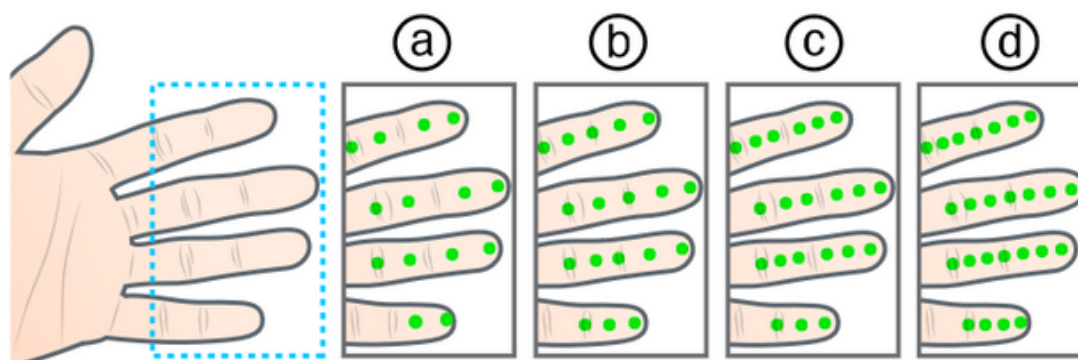


FIGURE 1.53 – Schéma représentant les zones à atteindre dans l'expérience de Huang et al. (2016). Illustration issue de [94].

très courte période d'apprentissage, et qu'ajouter des retours haptiques passifs (par exemple, relief, aimants) aux contacts n'est pas nécessaire.

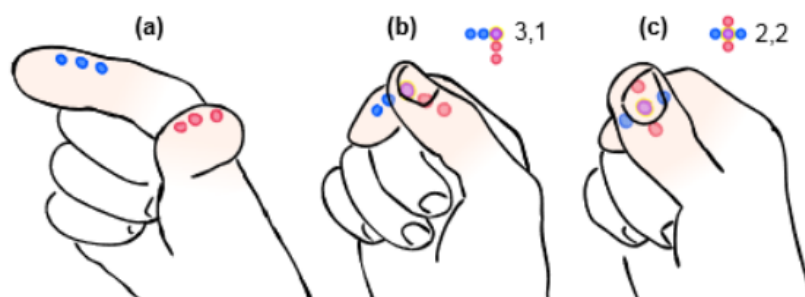


FIGURE 1.54 – Schéma représentant le dispositif Tip-Tap de Katsuragawa et al. (2019). Illustration issue de [112].

Ces travaux montrent que les μ Gestes TTF sont très efficaces pour les tâches de sélection. Dans ces travaux, le degré d'indirection est élevé sans que cela n'influence l'interaction. Le degré de compatibilité est intermédiaire pour les tâches de sélection - qui se font de manière indirecte, à l'aide de raccourcis arbitraires (c'est-à-dire une zone sur un doigt correspond à une fonction). Le degré de compatibilité est élevé pour les tâches de position et de trajectoire (dessin), car le geste est très similaire à l'effet sur l'objet d'intérêt. Étant des gestes effectués sur le corps, les μ Gestes TTF offrent un support physique, une tangibilité, qui permet d'utiliser la perception tactile en plus de la proprioception. En effet comme déjà expliqué dans la section sur le

paradigme gestuel, le retour d'information haptique qui en découle se fait à deux niveaux : celui de l'effecteur et celui du récepteur. Ce double retour d'information fait qu'il est possible d'agir et de percevoir en continu et simultanément dans l'espace l'endroit où s'effectue l'interaction, permettant donc d'être utilisé pour des tâches d'orientation, de position et de trajectoire en circuit fermé [75] même sans retour de la part du système - de façon très similaire à PalmGesture, présenté par Wang et al. [231].

Les μ Gestes TTF permettent aussi de réaliser des tâches de saisie de texte. Wong et al. [242] ont proposé FingerT9 en 2018, une modalité d'interaction permettant de saisir du texte avec le pouce à l'aide d'un clavier T9 associé aux phalanges des autres doigts (figure 1.56). Ils ont montré que la saisie de texte sans regarder (c'est-à-dire en regardant la sortie sur l'écran mais pas la main) est possible une fois que les participants sont familiers avec la disposition du clavier.



FIGURE 1.55 – Schéma représentant le dispositif FingerT9 de Wong et al. (2018). Illustration issue de [242].

Xu et al. [246] ont proposé TipText en 2019, une modalité d'interaction utilisable sans retour visuel permettant de saisir du texte en manipulant à l'aide du pouce un clavier miniature sous

la forme d'une grille 3x3 et située sur la première phalange de l'index. Leurs résultats indiquent que les participants ont atteint une vitesse de saisie de texte moyenne de 11,9 WPM. Dans ces travaux, le degré d'indirection est élevé sans que cela n'influence l'interaction et le degré de compatibilité est faible, la sélection des objets d'intérêts (c'est-à-dire les caractères) se fait de manière indirecte, à l'aide de raccourcis arbitraires (c'est-à-dire une zone sur un doigt correspond à un caractère).

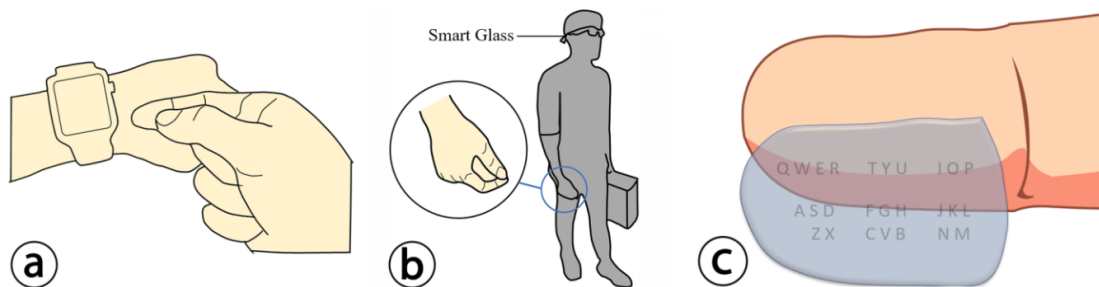


FIGURE 1.56 – Schéma représentant le dispositif TipText de Xu et al. (2019). Illustration issue de [246].

Le langage de ces modalités d'interaction repose sur la localisation sur laquelle est effectué le geste, et les études présentées montrent que la proprioception et la perception tactile peuvent être très précises lorsqu'il s'agit de cibler une partie spécifique de la main en situation de non-retour visuel [178, 94, 246, 112]. Cela permet un multiplexage spatial dense, augmentant l'expressivité : la plupart des modalités d'interaction présentées permettent de choisir parmi plus d'une dizaine d'objets d'intérêt à un instant donné. Par ailleurs, cette expressivité peut être augmentée en utilisant des gestes sémiotiques [43] (par exemple, dessiner un symbole) plutôt qu'ergotiques (c'est-à-dire appuyer sur un bouton - ou en l'occurrence, une phalange) comme avec DigitSpace de Huang et al. [94].

En synthèse, les modalités d'interaction utilisant les μ Gestes TTF sont similaires à celles utilisant les gestes sur le corps : elles respectent en majorité les principes de la manipulation directe des interfaces graphiques ; ont un niveau d'adéquation intermédiaire, quoique le degré de compatibilité soit faible pour les tâches de sélection ; et mieux encore, la proprioception et la

perception tactile font que les μ Gestes TTF peuvent être utilisés dans l'espace de façon précise ce qui permet de proposer une densité d'affichage de l'interface plus grande, augmentant ainsi l'expressivité et réduisant la charge perceptive. Par ailleurs, contrairement au paradigme Tangible, qui requiert souvent un équipement ad hoc et encombrant, les modalités μ Gestuelles ne requièrent pas de dispositif supplémentaire - sinon celui de captation, qui peut être intégré dans un dispositif grand public tel qu'un téléphone portable [195], une smartwatch [7], voire même un smartring [240]. Nous avons vu en outre que les modalités d'interaction basées sur des gestes sont considérées dans la littérature comme naturelles [126, 238] et intuitives [155, 239, 148], ce qui réduirait la charge cognitive.

De plus, les μ Gestes TTF ont généralement une amplitude très faible, du même ordre de grandeur que les gestes effectués sur téléphone portable, et inférieure aux autres gestes sur le corps. Ils sont de fait plus discrets, plus confidentiels et moins fatiguants.

Plusieurs études ont par ailleurs montré que les μ Gestes TTF étaient utilisables en parallèle d'une tâche principale [89, 212, 31]. Häulschmid et al. [89] les ont utilisés en situation de conduite automobile, Tan et al. [212] en conduite de vélo, et Boldu et al. [31] en situation d'activité sportive (c'est-à-dire course). Dans une étude d'élicitation de 2019, Sharma et al. [203] ont montré que les μ Gestes TTF sont utilisables même en situation de main occupée et donc utilisables dans une variété de situations du quotidien (par exemple, dans le métro en se tenant à une barre, avec un sac dans la main, en cuisine ...). S'il ne s'agit pas à proprement parler d'un usage en situation de non-retour visuel (bien que cela puisse l'être, puisque la vue est principalement affectée à une autre tâche), ces études (et d'autres [240, 244, 230]) montrent que les modalités utilisant les TTF μ G ont un faible impact sur la charge de travail des tâches d'interaction à effectuer, qu'elles s'inscrivent dans une activité principale ou secondaire.

Toutes ces propriétés font que, à l'instar des modalités basées sur les gestes effectués sur le corps vu précédemment, les modalités d'interaction basées sur les μ Gestes TTF semblent particulièrement prometteuses pour les PDV et c'est pourquoi nous choisissons de les étudier comme complément au paradigme Tactile.

Micro-Gestuel									
Articles	Modalité(s)	Fait pour / testé par des PDV	Adéquation	Charges de travail	Pénibilité	Portabilité	Versatilité	Discretion et Confidentialité	Discrimination
5	< [dispositif], grammaire de gestes du pouce sur les autres doigts >	Non	Indirection ✗ Intégration - Compatibilité ✓/✗	Perception ✓ Cognition ✓ Motricité ✓	--	++	++	++ ++	--

FIGURE 1.57

Micro-Gestuel									
Article	Modalité(s)	Fait pour / testé par PDV	Adéquation	Charges de travail	Pénibilité	Portabilité	Versatilité	Discretion et Confidentialité	Discrimination
Prătorius et al. (2014)	< caméra, position du pouce sur les autres doigts >	Non / Non (n=12)	Indirection ✗ Intégration - Compatibilité ✗	Perception ✓ Cognition ✓ Motricité ✓	--	++	++	++ ++	--
Huang et al. (2016)	< caméra, grammaire de gestes >	Non / Non (n=12)	Indirection ✗ Intégration - Compatibilité ✓/✗	Perception ✓ Cognition ✓ Motricité ✓	--	++	++	++ ++	--
Katsuragawa et al. (2019)	< RFID et capteurs de contact, position du contact entre le pouce et l'index >	Non / Non (n=12)	Indirection ✗ Intégration ✓ Compatibilité ✗	Perception ✓ Cognition ✓ Motricité ✓	--	++	++	++ ++	--
Wong et al. (2019)	< capteurs de contact, contacts sur les phalanges >	Non / Non (n=12)	Indirection ✗ Intégration ✓ Compatibilité ✗	Perception ✓ Cognition ✓/✗ Motricité ✓	--	++	++	++ ++	--
Xu et al. (2019)	< surface tactile, position du pouce sur l'index >	Non / Non (n=12)	Indirection ✗ Intégration ✓ Compatibilité ✗	Perception ✓ Cognition ✓ Motricité ✓	--	++	++	++ ++	--

FIGURE 1.58

1.8.1 Tactile + Micro-gestuel

Le paradigme Tactile combiné au Micro-Gestuel implique l'utilisation de gestes (par exemple, TAP, SWIPES) sur une surface tactile et des μ Gestes TTF présentés précédemment. Nous représentons ce paradigme comme utilisant les modalités d'interaction suivantes : <écran tactile, grammaire de gestes>+ <[dispositif], grammaire de gestes du pouce sur un autre doigt >.

Les μ Gestes TTF représentent une modalité prometteuse pour compléter le paradigme Tactile et augmenter l'expressivité. Par exemple, un appareil à écran tactile permettant uniquement 4 swipes directionnels, combiné à seulement deux μ Gestes TTF (par exemple, le pouce touchant soit l'index soit le majeur), pourrait fournir 12 types de swipes : les 4 swipes sans que le pouce ne touche aucun autre doigt, 4 swipes tandis que le pouce touche l'index, 4 swipes lorsqu'il touche le majeur.

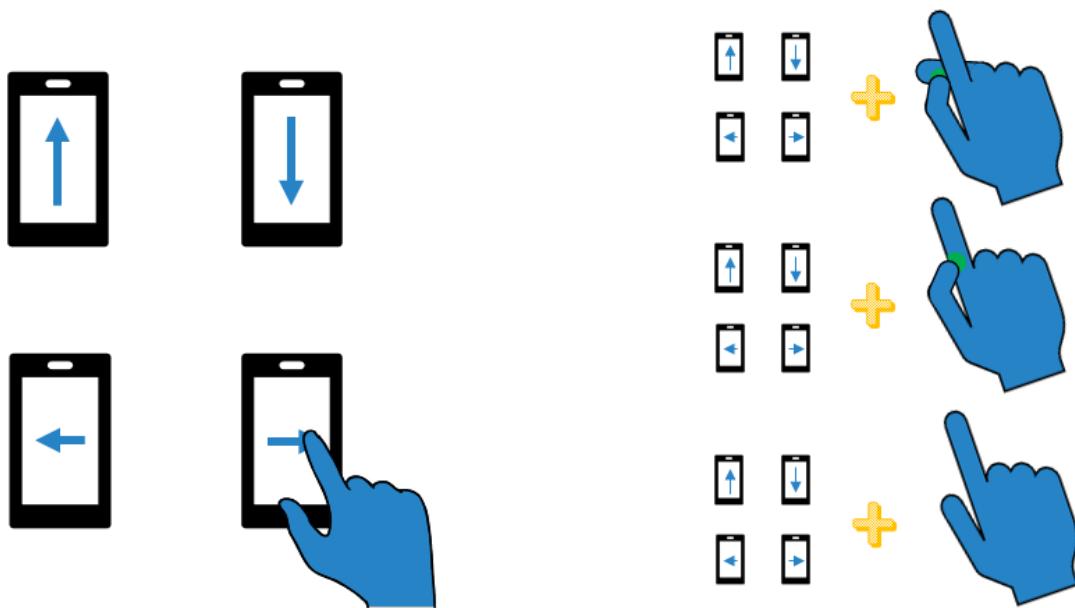


FIGURE 1.59 – Schéma illustrant l'utilisation des μ Gestes TTF pour augmenter l'expressivité.

Ainsi, plusieurs études ont utilisé des μ Gestes TTF pour augmenter l'expressivité du paradigme tactile [27, 220]. Par exemple, Tsai et al. [220] ont montré la faisabilité d'utiliser les

μ Gestes TTF en plus de l'écran tactile en situation de retour visuel sur téléphone portable dans deux scénarios : un contexte de navigation 3D et un contexte d'édition de texte. Les μ Gestes TTF sont alors utilisés pour basculer entre différents modes (par exemple, modes "écrire", "effacer" ou "surligner" dans le scénario d'édition de texte) (figure 1.60). Surale et al. [27] ont comparé six techniques pour basculer entre deux modes sur tablette en situation de retour visuel : 1) appui long, 2) main non-dominante, 3) deux doigts, 4) appui fort, 5) jointure des doigts, et 6) pouce-vers-doigt. Ils ont constaté que la technique pouce-vers-doigts (c'est-à-dire un TTF μ G) est la meilleure technique lorsque la précision tactile est requise. De plus, elle a été fortement appréciée par les participants, à la fois globalement et en termes de facilité d'utilisation.

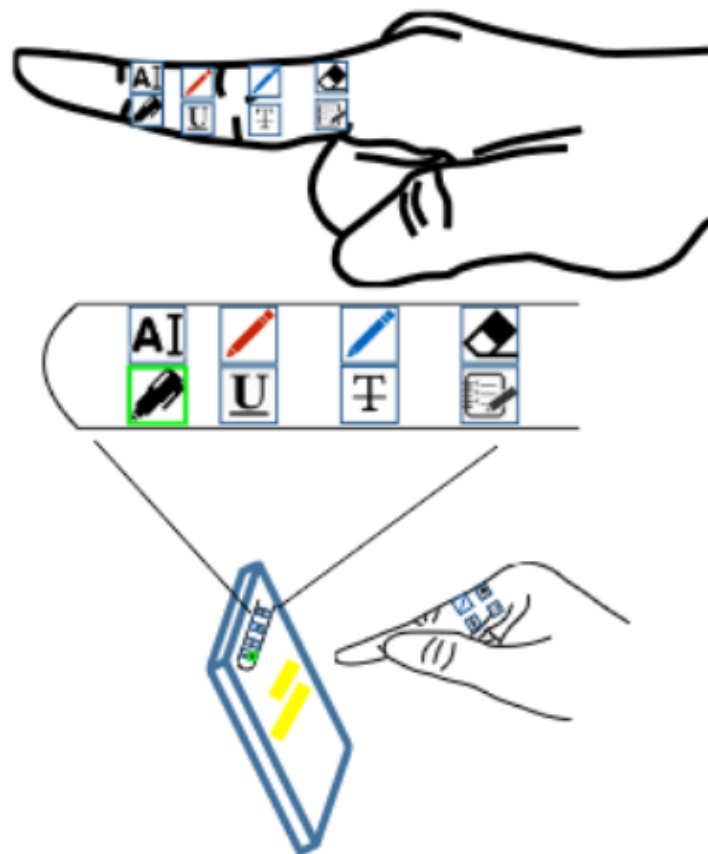


FIGURE 1.60 – Schéma illustrant le dispositif de Tsai et al. (2017). Illustration issue de [220].

Nous avons vu précédemment que les μ Gestes TTF étaient utilisables en parallèle d'une

tâche principale [89, 212, 31]. Dans une étude de 2019, Sharma et al. [203] ont montré que les μ Gestes TTF sont utilisables même en situation de main occupée et les deux études présentées dans le paragraphe précédent [27, 220] montrent que les μ Gestes TTF sont utilisables conjointement à un écran tactile et peuvent améliorer l'expressivité des modalités d'interaction basées sur le paradigme Tactile.

Ces études motivent notre travail sur la combinaison des μ Gestes TTF et des écrans tactiles pour les PDV. Nous avons vu en outre que les modalités d'interaction basées sur des gestes sont considérés dans la littérature comme naturelles [126, 238] et intuitives [155, 239, 148], qu'ils sont discrets, confidentiels et peu fatigants, et qu'ils respectent en majorité les principes de la manipulation directe des interfaces graphiques. Ainsi les modalités d'interaction basées sur les μ Gestes TTF semblent particulièrement prometteuses pour les PDV et c'est pourquoi nous choisissons de les étudier comme complément aux modalités du paradigme Tactile.

1.9 Synthèse

Les combinaisons de modalités d'interaction en entrée utilisables par les PDV sont multiples, cependant elles ne sont pas toutes efficaces. Selon l'objectif et les tâches d'interaction à effectuer pour y parvenir, un système peut-être plus ou moins adapté et accessible aux PDV.

Le paradigme Tactile est peu efficace en situation de non-retour visuel car sa bande-passante et son expressivité sont faibles - mais intégré dans les téléphones portables, il reste le paradigme le plus utilisé et est discret. Le paradigme Vocal est versatile, transportable, et particulièrement efficace pour les tâches de sélection directe et de saisie de texte - il manque toutefois de discrétion, exige de l'utilisateur la charge mentale de connaître et formuler ce qu'il veut faire, et n'est pas efficace pour les tâches simples et répétitives ou les tâches nécessitant un retour d'information continu. Utilisés conjointement, les paradigmes Vocal et Tactile souffrent des mêmes problèmes qu'utilisés de façon indépendante : leur multiplexage étant essentiellement temporel, leur bande-passante est très limitée, surtout lorsqu'il s'agit de tâches complexes.

Le paradigme Tangible est très efficace pour plusieurs types de tâches, cependant il nécessite des dispositifs spécifiques sur-mesure et souvent peu transportables, ce qui le rend peu compatible avec le paradigme Tactile. Si certains dispositifs ont été créés pour s'adapter à un téléphone portable ou une tablette par exemple, à minima, ils nécessitent d'utiliser des éléments supplémentaires (susceptibles d'être facilement perdus ou cassés) qui réduisent la praticité et la mobilité du téléphone portable utilisé seul. Ils sont en outre limités à des usages très spécifiques de par leur format et ne peuvent souvent pas être modifiés aisément. Étant donné que chaque application a une disposition et une utilité différente, cette solution nécessiterait l'utilisation de nombreux accessoires supplémentaires.

Le paradigme Gestuel - en particulier les gestes effectués sur le corps - est plutôt efficace, versatile et permet un usage multiplexé dans l'espace. Les modalités de ce paradigme offrent une bonne expressivité et bande-passante pour toutes les tâches d'interaction ainsi qu'un retour d'information continu. Le paradigme Gestuel peut être utilisé conjointement au paradigme Tactile mais pas toujours de façon simultanée. Par ailleurs, les gestes peuvent parfois être amples, fatigants et manquer de discrétion.

C'est pourquoi à ces gestes sur le corps nous faisons un focus sur le paradigme Micro-Gestuel : similaire au paradigme Gestuel, il est toutefois plus discret, plus confidentiel et moins fatiguant. Par ailleurs, les μ Gestes TTF sont utilisables conjointement et de façon simultanée avec le paradigme Tactile, et permettent de démultiplier l'expressivité et la bande-passante de ce dernier.

Les autres paradigmes ont leurs avantages mais nous estimons que le paradigme Micro-Gestuel est le plus pertinent à un usage en situation de non-retour visuel :

- contrairement au Vocal, il est discret et permet un usage multiplexé dans l'espace ;
- contrairement au Tangible, il ne repose pas sur des interfaces physiques externes, souvent encombrantes et peu versatiles ;
- similaire au Gestuel, il est toutefois plus discret et moins pénible de par son amplitude

plus réduite.

Les modalités d'interaction basées sur les μ Gestes TTF semblent particulièrement prometteuses pour les PDV et c'est pourquoi nous choisissons de les étudier comme complément au paradigme Tactile.

1.10 Conclusion

Dans ce chapitre nous avons introduit les principaux concepts liés aux paradigmes d'interaction en entrée dans un contexte d'usage par des PDV ayant une cécité totale et avons défini la terminologie adoptée dans nos travaux.

Nous avons étudié les modalités d'interaction existantes pour un contexte d'usage par des PDV que nous regroupons en 4 paradigmes d'interaction. Nous avons analysé ces modalités d'interaction selon une grille d'analyse construite à partir de modèles en IHM [75, 119, 40, 24] ... Cette analyse nous permet d'extraire de chaque paradigme traité un certain nombre de caractéristiques afin de déterminer leur adéquation avec un usage sans la vue, dans le but d'appliquer la plus adéquate à une utilisation sur téléphone portable, par des PDV, en complément du paradigme Tactile.

Suite à cette analyse, nous concluons que le paradigme Micro-Gestuel offre une potentialité intéressante à explorer : il s'agit de notre objet de recherche, celui que nous traitons dans les chapitres suivants et dans lequel nous exposons notre espace solution.

Nous n'avons pas connaissance d'étude réalisée évaluant et comparant l'utilisabilité des μ Gestes TTF les plus courants dans la littérature afin d'être utilisés en conjonction avec le paradigme Tactile en situation de non retour visuel. C'est pourquoi, dans le chapitre suivant (chapitre 2), nous présentons une première étude sur les μ Gestes TTF qui peuvent être utilisés sans retour visuel lorsque l'index est en contact avec un écran tactile.

Après avoir identifié l'ensemble des μ Gestes TTF possibles, nous présentons dans le chapitre 3 trois études portant sur l'utilité et l'utilisabilité de ces μ Gestes TTF dans plusieurs applications différentes aux regard des tâches considérées par Foley [75] : une exploration tactile de document, une tâche de copier/coller et une tâche de sélection de texte.

Chapitre 2

Faisabilité des microgestes

Thumb-To-Finger en complément du touché sur un écran tactile

Dans le chapitre précédent, nous avons étudié différents paradigmes d'interaction pouvant être utilisés en complément de l'écran tactile dans une situation de non-retour visuel. Notre analyse des travaux existants a montré que le paradigme Micro-Gestuel était particulièrement prometteur pour un usage multimodal en complément de l'écran tactile et sans retour visuel.

Dans ce chapitre, première étape de notre démarche, nous présentons nos travaux sur la faisabilité d'une interface multimodale utilisant conjointement l'écran tactile d'une tablette avec les microgestes (μ Gestes) Thumb-To-Finger (TTF) en l'absence de retour visuel. Ce chapitre reprend et présente des éléments issus d'un article publié et présenté dans la conférence ACM-ICMI 2022 : *"Keep in Touch : Combining Touch Interaction with Thumb-to-Finger μ Gestures for People with Visual Impairment"* [71].

2.1 Question de recherche

Nous avons vu dans le chapitre précédent que les microgestes (μ Gestes), de petits mouvements réalisés intentionnellement comme commandes d'entrée – et plus spécifiquement les μ Geste réalisés avec le pouce sur d'autres doigts de la même main (thumb-to-finger, TTF) – possèdent de nombreuses propriétés : ils offrent une bonne expressivité pour toutes les tâches d'interaction, peuvent être utilisés conjointement au paradigme Tactile de façon simultanée et sans retour visuel, sont discrets, peu fatiguants et impactent peu la charge de travail.

Les μ Gestes TTF offrent en effet une grande expressivité : de nombreux gestes sont possibles (e.g., appuyer, glisser, dessiner) et peuvent être effectués avec précision sur différentes parties des doigts en exploitant la dextérité de la main et la proprioception [178, 94, 246, 112], permettant un multiplexage spatial dense. La plupart des modalités d'interaction basées sur les μ Gestes TTF présentées au chapitre précédent permettent d'accéder à plus d'une dizaine d'objets d'intérêt à un instant donné [219] (voir figure 1.60).

De plus, les μ Gestes TTF peuvent être utilisés simultanément à un écran tactile, décuplant ainsi les possibilités quant aux commandes d'entrées de celui-ci [27, 220].

Enfin, la proprioception et la perception tactile au niveau des mains sont suffisamment précises pour permettre une utilisation des μ Gestes TTF en situation de retour visuel partiel, voire de non-retour visuel total [178, 94, 246, 112].

Ces trois propriétés notamment font du paradigme Micro-Gestuel un excellent candidat pour augmenter l'expressivité des PDV sur les dispositifs utilisant un écran tactile. Comme expliqué dans le chapitre précédent, les μ Gestes TTF ont été largement étudiés dans de nombreuses situations (e.g., en course d'athlétisme, en vélo, en voiture, dans le cockpit d'avion [89, 49, 240, 212, 31, 230], mais n'ont à notre connaissance que été que très peu étudiées avec les PDV, malgré les opportunités qu'elles pourraient offrir à ce public. En partie parce qu'ils permettent d'effectuer des μ Gestes TTF, Feng [73] a constaté dans un article de 2016

2.1. QUESTION DE RECHERCHE

que les contrôleurs montés sur la main (e.g., une bague) étaient préférés par les participants ayant une Déficience Visuelle (DV) à d'autres types de contrôleurs tels que ceux montés sur la hanche ou sur la tête (e.g., une ceinture, des lunettes). Pour leur part, Nanayakkara et al. [150] présentent en 2012 EyeRing, un dispositif porté au doigt utilisé dans de nombreuses applications pour les personnes déficientes visuelles, telles que la reconnaissance de billets de banque et la navigation. L'appareil en forme de bague est doté d'une caméra et d'un bouton sur le côté permettant d'activer le système à l'aide du pouce – un μ Geste TTF. Boldu et al. [32] ont proposé AiSee en 2020, un système de reconnaissance d'objet par caméra activé de façon similaire, à l'aide d'un bouton situé sur une bague. Dans cet article, les auteurs précisent qu'ils ont choisi cette technique d'interaction pour ses nombreux avantages :

“Pour maintenir l'interaction aussi discrète et peu intrusive que possible, nous avons intégré un interrupteur subtil dans une bague portée à l'index, qui déclencherait le processus de reconnaissance. Appuyer simplement sur l'interrupteur à l'aide du pouce constitue un microgeste minimal qui permet potentiellement une interaction réflexe et semble donc être une technique d'entrée optimale. Ce geste du pouce à l'index [...] peut être effectué avec peu d'effort en parallèle à d'autres tâches du monde réel.”

Ces études se contentent toutefois d'utiliser les μ Gestes TTF (souvent même, un seul) et non de les étudier. Pour ce que nous en savons, les études portant sur l'utilisation conjointe des μ Gestes TTF et de l'écran tactile pour en augmenter l'expressivité et la bande-passante en situation de non-retour visuel sont inexistantes. En effet, nous avons montré dans le chapitre précédent (Chapitre 1) qu'une utilisation conjointe – et notamment simultanée – du paradigme Tactile et Micro-Gestuel pourrait démultiplier le nombre de commandes possibles [27, 220].

Se pose alors la question d'étudier ces μ Gestes TTF en situation et de façon exhaustive : sont-ils faisables ? Sont-ils efficaces ?

Pour répondre à ces questions et avant de traiter de leur utilité et utilisabilité pour des tâches données, dans un premier temps, nous proposons d'évaluer la faisabilité des μ Gestes TTF en

2.1. QUESTION DE RECHERCHE



FIGURE 2.1 – **Gauche** : Photographie du dispositif EyeRing de Nanayakkara et al. (2012), issue de [150]. **Droite** : Photographie du dispositif AiSee de Boldu et al. (2020), issue de [32].

situation de non-retour visuel lorsque l'index est en contact avec l'écran tactile (figure 2.2). Dans le reste de ce chapitre, nous présentons une étude expérimentale destinée à évaluer systématiquement 33 μ Gestes parmi les plus courants de la littérature.

Cette étude expérimentale est inspirée d'études précédentes sur des μ Geste TTF effectués en l'air [94, 219, 101].



FIGURE 2.2 – Un TAP effectué sur l'index lorsque celui-ci est en contact avec l'écran tactile.

2.2 Étude expérimentale

Nous avons exposé dans la section précédente notre problématique concernant le degré de faisabilité des μ Gestes TTF lorsque l'utilisateur manipule un écran tactile en situation de non-retour visuel. Cette section introduit l'étude expérimentale nous permettant d'explorer cette faisabilité et dont l'objectif principal est d'identifier les meilleurs μ Gestes TTF dans cette situation, en termes de taux de réussite et de confort pour l'utilisateur, afin de fournir des lignes directrices de conception pour l'interaction sur les écrans tactiles lorsque l'index ou d'autres doigts restent en contact avec la surface du dispositif.

2.2.1 Protocole expérimental

Notre étude examine quels μ Gestes TTF peuvent être utilisés lorsque l'index est en contact avec un écran tactile. Pour étudier quels μ Gestes TTF sont les plus adaptés, les participants ont été invités à effectuer 33 μ Gestes TTF avec leur main gauche sans la regarder et en maintenant l'index gauche en contact avec l'écran d'une tablette posée sur une table devant eux (figure 2.2).

Neuf volontaires voyants (dont trois femmes) âgés de 20 à 34 ans (moyenne de 25,9, écart-type de 4,6) ont participé à cette expérience. Tous les participants (à l'exception d'un) sont droitiers et trois jouent d'un instrument. Le participant gaucher utilise des ordinateurs avec une souris et joue d'un instrument (guitare et basse) en tant que personne droitère. Aucun des participants n'a signalé de handicaps physiques, mentaux ou perceptuels. Tous utilisent quotidiennement un appareil tactile.

Comme l'étude se concentre sur l'aspect anatomique et le contrôle moteur des μ Gestes TTF, nous n'avons pas recouru à des participants PDV pour cette étude. En effet, la capacité à réaliser des μ Gestes TTF et le confort perçu sont liés à la biomécanique de la main plutôt qu'au degré de handicap visuel des participants. Notre raisonnement est basé sur les travaux de Palani et al. [168] : dans trois expériences, ils n'ont trouvé aucune différence entre les participants DV et les participants voyants ayant les yeux bandés dans la capacité à effectuer des

2.2. ÉTUDE EXPÉRIMENTALE

tâches perceptuelles kinésio-tactiles. De plus, dans une revue systématique sur les technologies destinées aux PDV, Brulé et al. [39] ont constaté que 178 études ont testé des techniques d'interaction "au niveau des actions" avec des participants voyants, ce qui montre que cela est généralement approprié. Enfin, ces auteurs ont mentionné que, lorsque c'est possible, s'appuyer sur des volontaires voyants réduit les contraintes liées à la participation des PDV aux expériences (telles que les problèmes de mobilité et disponibilité des volontaires). En effet, le problème d'accès à certains participants est bien connu dans la recherche universitaire, et découle d'une disponibilité limitée de ceux-ci : par exemple, sur 144 études en Interaction Humain-Machine (IHM) incluant des participants DV, le nombre moyen de participants est de seulement 9,7.

Étant donné que les PDV utilisent principalement leur main gauche lorsqu'elles explorent des graphiques [22], les participants sont invités à effectuer des μ Gestes TTF avec cette main, sans la regarder et avec l'index touchant l'écran de la tablette.

En nous basant sur des études antérieures évaluant des μ Gestes TTF dans des conditions différentes [94, 101, 219], nous avons sélectionné les μ Gestes TTF "TAP" (c'est à dire un appui court, comme un clic) et "SWIPE" (c'est à dire un mouvement glissant du pouce le long du doigt), car ils sont préférés par les utilisateurs [49]. Nous avons choisi les parties des doigts sur lesquelles effectuer les μ Gestes TTF au regard de cette même littérature : les phalanges [94, 101] et les ongles [110] car ceux-ci offrent des repères notables et accessibles sans retour visuel. Contrairement à Huang et al. [94], nous avons choisi de ne pas subdiviser les phalanges en zones plus petites car nous estimons que la position contrainte de la main et du doigt touchant l'écran tactile rendrait la tâche trop difficile.

Dans notre configuration expérimentale, l'index est en contact avec la surface, de sorte que la phalange supérieure et l'ongle de l'index ne sont pas accessibles avec le pouce. La phalange supérieure et l'ongle de l'index ont donc été exclues de l'ensemble, qui comprend un total de 33 μ Gestes TTF :

2.2. ÉTUDE EXPÉRIMENTALE

- 11 TAP (phalanges supérieure, moyenne, inférieure × index, majeur, annulaire, auriculaire – en excluant la phalange supérieure de l'index) ;
- 11 DOUBLE TAP (les mêmes zones) ;
- 3 TAP sur les ongles (majeur, annulaire, auriculaire) ;
- 8 SWIPE (du haut vers le bas, du bas vers le haut × index, majeur, annulaire, auriculaire).



FIGURE 2.3 – Les zones sur lesquels les gestes sont effectués.

L'expérience évalue et compare quels μ Gestes TTF sont réalisables, plus précis et plus confortables pour l'utilisateur dans notre contexte particulier. Par conséquent, la tâche consiste à effectuer l'ensemble des 33 μ Gestes TTF affichés à l'écran aussi rapidement et précisément que possible.

Nous avons installé les participants et les avons familiarisés avec l'ensemble de l'appareillage, puis nous leur avons donné les instructions. Nous avons insisté sur le fait qu'ils devaient "le faire aussi rapidement que possible, sans compromettre la précision".

Pour se familiariser avec la tâche, les participants ont suivi une phase d'entraînement en situation de retour visuel. Ils ont effectué chacun des 33 μ Gestes TTF dans l'ordre contrôlé

2.2. ÉTUDE EXPÉRIMENTALE

suivant :

- les TAP sur les phalanges puis les ongles, puis ;
- les DOUBLE TAP sur les phalanges, puis ;
- les SWIPE d'avant en arrière sur chaque doigt.

Ils devaient effectuer correctement chaque μ Geste TTF pour passer au μ Geste TTF suivant. Le plan expérimental est de type "intra-sujets" et "à mesures répétées". Les participants ont donc répété la tâche dans une série de huit blocs. Chaque bloc comprenait les 33 μ Gestes TTF présentés dans un ordre pseudo-aléatoire. Pour chaque essai (correspondant à l'exécution d'un μ Geste TTF), les participants étaient invités à effectuer le μ Geste TTF sur la zone de leur main affichée à l'écran (figure 2.4) avec la main gauche. Ils devaient commencer chaque essai avec leurs deux index sur la tablette. Ils pouvaient les placer où ils le souhaitaient tant qu'ils les maintenaient en contact avec la surface de la tablette pendant toute la durée de l'essai. Il était possible de les déplacer pendant l'essai, mais nous leur avons donné pour consigne de bouger le moins possible. Un essai se poursuivait jusqu'à ce que le participant effectue le μ Geste TTF attendu. Toutes les tentatives de μ Geste TTF ont été enregistrées. Les participants pouvaient prendre une pause chaque fois qu'ils le souhaitaient entre les essais.

Après l'expérience, les participants ont répondu à un questionnaire démographique, puis ils devaient évaluer le confort physique de chaque μ Geste TTF sur une échelle de Likert à 5 points, allant de 1 ("très inconfortable") à 5 ("très confortable").

2.2.2 Implémentation

Les participants sont assis devant une tablette de 10,1 pouces (Huawei MediaPad T5 Lite, 24*15,5 cm), posée sur une table devant eux. Un écran de 17 pouces est placé à hauteur des yeux devant eux pour afficher les instructions et les retours d'information (figure 2.4).

Pour capturer les μ Gestes TTF effectués, certaines techniques sont basées sur des caméras, cependant, elles souffrent énormément de l'occultation (c'est-à-dire que les articulations

2.2. ÉTUDE EXPÉRIMENTALE



FIGURE 2.4 – Exemples d’instruction affich e durant l’exp erience. **Gauche** : la zone   atteindre en vert repr esente un Tap sur la phalange basse de l’index. **Milieu** : la zone   atteindre en vert avec l’indication "x2" en rouge repr esente un Double Tap sur la phalange basse de l’index. **Droite** : la fl che rouge allant de la paume vers l’extr mit  du doigt repr esente un Swipe Bas-Haut sur l’index.

et les segments de doigts sont cach s   la vue de la cam ra) et n ecessitent un calibrage [178, 208]. D’autres approches bas es sur des capteurs   distance ont  t  con ues, notamment la tomographie [251], le radar [232], les IMU [219], le magn tom tre [49] et une combinaison d’un microphone et d’un gyroscope [248]. Ces m ethodes sont prometteuses pour la reconnaissance non intrusive des μ Gestes TTF, mais le nombre de gestes reconnus est encore limit . D’autres approches reposent sur des capteurs de contact [234, 235, 110]. Des approches plus facilement reproductibles sont bas es sur des gants avec des capteurs tactiles [101, 127, 219, 242] et c’est la solution que nous avons retenue pour l’ tude exp erimentale.

Notre prototype est un gant en tissu  lastique sur lequel sont brod s cinq connecteurs en fil conducteur par doigt (seulement quatre pour l’index), plus un connecteur en tissu conducteur sur l’ongle des trois derniers doigts. Chaque connecteur est reli    une entr e d’une carte Arduino Uno¹. Une zone en tissu conducteur reli e   la masse est plac e sur le pouce et permet de fermer le circuit lorsqu’elle se trouve en contact avec un ou plusieurs connecteurs.

A chaque  v nement, le programme de la carte Arduino envoie   l’ordinateur une liste de bool ens (0 et 1), chacun correspondant   l’ tat d’un connecteur (absence ou non de contact).

Des logiciels personnalis s ont  t  d velopp s pour l’exp erience. Un en Java pour la tablette, qui est utilis  pour suivre les positions des index et les envoyer   l’ordinateur via Wi-Fi.

1. Arduino Uno - site officiel : <https://www.arduino.cc/en/Guide/ArduinoUno>



FIGURE 2.5 – Photographique du premier prototype.

Un en Python pour l'ordinateur, qui reçoit et traite les données provenant de la tablette et du gant, et enregistre aussi les données expérimentales.

2.2.3 Résultats

Nous avons opté pour une conception intra-sujets avec mesures répétées, enregistrant ainsi un total de $9 \text{ participants} \times 33 \mu\text{Gestes TTF} \times 8 \text{ blocs} = 2376 \mu\text{Gestes}$. Nous écartons les valeurs aberrantes, que nous définissons comme les essais où les participants font plus de 5 tentatives erronées, en suivant la règle de distribution à 95%. Au total, 140 essais sur 2376 ont été retirés (soit 5,9% des données), aboutissant à 2236 essais. Nous utilisons une approche d'estimation [63] pour communiquer les résultats plutôt que les tests statistiques basés sur l'hy-

2.2. ÉTUDE EXPÉRIMENTALE

pothèse nulle.

Afin de déterminer quels μ Gestes TTF sont les plus adaptés, nous nous intéressons principalement à trois données : le taux d'erreur, le confort perçu et la dérive.

Le taux d'erreur est calculé comme la somme des erreurs divisée par le nombre total d'essais. Une erreur est comptabilisée lorsque la première tentative d'un essai est un échec – dit autrement, un essai qui n'est pas réussi du premier coup est compté comme une erreur. Pour rappel, durant l'expérience un essai se poursuit jusqu'à la réussite du μ Geste TTF attendu.

Le confort perçu est une évaluation subjective numérotée entre 1 (« très inconfortable ») et 5 (« très confortable »). Il s'agit du résultat par μ Geste TTF du questionnaire proposé aux participants à la fin de l'expérience.

La dérive correspond à l'écart de position de l'index sur la tablette lors de l'exécution du geste. Elle est calculée en mesurant la variation des coordonnées XY de l'index sur la tablette du moment où le stimulus apparaît jusqu'à la fin de l'essai.

Taux d'erreur : la figure 2.6 présente les taux d'erreur moyens et les intervalles de confiance à 95% (IC) pour les sessions d'entraînement et de l'expérience (tous les blocs, blocs 1 à 3 et blocs 4 à 8). De par la nouveauté des μ Gestes, nous nous attendions à ce que les participants s'améliorent après quelques blocs, c'est pourquoi nous avons considéré les trois premiers blocs comme une période où les participants pouvaient éventuellement s'améliorer avant que les participants n'atteignent un plateau de performance.

Conformément aux recommandations de Dragicevic [63], la zone en surbrillance et les encadrés en pointillés montrent que les intervalles de confiance des trois premiers blocs et des cinq derniers blocs ne se chevauchent pas, suggérant fortement un effet d'apprentissage. La figure 2.7 présente les taux d'erreur pour chaque type de μ Geste TTF (TAP, DOUBLE TAP, SWIPE), d'abord globalement, puis par doigt.

2.2. ÉTUDE EXPÉRIMENTALE

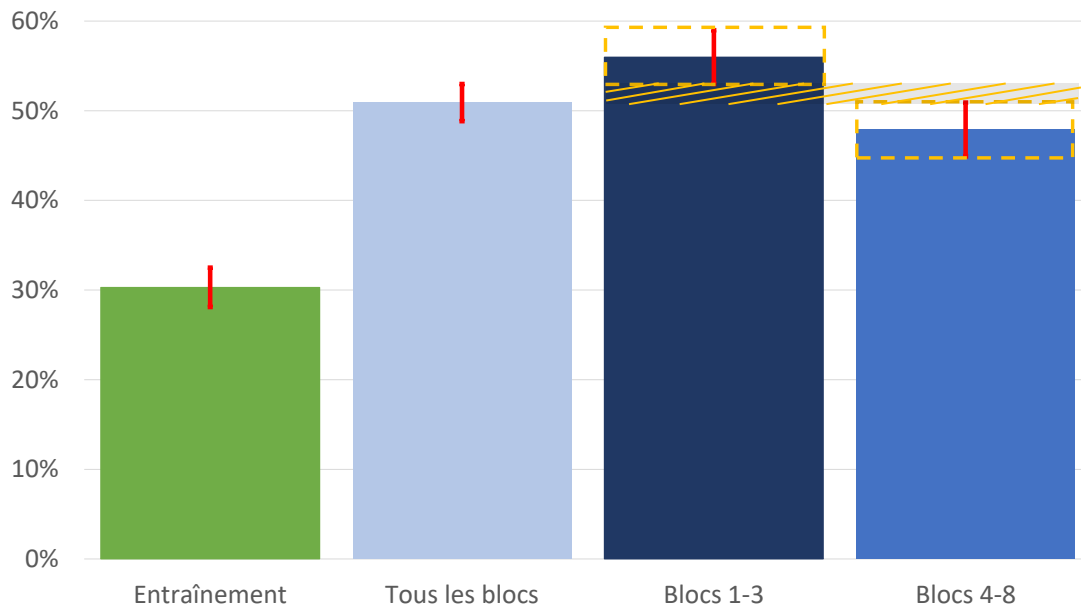


FIGURE 2.6 – Graphique illustrant le taux d'erreur pour la phase d'entraînement (colonne de gauche) et les phases de l'expérimentation (toutes les autres colonnes). Les zones en surbrillance et les encadrés en pointillés (en jaune) montrent que les intervalles de confiance des trois premiers blocs et des cinq derniers blocs ne se chevauchent pas.

Les intervalles de confiance des SWIPE effectués sur l'index et le majeur ne se chevauchent pas avec ceux des annulaires et des auriculaires. Cela suggère fortement que les SWIPE effectués sur l'index ou le majeur sont moins sujets aux erreurs que ceux réalisés sur les annulaires et les auriculaires (encadré en pointillé jaune à la figure 2.7). Les intervalles de confiance des TAP effectués sur l'index, le majeur et l'annulaire se chevauchent à peine avec ceux des DOUBLE TAP réalisés sur les mêmes doigts (encadré en pointillé orange à la figure 2.7). Cela suggère que les TAP effectués sur les phalanges de ces trois doigts sont significativement moins sujets aux erreurs que les DOUBLE TAP effectués sur les mêmes zones. L'intervalle de confiance du TAP effectué sur les ongles ne se chevauche avec rien (encadré en pointillé vert à la figure 2.7). Cela suggère fortement que les TAP sur les ongles sont significativement moins sujets aux erreurs que n'importe quel autre μ Geste TTF. La figure 2.7 montre également que le DOUBLE TAP est le μ Geste TTF le plus sujet aux erreurs : les participants ont échoué à le réaliser dès la première tentative dans 67,9% des cas. En revanche, les μ Geste TTF TAP

2.2. ÉTUDE EXPÉRIMENTALE

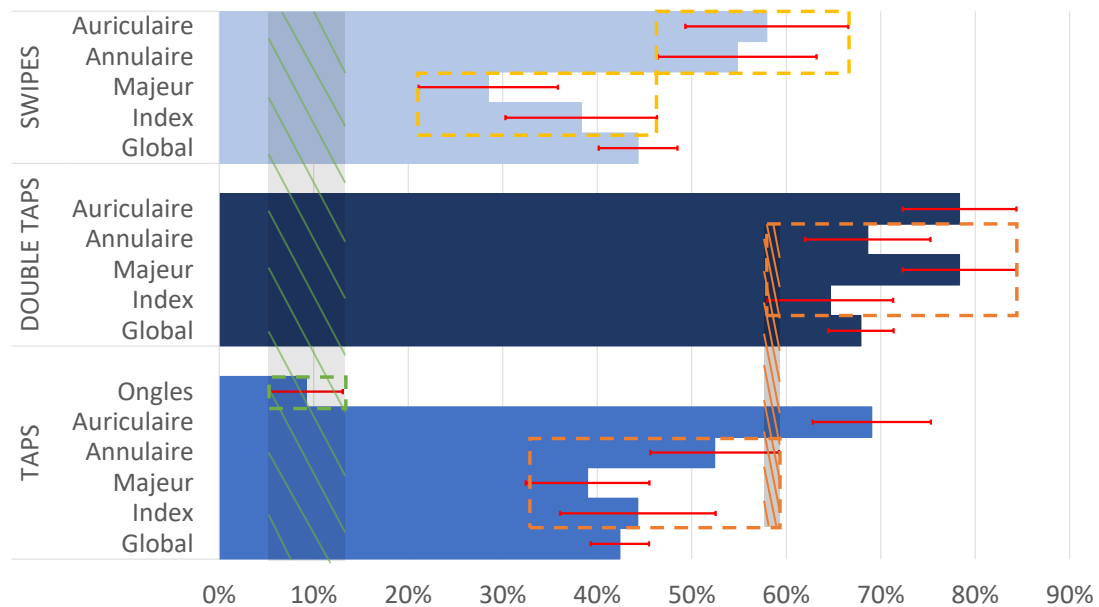


FIGURE 2.7 – Graphique illustrant le taux d’erreur pour chaque μ Geste TTF, globalement et pour chaque doigt. Les encadrés en pointillés (en jaune) montrent qu’il n’y a pas de chevauchement des intervalles de confiance entre le SWIPE de l’index et du majeur et le SWIPE de l’annulaire et de l’auriculaire. La zone en surbrillance et les encadrés en pointillés (en orange) montrent qu’il n’y a presque aucun chevauchement des intervalles de confiance entre le TAP et le DOUBLE TAP de l’index, du majeur et de l’annulaire. La zone en surbrillance et l’encadré en pointillés à gauche (en vert) montre qu’il n’y a pas de chevauchement des intervalles de confiance entre le TAP sur les ongles et le reste des μ Gestes TTF.

sont moins sujets aux erreurs avec seulement 42,4% de taux d’erreur global. Dans la figure 2.7), nous montrons séparément les TAP réalisés sur les ongles (étiquetés "Ongles") des TAP réalisés sur les phalanges des doigts (étiquetés par doigt), car ils sont nettement moins sujets aux erreurs.

Confort : la figure 2.8 présente la note moyenne de confort, comprise entre 1 (plus basse) et 5 (plus élevée), attribuée par les participants lors du questionnaire rempli après l’expérience, pour chaque μ Geste TTF, doigt et phalange cible. La zone en surbrillance et l’encadré en pointillés (en jaune) montrent un faible chevauchement des intervalles de confiance entre les TAP sur les ongles et le reste, suggérant fortement que les ongles sont perçus comme significativement plus confortables que les autres zones de la main pour réaliser des μ Gestes TTF. La zone en surbrillance et l’encadré en pointillés (en vert) montrent une absence de chevauche-

2.2. ÉTUDE EXPÉRIMENTALE

ment des intervalles de confiance entre les μ Gestes TTF de l'auriculaire et les autres doigts, celui-ci ayant un score beaucoup plus faible (2,4) que tous les autres. Cela montre que l'auriculaire est dans l'ensemble une zone peu appréciée pour réaliser des μ Gestes TTF. La zone en surbrillance et l'encadré en pointillés (en orange) montrent un faible chevauchement des intervalles de confiance entre les μ Gestes TTF effectués sur l'index et le majeur, et ceux effectués sur l'annulaire. Ces résultats montrent que les préférences des utilisateurs se portent vers les μ Gestes TTF effectués sur les doigts proches du pouce (c'est-à-dire l'index et le majeur) et sur les phalanges supérieures (c'est-à-dire les phalanges les plus éloignées de la paume) lorsque l'index est en contact avec une surface. Ceci est conforme à la littérature [94, 101, 219] et nous montrons ici que ces résultats existants sont confirmés lorsque l'index est en contact avec une surface.

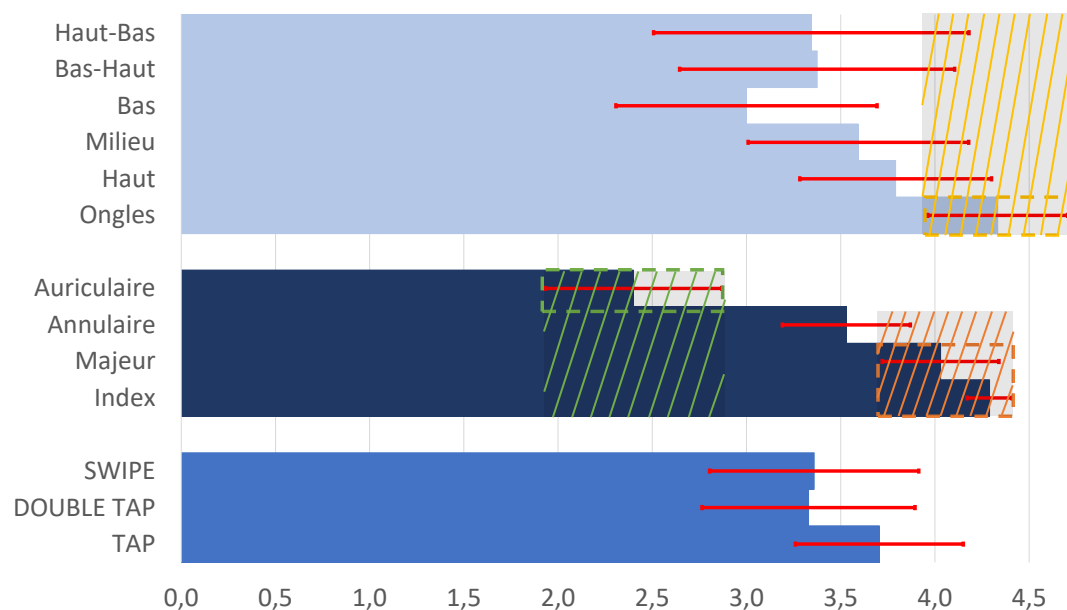


FIGURE 2.8 – Graphique illustrant la note de confort et les intervalles de confiance triés par μ Gestes TTF (en bas), doigt ciblé (au milieu) et zone ciblée (en haut). La zone en surbrillance et l'encadré en pointillés (en jaune) montrent qu'il y a très peu de chevauchement des intervalles de confiance entre le TAP sur les Ongles et le reste. La zone en surbrillance et l'encadré en pointillés (en vert) montrent qu'il n'y a pas de chevauchement des intervalles de confiance entre les μ Gestes TTF sur le petit doigt et le reste. La surbrillance et la boîte pointillée (en orange) montrent qu'il y a presque aucun chevauchement des intervalles de confiance entre les μ Gestes TTF sur l'index et le majeur d'une part et celui de l'annulaire d'autre part.

2.2. ÉTUDE EXPÉRIMENTALE

Dérive : nous rappelons que la dérive est la déviation de l'index de la main gauche sur la tablette pendant l'exécution du geste. Elle est calculée à l'aide de la variation des coordonnées XY de l'index sur la tablette depuis le moment où le stimulus est apparu jusqu'à la fin de l'essai (paragraphe 2.2.3). L'ellipse montre l'écart type autour du point de départ (figure 2.9). En moyenne, la dérive de l'index gauche a une ellipse avec un rayon x de 2,10 cm et un rayon y de 1,28 cm. La figure 2.9 montre l'écart type de la dérive globale, puis par type de μ Geste TTF et enfin par doigt tous types de μ Gestes TTF confondus.

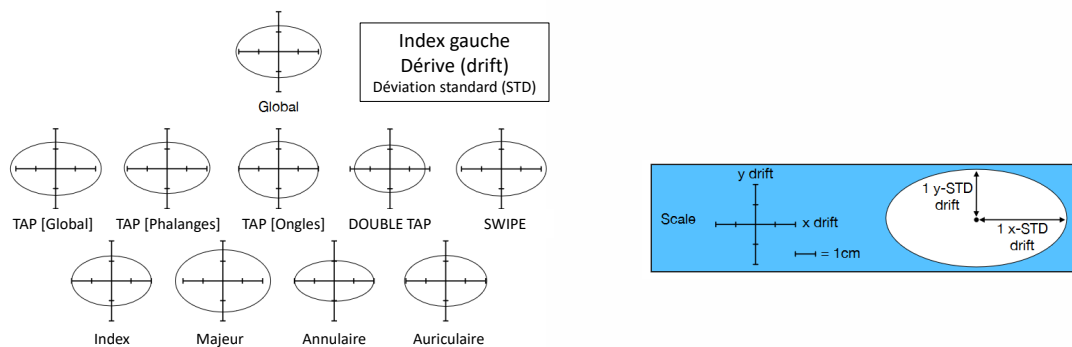


FIGURE 2.9 – **Gauche** : Dérive de l'index gauche : dérive globale, dérive par microgeste et par doigt cible. **Droite** : Schéma représentant l'échelle et expliquant comment nous avons calculé la dérive.

2.2.4 Analyse

Nos résultats démontrent que chacun des 33 μ Gestes TTF est physiquement réalisable même lorsque l'index est en contact avec une surface. Dans ce contexte, les préférences des utilisateurs se tournent vers les μ Gestes TTF réalisés sur les doigts proches du pouce (c'est-à-dire l'index et le majeur) ainsi que sur les phalanges supérieures (c'est-à-dire les phalanges les plus éloignées de la paume). Les TAP sur les ongles sont particulièrement populaires parmi les participants. Ces préférences concordent avec nos résultats expérimentaux de performance : les TAP sont les μ Gestes TTF les moins sujets aux erreurs (42,4%), suivis par les SWIPE (44,3%), et ils sont les plus utilisables (moins d'erreurs observées) lorsqu'ils sont effectués sur l'index et le majeur. Les ongles sont particulièrement utilisables, avec le taux d'erreur le plus faible parmi tous les μ Gestes TTF (9,2%). En revanche, effectuer des μ Gestes TTF sur l'annu-

2.2. ÉTUDE EXPÉRIMENTALE

laire et l'auriculaire est plus difficile que sur les autres doigts et est moins apprécié des participants. Les DOUBLE TAP entraînent significativement plus d'erreurs que les autres μ Gestes TTF (67,9%).

Notre objectif est de fournir des lignes directrices et un ensemble de μ Gestes TTF adaptés à une utilisation lorsque l'index touche l'écran en situation de non retour visuel. Nous cherchons donc à identifier l'ensemble de μ Gestes TTF présentant le taux d'erreur le plus faible et le meilleur niveau de confort : pour cela, nous affinons l'ensemble initial de 33 μ Gestes TTF en un ensemble plus utilisable en réduisant les facteurs d'erreur (par exemple, la nécessité d'un ciblage précis) et en supprimant les μ Gestes TTF les plus problématiques. Ce processus d'affinement est caractérisé par un compromis entre la résilience (c'est-à-dire un faible taux d'erreur) et la taille de l'ensemble des μ Gestes TTF considérés comme utilisables. Nous avons choisi d'ignorer la Dérive, car les données n'offrent pas de résultats permettant de discriminer les μ Gestes TTF. Notre analyse est basée sur une typologie des erreurs qu'un participant peut commettre lors d'un essai :

- Erreur de geste, le mauvais μ Geste TTF est effectué ;
- Erreur de doigt, le μ Geste TTF est effectué sur le mauvais doigt ;
- Erreur de zone, le μ Geste TTF est effectué sur le bon doigt mais sur la mauvaise partie du doigt.

Si au moins l'une de ces erreurs se produit avant la fin de l'essai, une occurrence d'erreur est comptée. Ces erreurs ne sont pas mutuellement exclusives. Un participant peut commettre plusieurs erreurs lors de l'exécution d'un seul μ Geste TTF. Par exemple, si un participant est invité à effectuer un TAP sur la zone de base du majeur mais effectue un DOUBLE TAP sur la zone supérieure du majeur, il y a une erreur de geste et une erreur de zone.

Sur 2249 essais, les erreurs de geste représentent un taux d'erreur de 30,3% (677 occurrences, CI 4,3%). Les erreurs de doigt représentent un taux d'erreur de 12,5% (297 occurrences, CI 2,0%). Les erreurs de zone représentent un taux d'erreur de 34,2% (765 occur-

2.2. ÉTUDE EXPÉRIMENTALE

rences, CI 3,7%).

2.2.5 Optimisation des μ Gestes TTF : TAP et DOUBLE TAP

Nous optimisons conjointement les TAP et les DOUBLE TAP, car il s'agit du même type d'entrée. Étant donné que les μ Gestes de type TAP sur les ongles sont – de loin – les moins sujets aux erreurs (figure 2.7) et les plus confortables (figure 2.8), nous avons décidé de les conserver tels quels dans notre ensemble de μ Gestes TTF.

		# Essais	Occurrence d'erreur	Taux d'erreur	CI	
Tous les blocs : TAP & DOUBLE TAP seulement	Global	1473	876	59,5%	±2,5%	
	Type d'erreur	Geste	1473	468	31,8%	±5,2%
		Doigt		197	13,4%	±2,6%
		Zone		699	47,4%	±4,9%
	Doigt cible	Index	278	142	51,1%	±5,9%
		Majeur	411	211	45,1%	±4,8%
		Annulaire	394	237	52,1%	±4,8%
		Auriculaire	390	286	63,3%	±4,4%

FIGURE 2.10 – Tableau récapitulatif par type d'erreur et par doigts pour les TAP et 2-TAP uniquement.

Lorsque nous prenons en compte uniquement les TAP et les DOUBLE TAP sur les phalanges (c'est-à-dire sans les TAP sur les ongles), le Tableau 2.10 montre un taux d'erreur de 59,5%. Il montre également que les erreurs liées à la zone ciblée sont la principale source d'erreurs (Erreurs de zone 47,4%). Le contexte de l'expérience pourrait être une explication : en raison de la position contrainte de la main et de la situation sans vision, les participants ont probablement des difficultés à cibler précisément une phalange. Une solution consiste à augmenter la taille de la cible en ne considérant qu'une zone sur chaque doigt : nous considérons que réa-

2.2. ÉTUDE EXPÉRIMENTALE

liser la tâche sur le doigt requis – indépendamment de la zone requise, à l'exception des ongles – est suffisant pour valider l'essai. Nous fusionnons donc les zones situées sur les phalanges en une seule et unique zone englobant tout le doigt en tant que première étape d'affinement. Dans cette configuration, le taux d'erreur global pour les TAP et les DOUBLE TAP passe de 59,5% à 41,1%, comme le montre la figure 2.11. Les TAP et DOUBLE TAP sont réduits de 25 (11 TAP et 11 DOUBLE TAP + 3 TAP sur les ongles) à 11 (4 TAP et 4 DOUBLE TAP, un par doigt + 3 TAP sur les ongles).

Les μ Gestes TTF TAP réalisés sur les ongles sont comparativement moins sujets aux erreurs (9,2%), comme l'indique la figure 2.6, ce qui justifie de les traiter séparément.

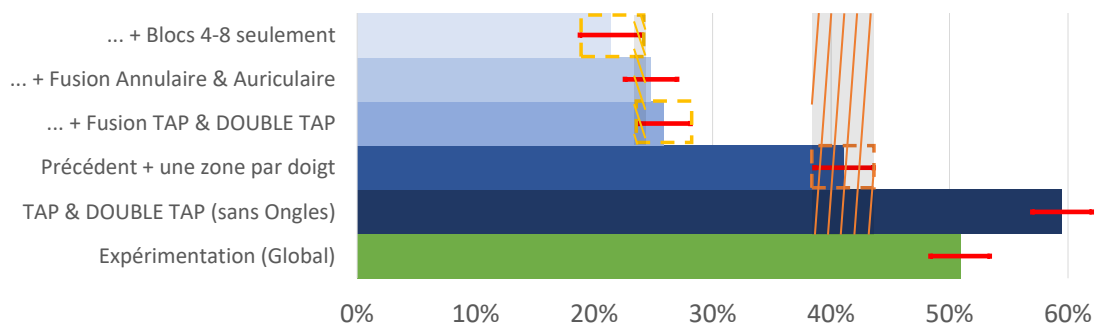


FIGURE 2.11 – Graphique illustrant le taux d'erreur avec intervalle de confiance pour les TAP et 2-TAP lors des sessions de test, pour chaque étape du processus de raffinement. La zone en surbrillance et les encadrés en pointillés (en jaune) montrent que les intervalles de confiance se chevauchent à peine entre la configuration d'une zone par doigt, la configuration fusionnée TAP et 2-TAP et la dernière étape de raffinement. La zone en surbrillance et l'encadré en pointillé (en orange) montrent qu'il n'y a aucun chevauchement d'intervalle de confiance entre la configuration "Une zone par doigt" et le reste.

Les erreurs liées aux gestes sont la deuxième source d'erreurs (31,8%). Étant donné que les TAP et les DOUBLE TAP sont du même type d'entrée, ils peuvent être fusionnés pour éviter toute confusion entre les deux μ Gestes. La figure 2.7 montre que les intervalles de confiance des TAP réalisés sur l'index, le majeur et l'annulaire se chevauchent à peine avec les intervalles de confiance des DOUBLE TAP réalisés sur les mêmes doigts. Cela suggère que les TAP réalisés sur les phalanges de ces trois doigts sont significativement moins sujets aux erreurs que les DOUBLE TAP réalisés sur les mêmes zones, et que la réalisation correcte des DOUBLE TAP est plus difficile. Nous fusionnons donc les DOUBLE TAP avec les TAP en tant que deuxième

2.2. ÉTUDE EXPÉRIMENTALE

étape d'affinement : nous considérons que toutes les DOUBLE TAP réalisées sont des TAP. Dans cette configuration, la figure 2.11 montre que le taux d'erreur global passe de 41,1% à 25,9%. Les TAP sont réduits de 11 (4 TAP et 4 DOUBLE TAP, un par doigt + 3 TAP sur les ongles) à 7 (4 TAP, un par doigt + 3 TAP sur les ongles).

Comme le montre le tableau 2.10 et la figure 2.7, les μ Gestes TTF réalisés sur l'auriculaire sont unanimement les plus sujets aux erreurs. Les participants leur attribuent également le plus bas niveau de confort, comme le montre la figure 2.8. Cela est conforme aux résultats de la littérature [94, 101, 178]. Au lieu de supprimer complètement les μ Gestes TTF sur l'auriculaire, nous fusionnons l'auriculaire avec l'annulaire : nous considérons ces deux doigts comme un seul doigt (et une seule zone), à l'exception des μ Gestes TTF sur les ongles. Si un TAP sur l'annulaire ou l'auriculaire est requis, il peut être réalisé sur l'un ou l'autre de ces deux doigts. Ce faisant, le taux d'erreur passe de 25,9% à 24,8%, comme le montre la figure 2.11. Les TAP sont réduits de 7 (4 TAP et 4 DOUBLE TAP, un par doigt + 3 TAP sur les ongles) à 6 (3 TAP + 3 TAP sur les ongles).

Enfin, la figure 2.6 montre que les blocs 1 à 3 ont un taux d'erreur moyen plus élevé que les blocs 4 à 8. Cela suggère un effet d'apprentissage, car les participants deviennent progressivement meilleurs dans la réalisation des μ Gestes TTF. Nous pouvons prédire un taux d'erreur moyen pour un utilisateur entraîné en n'utilisant que les derniers blocs. Comme le montre la figure 2.11, en ne tenant compte que des blocs 4 à 8, le taux d'erreur diminue de 24,8% à 21,4%. Les encadrés en surbrillance et en pointillés (en jaune) en haut à gauche de la figure 2.11 se chevauchent peu, ce qui suggère que les deux dernières étapes d'affinement (c'est-à-dire la fusion de l'annulaire et de l'auriculaire, ainsi que la prise en compte des blocs 4-8 uniquement) ont un impact significatif sur le taux d'erreur.

Après avoir affiné l'ensemble des μ Gestes TTF TAP et DOUBLE TAP, nous disposons de 6 μ Gestes TTF TAP : 3 sur les doigts (index, majeur, annulaire / auriculaire) et 3 sur les ongles (majeur, annulaire, auriculaire). Comme le montre la figure 2.11, ce processus d'affinement réduit le taux d'erreur de 59,5% à 21,4%.

2.2.6 Optimisation des μ Gestes TTF : SWIPE

Nous affinons les gestes SWIPE séparément, car il s'agit d'un type de gestes différent : ce sont notamment les seuls gestes bidirectionnels. Lorsque nous ne prenons en compte que les gestes SWIPE, le tableau 2.12 montre un Taux d'erreur de 44,4% et que 11,7% des erreurs en phase de test sont des erreurs de doigt.

		# Essais	Occurrence d'erreur	Taux d'erreur	CI	
Tous les blocs : SWIPE seulement	Global	547	243	44,4%	±4,2%	
	Type d'erreur	Geste		208	38%	±4,1%
		Doigt	547	64	11,7%	±2,7%
		Zone		58	10,6%	±2,6%
	Doigt cible	Index	141	54	38,3%	±5%
		Majeur	144	41	28,5%	±4,3%
		Annulaire	136	75	55,1%	±5,8%
		Auriculaire	126	73	57,9%	±6%

FIGURE 2.12 – Tableau récapitulatif par type d'erreur et par doigts pour les SWIPE uniquement.

Pour éviter de réaliser accidentellement un geste SWIPE sur un autre doigt, nous gardons seulement un SWIPE actif à la fois, sur un doigt qui génère le moins d'erreurs. Cela signifie que nous considérons qu'un SWIPE effectué sur un autre doigt ne se produit pas et ne déclenche donc pas d'erreur. Le Tableau 2.12 montre que pour les SWIPE, le majeur est le doigt qui génère le moins d'erreurs (28,5%), suivi par l'index (38,3%).

Dans la figure 2.6, les encadrés en pointillés (en jaune) montre que les intervalles de confiance des SWIPE effectués sur l'index et le majeur ne se chevauchent pas avec ceux des doigts annulaire et auriculaire. Cela suggère fortement que les SWIPE effectués sur l'index ou

2.2. ÉTUDE EXPÉRIMENTALE

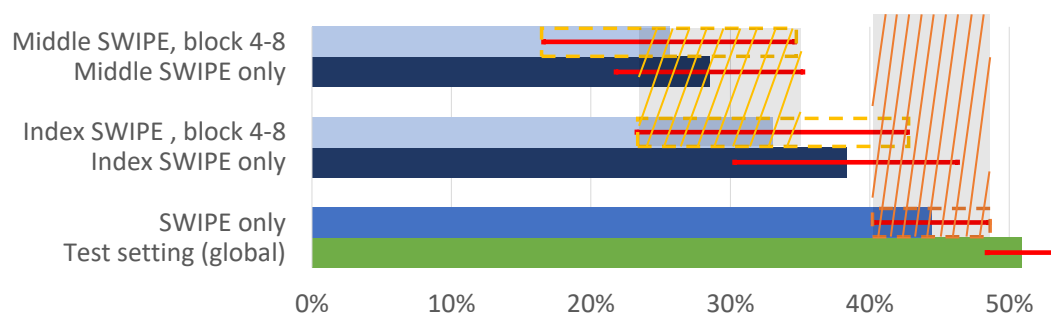


FIGURE 2.13 – Graphique illustrant le taux d'erreur avec intervalle de confiance pour les SWIPE lors des sessions de test, pour chaque étape du processus de raffinement. La zone en surbrillance et les encadrés en pointillés (en jaune) montrent que les intervalles de confiance se chevauchent entre les SWIPE de l'index et du majeur pour les blocs 4 à 8. La zone en surbrillance et l'encadré en pointillé (en orange) montrent qu'il y a un chevauchement entre les intervalles de confiance des SWIPE (tous doigts confondus) et les SWIPE effectués sur l'index. Il n'y a cependant qu'un faible chevauchement avec les SWIPE effectués sur l'index dans les blocs 4-8 et aucun chevauchement avec les SWIPE effectués sur le majeur..

le majeur sont moins sujets aux erreurs que les SWIPE effectués sur l'annulaire et l'auriculaire. De plus, dans la figure 2.8, la zone en surbrillance et l'encadré en pointillé (en orange) montrent que l'index et le majeur sont les préférés des utilisateurs. Par conséquent, pour éviter toute confusion et fausse reconnaissance, nous recommandons d'utiliser uniquement un doigt sur lequel effectuer le SWIPE, soit l'index soit le majeur. La zone en surbrillance et l'encadré en pointillé (en orange) dans la figure 2.13 montrent que l'intervalle de confiance des SWIPE (tous doigts confondus) se chevauche avec l'intervalle de confiance des SWIPE effectués sur l'index, mais pas sur le majeur. Cela suggère que n'utiliser que l'index pour effectuer des SWIPE n'est pas significativement différent que d'utiliser SWIPE avec tous les doigts activés. Cependant, ce n'est pas le cas après que les participants sont montés en compétence (c'est-à-dire après les blocs 1 à 3) : la zone en surbrillance et les encadrés en pointillé (en jaune) montrent que les intervalles de confiance des SWIPE effectués sur l'index et le majeur se chevauchent, suggérant fortement que les deux ne sont pas significativement différents après que les participants soient montés en compétence. Dans cette configuration, le taux d'erreur du SWIPE diminue de 44,4% à 38,3% avec seulement l'index activé, et à 28,5% pour le majeur. Puis, comme nous l'avons fait pour les TAP, nous ne prenons en compte que les blocs 4 à 8. La figure 2.13 montre que cela diminue le Taux d'erreur de 38,3% à 33% pour l'index, et de 28,5% à 25,6% pour le

2.3. SYNTHÈSE

majeur. L'ensemble de gestes SWIPE est ainsi réduit de 8 (1 SWIPE par doigt dans chaque direction) à 2 (1 SWIPE sur l'index ou le majeur, dans chaque direction).

Nous ne pouvons pas améliorer davantage le taux d'erreur : le Tableau 2.6 montre que la plupart des erreurs sont des erreurs de geste (38%), ce qui signifie que les SWIPE sont détectés comme des TAP ou des DOUBLE TAP. Le contexte de l'expérience pourrait être une explication. En raison de la position contrainte de la main et de la situation sans retour visuel, les participants ont probablement effectué les SWIPE en deux étapes : d'abord, ils ont ciblé la zone et établi le contact, puis ils ont effectué le mouvement de balayage. Cependant, le prototype de gant était conçu avec un seuil temporel de 500 ms initié au premier contact, après quoi, en l'absence d'autres informations, le contact était considéré comme un TAP. Ainsi si le participant effectuait son geste SWIPE en deux étapes et de manière lente (> 500 ms), le geste n'était pas reconnu comme un SWIPE. Nous soulignons ici que l'erreur n'est donc pas uniquement due aux participants mais peut être aussi due aux limites du système de captation et d'interprétation.

Après avoir affiné l'ensemble des μ Gestes TTF SWIPE, l'ensemble final comprend 2 μ Gestes TTF SWIPE disponibles, Haut-Bas et Bas-Haut, soit sur l'index, soit sur le majeur. Comme le montre la figure 2.13, le processus d'affinement réduit le taux d'erreur de 44,4% à 33% pour l'index, et à 28,5% pour le majeur. Les encadrés en surbrillance et en pointillés supérieures gauche (en jaune) dans la figure 2.13 montrent que les intervalles de confiance pour les SWIPE effectués sur l'index et le majeur pour les blocs 4 à 8 se chevauchent beaucoup. Bien que nous ne puissions pas affirmer qu'ils sont strictement équivalents, nous croyons qu'ils peuvent être utilisés de manière interchangeable avec des différences minimales entre les deux.

2.3 Synthèse

Notre processus d'affinement appliqué à notre ensemble initial de 33 μ Gestes TTF nous conduit à définir un ensemble final de 8 μ Gestes TTF (figure 2.14) : 3 TAP sur les doigts (c'est-à-dire index, majeur, annulaire / auriculaire), 3 TAP sur les ongles (majeur, annulaire, auriculaire)

2.3. SYNTHÈSE

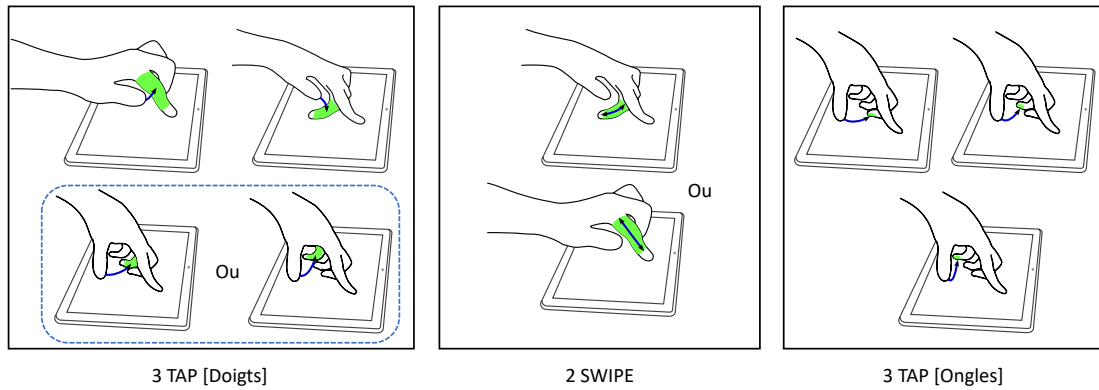


FIGURE 2.14 – Ensemble de μ Gestes TTF pouvant être effectués tout en touchant un écran tactile. **À gauche** : 3 TAP effectués par le pouce sur l'index, le majeur et soit l'annulaire, soit l'auriculaire. **Au centre** : 2 SWIPE le long de l'index ou du majeur, dans les deux directions. **À droite** : 3 TAP sur l'ongle du majeur, de l'annulaire et de l'auriculaire.

et 2 SWIPE (sur l'index ou le majeur, dans deux directions) qui peuvent être utilisés simultanément à l'écran tactile – donc lorsque l'index est en contact avec une surface.

Afin de communiquer précisément de quel μ Gestes TTF il s'agit, nous utilisons la notation μ Glyph de Chaffangeon-Caillet et al. [48]. Les 6 TAP sur les doigts et les ongles sont notés ainsi : $x \in \{i|m|r/p\}$, $y \in \{f | n\}$, $\left(\overset{\circ}{\underset{\circ}{\nabla}}(x \bullet y); \overset{\circ}{\underset{\circ}{\blacktriangle}} \right) \parallel \overset{\circ}{\underset{\circ}{\blacksquare}}_{\text{tablette}} \bullet$

Les 2 SWIPE sur l'index ou le majeur (dans les deux sens) sont noté ainsi : $x \in \{i|m\}$, $\left(\overset{\circ}{\underset{\circ}{\nabla}}(x \bullet); \overset{\circ}{\underset{\circ}{\nabla}}(x \bullet); \overset{\circ}{\underset{\circ}{\blacktriangle}} \right) \parallel \overset{\circ}{\underset{\circ}{\blacksquare}}_{\text{tablette}} \bullet$ $x \in \{i|m\}$, $\left(\overset{\circ}{\underset{\circ}{\blacktriangle}}(x \bullet); \overset{\circ}{\underset{\circ}{\blacktriangle}}(x \bullet); \overset{\circ}{\underset{\circ}{\nabla}} \right) \parallel \overset{\circ}{\underset{\circ}{\blacksquare}}_{\text{tablette}} \bullet$

L'utilisation de ces 8 μ Gestes TTF donne un taux d'erreur de 27,9% (450 erreurs sur 1616 essais, IC 2,2%). Nous tenons à souligner que le taux d'erreur dépend de notre configuration expérimentale et qu'il sert uniquement à assurer une cohérence interne. Notre prototype de gant et son logiciel de traitement des événements présentaient certaines lacunes. Un prototype plus abouti de captation et d'interprétation des μ Gestes TTF aurait sans doute permis un score plus bas. En effet, ayant démontré le potentiel des μ Gestes TTF comme paradigme d'entrée combiné au Tactile, il convient d'approfondir le système de captation et d'interprétation des μ Gestes TTF pour pouvoir les capturer avec précision en fonction du contexte d'utilisation (par

exemple, une caméra dans un environnement fixe [178], une bague dans un environnement mobile [31]).

2.4 Conclusion

Dans ce chapitre nous avons présenté une première étude expérimentale portant sur 33 μ Gestes TTF parmi les plus courants de la littérature afin d'évaluer leur faisabilité et le confort perçu lorsque l'index est en contact avec une surface – et ce dans une situation sans retour visuel. À partir de cette première étude, nous identifions un ensemble de 8 μ Gestes TTF (c'est-à-dire 6 TAP et 2 SWIPE) qui peuvent être utilisés sans retour visuel et simultanément à l'écran tactile afin d'en compléter l'usage, permettant notamment d'augmenter la bande-passante et l'expressivité du système interactif. En effet, l'ensemble de 8 μ Gestes TTF que nous proposons constitue un compromis entre la résilience (c'est-à-dire un faible taux d'erreur) et l'expressivité (c'est-à-dire le nombre d'entrées possibles) : une modalité d'interaction utilisant ces 8 μ Gestes TTF, combinée à la modalité sur écran Tactile, multiplierait de façon réaliste le nombre de commandes au moins par 8.

Par ailleurs, bien que notre contexte soit particulier, nos résultats sont cohérents avec la littérature sur les points suivants : 1) effectuer des μ Gestes TTF sur l'index et le majeur est plus facile que sur les autres doigts et est davantage préféré par les participants [94, 101, 110, 219, 230] ; 2) les TAP sur les ongles du majeur, de l'annulaire et de l'auriculaire sont particulièrement utilisables et populaires auprès des participants [110, 124].

Dans le chapitre suivant, nous illustrons l'utilisation de ces μ Gestes TTF à destination des PDV en concevant trois applications multimodales combinant l'écran tactile avec les μ Gestes TTF. Ces applications traitent de tâches variées que sont une exploration de carte tactile, une tâche de copier/coller et une tâche de sélection de texte.

Chapitre 3

Utilité et Utilisabilité des microgestes Thumb-To-Finger en complément du touché sur un écran tactile

Dans le chapitre précédent, nous avons présenté nos travaux sur la faisabilité d'une interface multimodale utilisant conjointement l'écran tactile d'une tablette avec les microgestes (μ Gestes) Thumb-To-Finger (TTF) en l'absence de retour visuel. Notre processus de raffinement appliqué aux 33 μ Gestes TTF étudiés nous a conduit à définir un ensemble utilisable comprenant 6 TAP et 2 SWIPE.

Dans ce chapitre, nous présentons trois applications afin d'évaluer l'utilité et l'utilisabilité de ces μ Gestes TTF pour des Personnes ayant une Déficience Visuelle (PDV) dans des scénarios concrets : une exploration de carte tactile, une tâche de couper/coller et une tâche de sélection de texte. Nous présentons aussi 2 nouvelles implémentations de gants permettant de reconnaître les ensembles de μ Gestes TTF utilisés dans ces applications. Ce chapitre reprend

et présente des éléments issus d'un article publié et présenté dans la conférence ACM-ICMI 2023 : “ *μ GeT : Multimodal eyes-free text selection technique combining touch interaction and microgestures*” [72].

3.1 Question de recherche

Nous avons montré dans le chapitre précédent que les μ Gestes TTF étaient faisables lorsque l'index reste en contact avec la surface de l'écran tactile, et ce dans une situation de non-retour visuel. Après avoir étudié 33 μ Gestes TTF parmi les plus courants dans la littérature, nous avons défini un ensemble de 8 μ Gestes TTF que nous considérons comme les plus performants et les plus confortables à utiliser : 3 TAP sur les doigts (l'index, le majeur, l'annulaire ou l'auriculaire), 3 TAP sur les ongles (du majeur, de l'annulaire, de l'auriculaire) et 2 SWIPE (sur l'index ou le majeur, dans deux directions haut/bas) qui peuvent être utilisés simultanément à l'écran tactile, dans le but d'augmenter l'utilisabilité des dispositifs dotés d'écrans tactiles pour les PDV, en améliorant notamment l'accessibilité, la bande-passante et l'expressivité du système interactif.

Pourtant, bien que ces μ Gestes TTF soient faisables, permettent-ils d'améliorer l'interaction sur écran tactile? Pour répondre à cette question, nous proposons d'évaluer l'utilité et l'utilisabilité perçue des μ Gestes TTF lorsqu'ils sont utilisés conjointement à un dispositif tactile en situation de non-retour visuel, et ce au travers de trois applications permettant de placer les PDV dans des situations d'usage courantes.

La première application est basée sur l'exploration d'un document audio-tactile [37, 38] composée d'une tablette et d'une surcouche en papier thermorelief (autrefois appelé papier Zytex¹) permettant de créer du relief perceptible de façon tactile (figures 3.1, 3.2). Les PDV ont une façon d'interagir avec les documents tactiles différente par rapport aux utilisateurs voyants utilisant des documents graphiques. Étant donné qu'ils explorent le document avec leurs mains, l'intégration du contenu doit être réalisée de manière séquentielle, contrairement aux utilisateurs voyants, en agrégeant une succession de détails incomplets jusqu'à ce qu'ils aient suffisamment

1. Papier Zytex : <https://zychemltd.com/swell-touch-paper-for-tactile-graphics/>

3.1. QUESTION DE RECHERCHE

d'informations pour synthétiser et comprendre l'image d'ensemble [132]. Lorsqu'elles explorent des documents audio-tactiles, les PDV s'appuient sur des stratégies bimanuelles principalement basées sur les mouvements des deux index touchant le document tactile [170, 21, 22, 252].

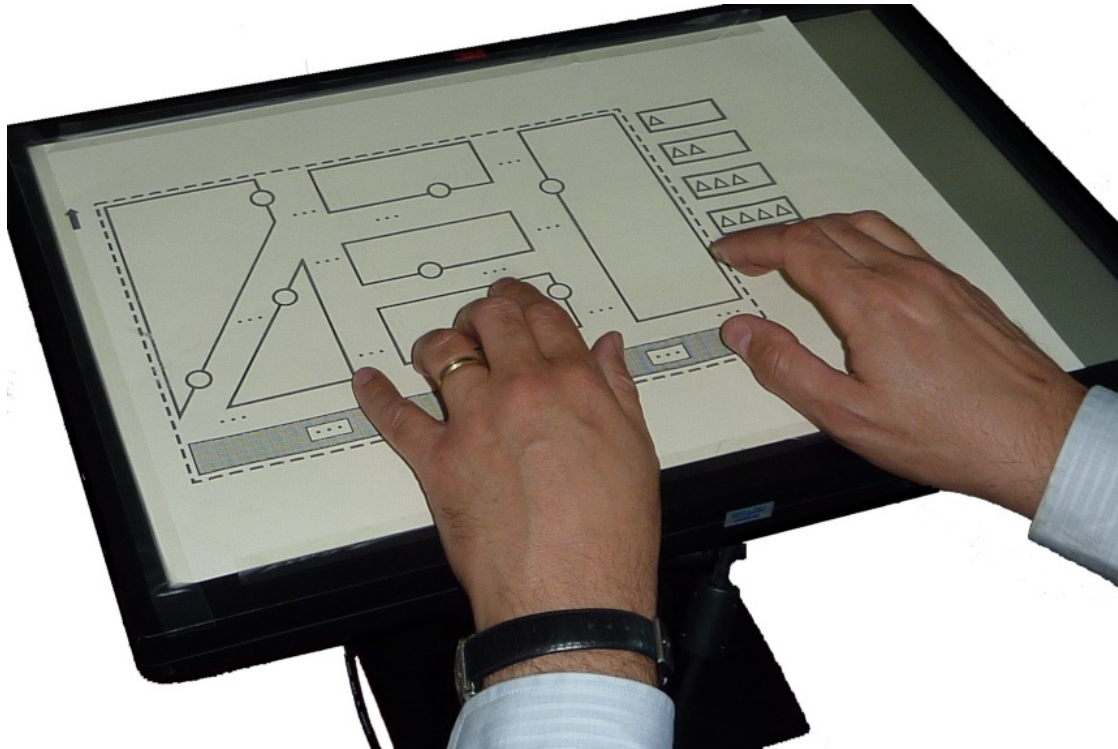


FIGURE 3.1 – Photographie d'un utilisateur explorant une carte interactive audio-tactile. Une feuille en papier thermorelief est superposée à l'écran tactile du dispositif. Sur la droite du dispositif, une zone de menu composée de boutons rectangulaires (contenant un ou plusieurs triangles pour les différencier). Photographie issue de INRIA - Potioc

Cependant, les dispositifs audio-tactiles existants nécessitent souvent des allers-retours entre le document exploré et un menu de commandes placé sur le côté (figures 3.1, 3.2), ce qui interrompt les stratégies d'exploration et augmente la charge cognitive [37]. Dans ce scénario, les μ Gestes TTF visent à résoudre ce problème : notre objectif est d'utiliser les μ Gestes TTF pour déclencher des commandes et des retours d'informations localisés sans interrompre le contact avec le graphique exploré.

Pour concevoir cette application nous nous sommes inspirés d'exemples tirés de la littérature [37, 38, 65, 144]. Nous avons notamment implémenté cette application à l'aide d'une version

3.1. QUESTION DE RECHERCHE

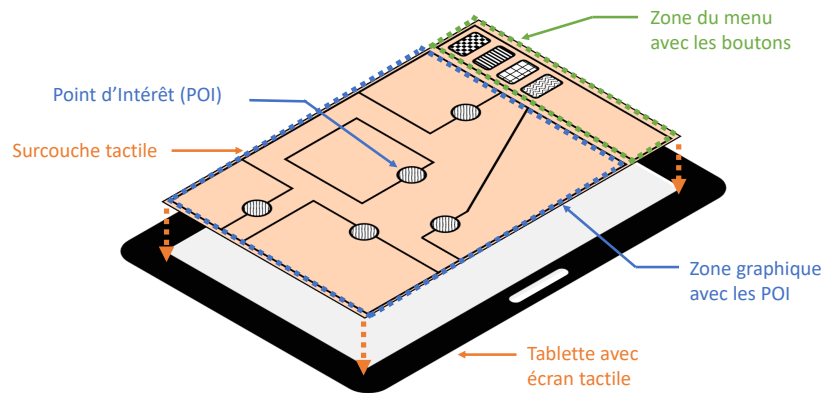


FIGURE 3.2 – Illustration schématisée d'un dispositif audio-tactile composé d'une tablette et d'une surcouche en relief.

personnalisée du logiciel DERi² développé par le laboratoire Cherchons Pour Voir à Toulouse [37, 38], sur laquelle nous avons ajouté la possibilité d'interagir avec les μ Gestes TTF.

Nous avons dans un premier temps testé les μ Gestes TTF auprès des PDV à l'aide de cette application. Par la suite, nous avons recueilli des idées auprès d'autres PDV lors d'un atelier de groupe et d'entretiens individuels. En effet, les deux applications suivantes sont le résultat d'un atelier d'une heure mené avec 11 PDV (dont 5 femmes) âgés de 24 à 58 ans recrutés dans un centre d'éducation spécialisé. Nous avons également mené quatre entretiens d'une heure : trois sur l'utilisation générale du téléphone (avec un expert en VoiceOver (VO) et deux novices) et un sur la manipulation de texte (avec un expert en VO). Les participants avaient des antécédents différents et des niveaux de déficience visuelle différents. Cet atelier visait à comprendre l'utilisation des écrans tactiles dans leur vie quotidienne et à identifier les obstacles d'utilisabilité qui pourraient être évités en utilisant les μ Gestes TTF. L'atelier comprenait une discussion de 10 minutes sur leur utilisation quotidienne, suivie de trois discussions de groupe parallèles

2. DERi (Cherchons Pour Voir) : <https://cherchonspourvoir.org/projet/nos-projets/deri-la-surface-interactive>

3.1. QUESTION DE RECHERCHE

de 25 minutes (chacune dirigée par un organisateur). Au cours de ces sessions de groupe, nous avons présenté le concept de μ Geste TTF (qui était nouveau pour chaque participant) et leur avons demandé comment ils pourraient les utiliser avec leurs applications préférées. Nous avons terminé l'atelier par une session de 25 minutes, au cours de laquelle nous avons compilé et discuté de toutes les idées générées. Les résultats de cet atelier sont les suivants : tous utilisaient des applications de messagerie (comme WhatsApp ou Telegram) et des applications multimédias (principalement YouTube et Spotify). Tous utilisaient un lecteur d'écran. Bien que ceux ayant des restes visuels regardaient des vidéos et lisaient, presque aucun d'entre eux n'utilisait des applications qui nécessitaient de parcourir de grands documents visuels (comme Google Maps). Seuls deux participants jouaient à des jeux sur leur appareil, bien que quatre d'entre eux ont affirmés qu'ils joueraient si les jeux étaient plus accessibles.

Tous les groupes ont indépendamment proposé d'utiliser les μ Geste TTFs comme raccourcis définis par l'utilisateur au niveau du système d'exploitation (par exemple, ouvrir une application spécifique, changer le volume, passer d'une application ouverte à l'autre) ou des raccourcis courants au niveau de l'application (par exemple, couper/coller, page précédente/suivante), qui sont faciles à utiliser sur un ordinateur de bureau mais peuvent être très difficiles à utiliser avec des outils d'accessibilité sur écran tactile en raison de la longueur de la trajectoire d'interaction et de la complexité des menus hiérarchiques. Ils étaient également préoccupés par le maintien de la cohérence de ces raccourcis entre les applications, afin d'éviter d'apprendre de nouvelles correspondances pour chaque nouvelle application. Cela était principalement motivé par une stratégie à laquelle ils adhéraient tous : contourner les tâches difficiles à effectuer avec les outils d'accessibilité. Par exemple, ils préféreraient réécrire un message à partir de zéro au lieu de le copier-coller, car cela nécessite une douzaine d'opérations à effectuer. En effet, pour copier un mot avec VO (sans utiliser de raccourci), il faut effectuer une douzaine d'opérations, y compris des mouvements de rotation avec deux doigts pour naviguer dans le menu rotatif. La tâche est longue, pénible et difficile - ou du moins inconfortable - pour les utilisateurs.

Concrètement, ils ont suggéré que l'approche multimodale pourrait être prometteuse pour accéder plus facilement aux fonctionnalités du téléphone. En effet, ils ont mentionné comment

3.1. QUESTION DE RECHERCHE

les outils d'accessibilité actuels et leurs menus les contraignent à des interactions linéaires, qu'ils trouvent longues et fastidieuses, au point que la plupart d'entre eux refusent d'utiliser les fonctionnalités courantes de leur téléphone comme le couper/coller. Nous avons donc choisi pour les applications suivantes d'utiliser les μ Gestes TTF pour augmenter l'expressivité de l'écran tactile afin de raccourcir les trajectoires d'interaction, jugées trop complexes par les PDV, et ce dans deux tâches courantes : le couper/coller et la sélection de texte.

La seconde application est basée sur une tâche de réorganisation d'une grille d'icônes. Comme rappelé ci-dessus et exposé dans le chapitre 1, les PDV ont une façon d'interagir différente avec les téléphones portables et les tablettes par rapport aux utilisateurs voyants. Ils doivent utiliser des Technologies d'Assistance (TA), et notamment des lecteurs d'écran tels que VoiceOver (VO) ou TalkBack (TB) pour accéder aux fonctionnalités du dispositif, ce qui rallonge considérablement les trajectoires d'interaction, y compris pour accéder à des fonctionnalités basiques telles que le couper/coller. Dans ce scénario, les μ Gestes TTF visent à résoudre ce problème : notre objectif est d'utiliser les μ Gestes TTF comme une commande supplémentaire afin d'augmenter l'expressivité du système. La finalité est double. D'une part, les μ Gestes TTF permettent à l'utilisateur d'accéder à une fonctionnalité courante à l'aide d'un μ Geste TTF facile à réaliser (fonctionnant comme un raccourci), et d'éviter de passer par l'étape de navigation dans un menu comme le proposent VO et TB. D'autre part, durant l'interaction, le μ Geste TTF permet de discriminer entre deux commandes différentes effectuées par le même geste, en l'occurrence un SWIPE horizontal (c'est à dire qu'un SWIPE horizontal effectué avec le μ Geste TTF ne déclenche pas la même commande qu'un SWIPE identique effectué sans μ Geste TTF). Dans ce scénario, les μ Gestes TTF ajoutent donc de l'expressivité de manière additive (sous la forme de raccourcis) et multiplicative (en changeant les commandes effectuées par un même geste sur l'écran tactile).

La troisième application consiste en une tâche de sélection de texte. À l'instar du second scénario, les μ Gestes TTF servent à augmenter l'expressivité de l'écran tactile et à éviter d'utiliser les menus hiérarchiques des TA telles que VO et TB qui allongent les trajectoires d'interaction, notamment lorsqu'il s'agit de faire des sélections qui requièrent différents niveaux de

3.1. QUESTION DE RECHERCHE

granularité (par exemple, une phrase et quelques mots). Dans ce scénario, les μ Gestes TTF visent à résoudre ce problème : notre objectif est de supprimer l'usage des menus hiérarchiques en augmentant l'expressivité de l'écran tactile. Dans ce scénario, les μ Gestes TTF permettent de discriminer entre plusieurs commandes effectuées par le même geste (à savoir, des SWIPE), multipliant par 3 le nombre de commandes possibles sur l'écran tactile.

3.1.1 Dispositif de captation

Dans le cadre de ces trois applications, nous avons utilisé la même approche que pour l'étude expérimentale présentée dans le chapitre précédent (Chapitre 2) pour capter les μ Gestes TTF : nous avons utilisé un gant avec des capteurs tactiles. Cependant, ce premier prototype devait capturer les 33 μ Gestes TTF de l'étude et embarquait pour ce faire beaucoup de capteurs, ce qui le rendait plutôt encombrant. Dans les trois applications présentées dans ce chapitre, nous n'utilisons qu'un sous-ensemble de μ Gestes TTF issus de l'ensemble de 8 μ Gestes TTF identifiés dans le chapitre précédent. Nous avons donc construit un prototype de gant modulaire, basé sur une technologie capacitive. Techniquement, il s'agit d'une carte Arduino Uno³ reliée à une carte capacitive permettant de détecter des contacts (Adafruit MPR121 12-Key Capacitive Touch Sensor⁴). Ce nouveau prototype permet d'y raccorder un ou plusieurs manchons en tissu à placer sur les doigts (figure 3.3) et comportant seulement les capteurs nécessaires à l'interaction - diminuant ainsi l'encombrement et améliorant l'ergonomie du gant. Utiliser une technologie capacitive nous a aussi permis de faire l'impasse sur la partie conductrice qui habillait le pouce dans le précédent prototype, et qui diminuait la perception tactile et la proprioception des participants (en particulier pour la première application d'exploration d'une carte)

3. Arduino Uno - site officiel : <https://www.arduino.cc/en/Guide/ArduinoUno>

4. Carte capacitive Adafruit MPR121 - site officiel : <https://learn.adafruit.com/adafruit-mpr121-12-key-capacitive-touch-sensor-breakout-tutorial>

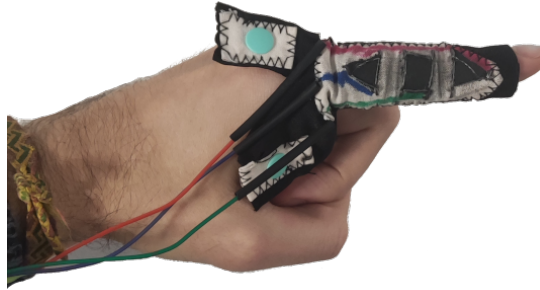


FIGURE 3.3 – Exemple de manchon pour l'index, doté de 3 capteurs en tissu conducteur et permettant de capter un TAP et les SWIPE dans les deux sens. Le TAP peut avoir lieu n'importe quel capteur tandis que les SWIPE doivent toucher successivement deux capteurs.

3.2 Études expérimentales

Nous avons exposé dans la section précédente nos motivations quant à l'étude de l'utilité et l'utilisabilité des μ Gestes TTF pour les PDV dans plusieurs contextes définis par trois applications. Cette section introduit les études expérimentales que nous avons menées pour tester ces trois applications.

3.2.1 Étude 1 : exploration de carte

Scénario d'usage

Jean est un collégien ayant une déficience visuelle. Il apprend la géographie de la Moldavie. Sa classe utilise une application spécialisée sur tablette permettant l'usage d'une surcouche tactile superposable à l'écran, permettant ainsi à des PDV d'interagir avec des graphiques numériques de façon tactile, comme présenté dans la section précédente.

Avec cette application, l'objectif de John est de récupérer plusieurs informations sur les villes et les régions de la France. Dans cette application, 4 μ Gestes TTF issus de l'ensemble de 8 identifiés dans le chapitre précédent sont intégrés. Deux TAP : un sur l'index pour déclencher une action fréquente et un sur l'ongle du petit doigt pour en déclencher une plus ponctuelle ; et deux SWIPE sur l'index (dans les deux sens) pour naviguer dans un menu. Aucun autre geste n'est reconnu. Avec cet ensemble de μ Gestes TTF, John peut sélectionner un point d'intérêt

3.2. ÉTUDES EXPÉRIMENTALES

(POI) en le touchant sur l'écran tactile avec son index. Il peut ensuite naviguer avec le SWIPE pour faire défiler un menu cyclique⁵ avec plusieurs catégories ("Nom", "Population", "Superficie" ou "Dirigeant"). Il peut changer le niveau d'abstraction de l'information ("Ville" ou "Région") en effectuant un TAP sur l'ongle du petit doigt. À tout moment, un TAP sur son index déclenche un retour audio lui fournissant l'information correspondant au POI qu'il touche avec son index, et à la catégorie et au niveau d'information actuellement sélectionnés (par exemple, la superficie de la ville de Chisinau).

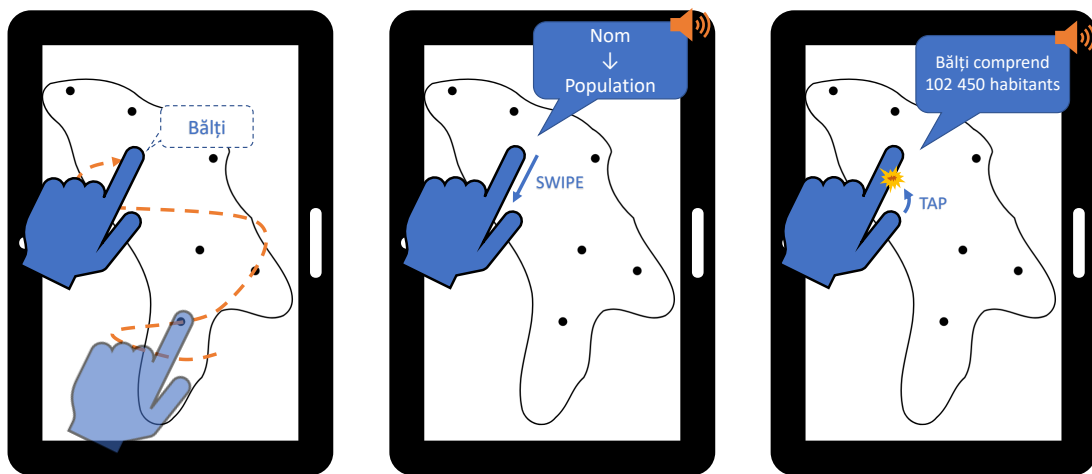


FIGURE 3.4 – Exemple d'utilisation de la carte audio-tactile : toucher un Point d'Intérêt (POI) le sélectionne ; effectuer un SWIPE de l'index change de catégorie ; effectuer un TAP de l'index déclenche un retour audio.

Protocole expérimental

Notre étude vise à illustrer l'utilisation des μ Gestes TTF dans le cadre d'une interaction multimodale combinant μ Gestes TTF et écran tactile à destination de PDV, et ce afin d'obtenir leur retour qualitatif.

Nous avons recruté 7 participants (dont une femme) ayant une déficience visuelle, tous droitiers, âgés de 23 à 63 ans (moyenne de 36,3, écart type de 14). Nous les avons soigneusement sélectionnés de manière à ce qu'ils aient des expériences différentes et que leurs niveaux de

5. Menu dont le dernier élément boucle avec le premier.

3.2. ÉTUDES EXPÉRIMENTALES

déficience visuelle varient entre la catégorie 1 (c'est-à-dire une déficience visuelle légère) et la catégorie 5 (c'est-à-dire une cécité complète) de la classification de l'OMS [76]. Les participants 2, 4, 5, et 7 ont déjà utilisé des documents tactiles. Aucun des participants n'a signalé d'autres handicaps mentaux ou perceptuels que la déficience visuelle. Tous utilisent quotidiennement un appareil tactile.

Chaque session a commencé par une présentation des μ Gestes TTF et du système. Les participants ont essayé les 4 μ Gestes TTF : d'abord dans l'air, puis avec l'index sur la carte, et enfin avec le gant. Après avoir essayé toutes les interactions plusieurs fois, nous leur avons demandé d'explorer la carte pour répondre à une série de 14 questions telles que "Combien d'habitants y a-t-il dans la ville X?".

Ces questions ont permis de garantir une grande variété d'interaction multimodale (c'est-à-dire en changeant de "POI", de "catégorie" et/ou de "niveau d'abstraction"). Nous leur avons demandé de penser à haute voix. Après la tâche d'exploration, les participants ont répondu à un questionnaire sur l'utilité et l'utilisabilité perçues. Les questions sur l'utilisabilité portaient sur la facilité d'exécution et d'apprentissage des μ Gestes, ainsi que sur d'éventuels inconforts. Les questions sur l'utilité couvraient l'utilité perçue des μ Gestes TTF et les qualités de la technique d'interaction propre au scénario, ainsi que leur avis sur d'autres scénarios d'utilisation potentiels. Le questionnaire se trouve en annexe (Annexe A).

Implémentation

Nous avons mis en œuvre les interactions décrites ci-dessus en utilisant une version personnalisée du logiciel DERi développé par le laboratoire Cherchons Pour Voir à Toulouse [37, 38]. Nous avons déployé ce système sur une tablette de 10,1 pouces (Huawei MediaPad T5 Lite, 24*15,5 cm), posée sur une table devant les participants.

Pour capturer les μ Gestes TTF effectués, nous avons utilisé le concept de gant modulaire

3.2. ÉTUDES EXPÉRIMENTALES

présenté dans le paragraphe 3.1.1. Le gant est composé d'un manchon en tissu élastique sur lequel sont brodés trois connecteurs en fil et en tissu conducteur, chacun relié à une entrée de la carte capacitive. Ce manchon permet de détecter les 4 μ Gestes TTF nécessaires à l'expérimentation (figure 3.3).

À chaque événement, le programme de la carte Arduino envoie à l'ordinateur une liste de booléens (0 et 1), chacun correspondant à l'état d'un connecteur (absence ou non de contact). Un logiciel personnalisé développé en Python tournant sur un ordinateur reçoit et transforme les données provenant du gant en événements représentant les 4 μ Gestes TTF utilisés. Ces événements sont enregistrés et transmis à la tablette en WiFi.

Résultats

Nous avons analysé les questionnaires et les commentaires libres pour extraire les éléments clés et synthétiser les résultats.

Concernant l'**utilisabilité**, tous les participants ont convenu que le TAP de l'index était facile à réaliser. P2 et P7 ont eu quelques problèmes avec le TAP sur l'ongle du petit doigt. P3 a trouvé que "bien que facile à faire, c'était le μ Geste TTF le plus difficile de l'ensemble". Tous sauf P4 ont trouvé les gestes SWIPE faciles à réaliser. P5 et P7 ont déclaré "une fois qu'on a compris, c'est facile" et "tout le monde peut le faire". Trois participants ont déclaré que les μ Gestes TTF étaient rapides à exécuter. Cinq ont déclaré qu'ils étaient pratiques. Deux ont dit qu'ils étaient facilement accessibles.

Tous les participants ont fortement convenu de la facilité d'apprentissage de la technique d'interaction. Les observations in situ appuient davantage leurs commentaires : tous sont devenus indépendants après quelques minutes d'utilisation, seul P2 avait besoin de quelques rappels tout au long de la session. P1, P3 et P6 ont trouvé la technique d'interaction "très intuitive". P3 a dit qu'il pouvait "imaginer le menu le long de [son] doigt". P6 a dit que l'ensemble de

3.2. ÉTUDES EXPÉRIMENTALES

μ Gestes peut être "facilement mémorisé".

Mis à part les problèmes avec le TAP sur l'ongle du petit doigt (P2, P7), aucun d'entre eux n'a mentionné quoi que ce soit de négatif sur la technique d'interaction. Ils ont plutôt souligné des problèmes liés au système développé, en particulier la latence et la taille du gant.

Concernant l'utilité, tous les participants ont déclaré que cette technique d'interaction peut être utile pour interagir avec un appareil disposant d'un écran tactile en situation de non-retour visuel. P1, P3 et P7 ont spontanément mentionné que les μ Gestes TTF permettent d'interagir tout en gardant la main en place au lieu de devoir la déplacer (par exemple, pour effectuer un geste à l'écran ou pour interagir avec un menu latéral) et de perdre le contexte. P2 et P7 ont également déclaré que les μ Gestes TTF pourraient être utiles pour modifier les paramètres à la volée (par exemple, arrêter une rétroaction audio longue ou indésirable, changer la vitesse de sortie de l'audio).

En ce qui concerne l'utilisation des μ Gestes TTF dans d'autres situations, six participants ont mentionné d'abord un contexte éducatif (par exemple, apprendre la géographie, l'anatomie ou la géométrie à l'école). P3 a déclaré qu'ils pourraient être utilisés sur n'importe quelle surface. Il a également mentionné qu'ils pourraient être utilisés comme entrées supplémentaires avec un ordinateur (par exemple, comme raccourcis ou macro⁶) ou comme des touches de clavier dite "modifiantes" (par exemple, CTRL, ALT) mais pour les écrans tactiles. Il a également déclaré que les techniques basées sur les μ Gestes TTF pourraient être particulièrement utiles pour les personnes souffrant d'un handicap moteur (par exemple, la tétraplégie). P5 et P6 ont mentionné que cela serait intéressant pour les jeux.

En résumé, tous les participants ont trouvé les μ Gestes TTF utiles, faciles à utiliser, intuitifs et confortables lorsqu'ils sont utilisés pour explorer un document audio-tactile. Ils étaient très intéressés par les possibilités d'interaction que les μ Gestes TTF peuvent offrir et étaient particulièrement curieux des scénarios concrets que nous développerions à l'avenir.

6. Macro - Wikipédia : <https://fr.wikipedia.org/wiki/Macro-d%C3%A9finition>

3.2.2 Étude 2 : couper / coller

Scénario d'usage

Émilie est une PDV qui a de nombreuses applications sur sa tablette. Elle télécharge généralement une application par semaine pour tester les nouveautés, elle a donc souvent besoin de les réorganiser au sein de son écran d'accueil. Bien qu'elle puisse lire son écran avec un zoom suffisant, elle préfère utiliser les lecteurs d'écran (comme celui de VO) lorsque cela est possible. Dans ce scénario, seuls deux μ Gestes TTF sont utilisés : les SWIPE dans les deux directions sur l'index. Aucun autre μ Geste TTF n'est reconnu. Ces deux μ Gestes TTF sont issus de l'ensemble de 8 identifiés dans le chapitre précédent.

Avec cet ensemble de μ Gestes TTF, Émilie peut sélectionner l'icône A qu'elle souhaite déplacer en la touchant sur l'écran (figure 3.5), puis la couper avec un SWIPE Haut-Bas (c'est-à-dire du haut du doigt vers la base du doigt). Métaphoriquement, Émilie "glisse l'icône dans sa manche". Elle déplace ensuite son doigt sur l'icône B pour coller l'icône A avec un SWIPE Bas-Haut (c'est-à-dire de la base du doigt vers le haut du doigt) tout en maintenant son doigt sur l'écran. Métaphoriquement, Émilie "glisse l'icône de sa manche vers l'espace de travail". Pour décider si l'icône A sera placée avant ou après l'icône B, elle peut, après avoir effectué le SWIPE, déplacer son doigt dans l'une des trois zones disponibles (figure 3.5, étape 6). La zone de gauche, qui s'étend du bord gauche de la tablette jusqu'au bord gauche de l'icône B, permet de déplacer l'icône A avant l'icône B. La zone de droite, qui s'étend du bord droit de la tablette jusqu'au bord droit de l'icône B, permet de déplacer l'icône A après l'icône B. La zone centrale permet d'annuler l'action, de sorte à ce qu'aucune réorganisation de la grille n'ait lieu. En levant son doigt de l'écran, Émilie confirme l'action choisie selon la zone sur laquelle se trouve son doigt et la grille d'icônes sera réorganisée.

3.2. ÉTUDES EXPÉRIMENTALES

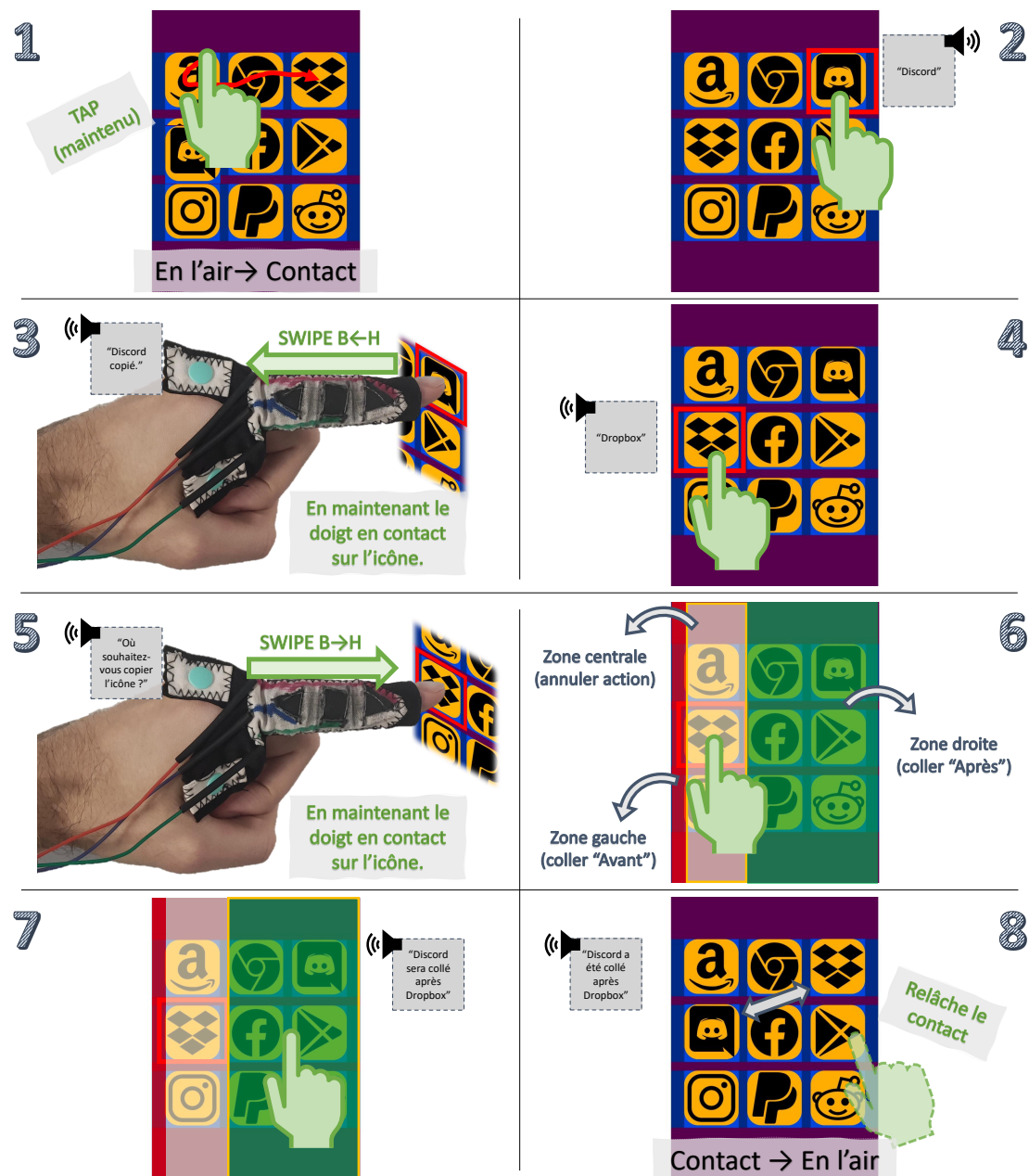


FIGURE 3.5 – Exemple d’une interaction complète. (1) L’utilisateur touche d’abord l’écran avec son doigt. (2) Lorsqu’elle franchit la limite d’une nouvelle icône, un retour audio nommant l’icône actuellement sélectionnée est joué. (3) Après avoir positionné son doigt sur l’icône désirée, un balayage de haut en bas de son doigt, tout en maintenant le contact, découpe l’icône actuelle. Un retour audio confirme la découpe. (4) Elle déplace maintenant son doigt vers une autre icône. (5) Un balayage de bas en haut initie une action de collage. Un retour audio est joué. (6) Elle peut maintenant déplacer son doigt sur un panneau à gauche de l’icône pour coller avant, sur un panneau à droite de l’icône pour coller après, ou laisser son doigt sur le panneau central pour annuler l’action de collage. Un retour audio l’informe chaque fois qu’elle franchit la limite d’un panneau. (7) Elle choisit de coller l’icône après. (8) Enlevant son doigt confirme et effectue l’action de collage.

Protocole expérimental

Ce scénario est basé sur la simplification d'un processus extrêmement courant - à savoir la fonction "couper/coller" - en utilisant un ensemble minimal de 2 μ Gestes TTF. Afin d'en démontrer la faisabilité et d'obtenir un retour qualitatif de la part d'utilisateurs PDV, nous avons mis en place une étude expérimentale.

Nous avons testé notre application avec 6 personnes ayant une déficience visuelle (2 hommes, 4 femmes, 2 gauchers, 4 droitiers, âgés de 18 à 49 ans, âge moyen de 30,5 ans, niveaux de déficience visuelle allant de la catégorie 1 à la catégorie 2 [163]). Aucun d'entre eux ne connaissaient les μ Gestes TTF avant cette étude. Aucun des participants n'a signalé d'autres handicaps mentaux ou perceptuels que la déficience visuelle. Tous utilisent quotidiennement un appareil tactile.

Nous avons brièvement expliqué comment fonctionnait le gant et les avons laissés s'entraîner aux 2 μ Gestes TTF SWIPE sur l'index, en utilisant une application simple qui produisait une rétroaction audio "Balayage vers le bas" ou "Balayage vers le haut" lorsque le μ Geste TTF correspondant était effectué (respectivement un SWIPE de la base de la main vers le haut du doigt et vice-versa). Après cette phase d'entraînement, nous avons présenté l'application, composée de 2 pages. Chaque page comprend une grille d'icônes de 3x3 (toutes les icônes correspondent à des applications grand public, voir figure 3.5), imitant l'écran d'accueil d'une tablette. Des SWIPE horizontaux (sans μ Gestes TTF) permettent de naviguer entre les deux pages.

Nous avons également expliqué comment réorganiser les grilles d'icônes et les avons laissés utiliser le système librement pour qu'ils s'y habituent. La figure 3.5 illustre une trajectoire d'interaction type à effectuer pour couper/coller une icône. La tâche consistait à effectuer six réorganisations (quatre dans la même page et deux d'une page à l'autre). Les instructions étaient données verbalement (et pouvaient être répétées) sous forme de : "Coupez [nom de l'icône] et collez-la [avant / après] l'icône [nom de l'icône cible]".

3.2. ÉTUDES EXPÉRIMENTALES

Après l'expérience, les participants ont répondu à un questionnaire Me-CUE (annexe A) adapté aux personnes ayant une déficience visuelle [146] et ont donné une note globale de 1 à 10 pour la technique d'interaction qui leur a été présentée. Pour les données qualitatives, toutes les tâches ont été enregistrées en vidéo et ultérieurement catégorisées. Nous avons également réalisé une analyse des échanges libres avec les participants et autres commentaires qui ont eu lieu pendant l'étude.

Implémentation

Nous avons développé une application Ionic fonctionnant sur une tablette Samsung Galaxy Tab S2 de 9,7 pouces (taille de l'écran de 23,7x16,9 cm), posée à plat sur une table, et utilisant un prototype de manchon pour les doigts simple, ne disposant que de 3 capteurs, comme expliqué précédemment (voir paragraphe 3.1.1). Comme illustré dans la description du scénario, l'application ne reconnaît que 2 μ Gestes TTF (c'est-à-dire les SWIPE de l'index) pour simplifier considérablement le processus d'interaction par rapport aux outils d'accessibilité courants tels que VO ou TB.

Bien que les SWIPE utilisant uniquement l'index ou le majeur ne sont pas significativement différents en terme de performance (voir Chapitre 3 - Résultats), nous avons opté pour les SWIPE de l'index car ils sont préférés et plus confortables (voir Chapitre 2 et [94, 101]).

Résultats

Nous avons compilé les réponses des questionnaires Me-CUE, l'analyse des vidéos et les commentaires des participants pour couvrir les principaux aspects de l'expérience utilisateur concernant la technique d'interaction par μ Gestes TTF. Nous présentons d'abord les résultats globaux, puis nous présentons plus précisément les résultats dans le même ordre que le questionnaire, qui se compose de quatre parties :

- les qualités instrumentales de l'application, c'est-à-dire l'efficacité de l'application et sa

3.2. ÉTUDES EXPÉRIMENTALES

facilité d'utilisation (utilisabilité et utilité, partie 1 du questionnaire) ;

- les qualités non instrumentales, c'est-à-dire l'attrait et l'attractivité (statut, partie 2 du questionnaire ; en raison de la nature du prototype et des utilisateurs ciblés, nous avons décidé de ne pas prendre en compte l'esthétique visuelle et l'engagement) ;
- les émotions (partie 3 du questionnaire) ;
- les conséquences (intentions d'utilisation future, partie 4).

Les six tâches ont pris en moyenne 6 minutes et 50 secondes pour être effectuées, en comptant les instructions. La participante la plus rapide a mis 2 minutes et 30 secondes pour effectuer toutes les tâches, tandis que la plus longue a mis 10 minutes et 30 secondes. La tâche la plus rapide a été résolue en seulement 15 secondes. Un participant (P5) était dyspraxique et s'est arrêté au bout de 7 minutes et 15 secondes après avoir terminé seulement trois des six tâches. Les participants ont attribué à l'expérience qu'ils ont eue avec la technique d'interaction une note globale de 7,7 sur 10 (écart type de 1,5).

Concernant l'utilisabilité, tous les participants étaient d'accord sur la facilité d'utilisation de la technique d'interaction. Quatre d'entre eux ont mentionné qu'il fallait un certain temps pour s'y habituer. Tous ont trouvé les μ Gestes TTF SWIPE faciles à effectuer. Les observations vidéo confirment également leurs commentaires, montrant que cinq participants sont devenus autonomes (c'est-à-dire qu'ils ont enchaîné plusieurs tâches à la suite sans erreur ni hésitation) après seulement 3 minutes de pratique. Cependant, un participant a encore eu besoin de quelques rappels tout au long de l'expérience. Nous n'avons pas pu déterminer avec certitude quand P5 a acquis le modèle mental pour effectuer à la fois les gestes et utiliser la technique. Néanmoins, lors de notre séance de débriefing, elle a prouvé qu'elle avait compris les deux, même si elle avait des difficultés à les exécuter. Deux participants ont rencontré des difficultés spécifiques : par moment, P2 confondait les actions déclenchées par les deux μ Gestes TTF SWIPE (pour rappel, du bout du doigt vers la base pour copier et de la base vers le bout du doigt pour coller), et P1 avait du mal à comprendre comment les panneaux latéraux étaient liés à la position de l'icône cible (c'est-à-dire qu'elle ne comprenait pas ce que signifiait coller l'icône copiée "avant" ou "après" l'icône cible).

3.2. ÉTUDES EXPÉRIMENTALES

Concernant l'utilité perçue, tous les participants sauf un ont convenu que cette technique d'interaction était utile. Tous ont déclaré qu'elle était rapide, pratique et peu fatigante. Un participant a mentionné qu'elle était discrète et facilement disponible. P5 a déclaré que cela ne serait pas utile pour elle, bien qu'elle reconnaisse que cela pourrait l'être pour d'autres personnes ayant une déficience visuelle, en fonction de leur pathologie. Cependant, elle nous a admis que cela est motivé par sa réticence à changer ses habitudes. Quatre participants ont déclaré ne pas utiliser le mécanisme de copier/couper/coller sur leur téléphone en raison de sa complexité. Un participant ne savait tout simplement pas que c'était possible. Après l'expérience, tous les participants ont convenu que copier et coller serait beaucoup plus facile, l'une d'entre eux admettant même qu'elle "utiliserait [la technique par μ Geste TTF] tout le temps, pour couper/coller et plus encore".

Nous avons évalué la manière dont les participants pensaient être perçus par les autres lorsqu'ils utilisent cette technique d'interaction, et comment cela refléterait leur identité sociale. Cinq participants étaient fortement d'accord pour dire qu'ils l'utiliseraient à la maison, au travail, à l'extérieur ou dans les transports en commun. Il y avait aussi un fort consensus parmi les participants de penser que non seulement les PDV voudraient utiliser cette technique d'interaction, mais aussi des personnes voyantes.

Concernant les émotions ressenties, tous les participants ont déclaré se sentir heureux lorsqu'ils utilisaient la technique d'interaction. Quatre d'entre eux étaient curieux d'essayer quelque chose de nouveau, P3 était particulièrement enthousiaste et a dit "Génial! Excellente idée!", P4 était "ravi de pouvoir faire les choses en autonomie", ajoutant également que "cela ressemble à un jeu et [qu'il] pourrait ainsi réellement accomplir des choses". Lorsqu'on leur a demandé quelles émotions négatives ils avaient ressenties, quatre participants ont affirmé n'en avoir éprouvé aucune, P1 a ressenti de la frustration au début parce qu'elle avait du mal à comprendre comment fonctionnait la technique d'interaction, et P5 était inquiète et confuse quant à l'origine des erreurs qu'elle pouvait rencontrer (c'est-à-dire si l'erreur provenait d'une mauvaise manipulation, d'un mauvais usage de l'application, d'un bug ...).

Concernant les intentions d'usage, cinq participants ont déclaré qu'ils utiliseraient la technique si elle était disponible sur leur téléphone portable. P3 a ajouté : "Si ça marche, j'en veux un [maintenant]!", et P4 a déclaré : "Maintenant, tout ce que vous avez à faire, c'est de le commercialiser". Lorsqu'on leur a demandé pourquoi ils continueraient à l'utiliser, ils ont réitéré ce qu'ils avaient dit précédemment : c'est rapide, accessible, non fatigant et permet de réaliser des tâches. Du fait de sa dyspraxie, P5 a confirmé qu'elle ne l'utiliserait que si nécessaire, car elle pensait que la technique n'était pas suffisamment accessible dans sa situation.

3.2.3 Étude 3 : sélection de texte

Scénario d'usage

Abigaïl est une PDV qui aime écrire des messages sur les réseaux sociaux. Elle a généralement besoin d'éditer son texte plusieurs fois avant de l'envoyer, afin de s'assurer que celui-ci est bien écrit et dépourvu de fautes. Elle doit donc sélectionner les portions de textes qu'elle souhaite éditer (par exemple, copier, remplacer ou supprimer, mettre en gras / italique . . .). Pour ce faire, elle peut utiliser la solution fournie par VoiceOver (VO). Pour rappel, VO utilise des entrées tactiles discrètes pour naviguer dans un menu radial appelé rotor (figure 3.6) avec des commandes contextuelles, comme décrit dans le Chapitre 2, à l'aide de gestes de rotation rapides dans le sens des aiguilles d'une montre (ClockWise – CW) et dans le sens inverse des aiguilles d'une montre (Counter-ClockWise – CCW) avec deux doigts.

Il y a par défaut 8 éléments dans le rotor. Chaque élément du rotor est un mode persistant. Dans ce qui suit, nous nous référons au premier élément comme étant celui du "Nord", et nous comptons dans le sens des aiguilles d'une montre.

Pour sélectionner une partie de texte sur téléphone portable ou tablette, les utilisateurs doivent d'abord mettre en surbrillance une entité (par exemple, un caractère, un mot, une phrase ... en surbrillance grise **IPSUM** dans l'exemple suivant). Il peuvent le faire soit en pointant sur l'écran pour mettre en surbrillance l'entité survolée, soit en effectuant un glissement à un doigt vers la gauche/droite (appelé également SWIPE), déplaçant la surbrillance vers l'entité précé-

3.2. ÉTUDES EXPÉRIMENTALES

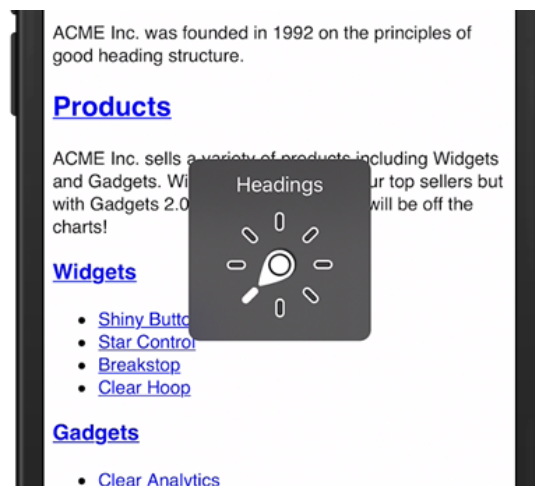


FIGURE 3.6 – Capture d'écran du rotor de VoiceOver, sous iOS.

dente (ou suivante). Placer le rotor en position 3 (Caractère), 4 (Mot) ou 5 (Phrase) change le type d'entité mise en surbrillance à l'aide du pointage directe ou du SWIPE. Par défaut, les mots sont mis en surbrillance.

L'ancre (c'est-à-dire le marqueur de sélection le plus à gauche, en bleu ↓ dans l'exemple suivant) est placée en premier, sa position étant réglée au début de l'entité mise en surbrillance (c'est-à-dire à gauche de son premier caractère). L'ancre ne peut pas être déplacée autrement. Pour sélectionner du texte, les utilisateurs déplacent ensuite la poignée de sélection (c'est-à-dire le marqueur de sélection le plus à droite, en bleu ↑ dans l'exemple suivant) initialement placée à la position de l'ancre. Pour déplacer la poignée de sélection, le rotor doit être placé en position #8 ("Sélection de texte"). Dans ce mode, la poignée est déplacée vers la gauche ou vers la droite en utilisant un SWIPE vers la gauche ou la droite. Tant qu'Abigaïl reste dans ce mode, elle peut également modifier la granularité à laquelle la poignée est déplacée en utilisant un SWIPE vers la haut ou le bas, en passant par 5 sous-modes (figure 3.7) : "caractère" (Character by Character – CbC), "mot" (Word by Word – WbW), "phrase" (Sentence by Sentence – SbS), "paragraphe" ou "sélectionner tout". La granularité est initialement réglée sur "mot".

Chaque changement (rotor, sous-modes, sélection mise à jour de la sélection de texte cou-

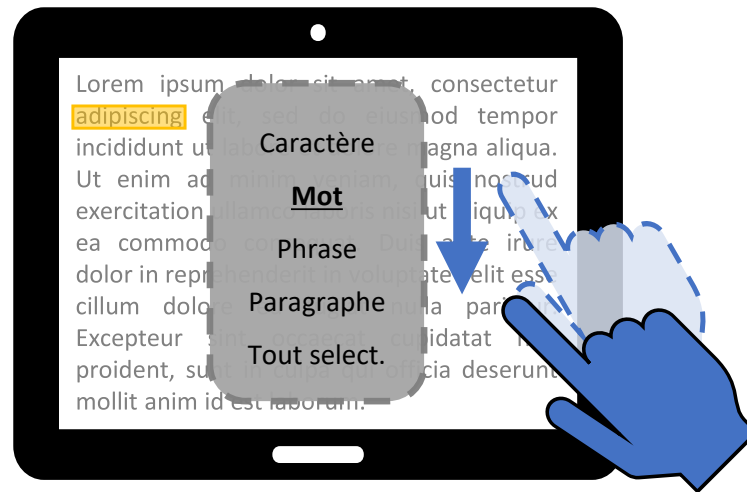


FIGURE 3.7 – Schéma du sous-menu accessible depuis la position #8 (Selection de texte) du rotor.

rante) déclenche un retour audio qui lit l'élément du menu ou la dernière entité qui a été mise en surbrillance ou sélectionnée. Tout ce qui est compris entre l'ancre et la poignée de sélection est considéré comme sélectionné (en surbrillance jaune **IPSUM** dans l'exemple suivant).

Voici un exemple de l'enchaînement des étapes de sélection :

- Surbrillance : LOREM **IPSUM** ET DOLOR SIC AMET
- Position des marqueurs de sélection : LOREM ↑↓ IPSUM ET DOLOR SIC AMET
- *Commande de sélection, qui déplace la poignée droite d'un mot vers la droite.*
- Sélection : LOREM ↑ **IPSUM** ↓ ET DOLOR SIC AMET

Pour aider les utilisateurs, nous avons ajouté un geste de DOUBLE TAP à effectuer sur l'écran, et qui déclenche un retour audio lisant la sélection en cours, s'il y a lieu, ou simplement arrêtant le retour audio en cours.

Avec cette technique d'interaction, Abigaïl peut sélectionner du texte mais elle doit interrompre sa tâche de sélection pour naviguer dans un menu afin de choisir les options de sélection, ce qui rallonge la trajectoire d'interaction et peut complexifier la tâche.

3.2. ÉTUDES EXPÉRIMENTALES

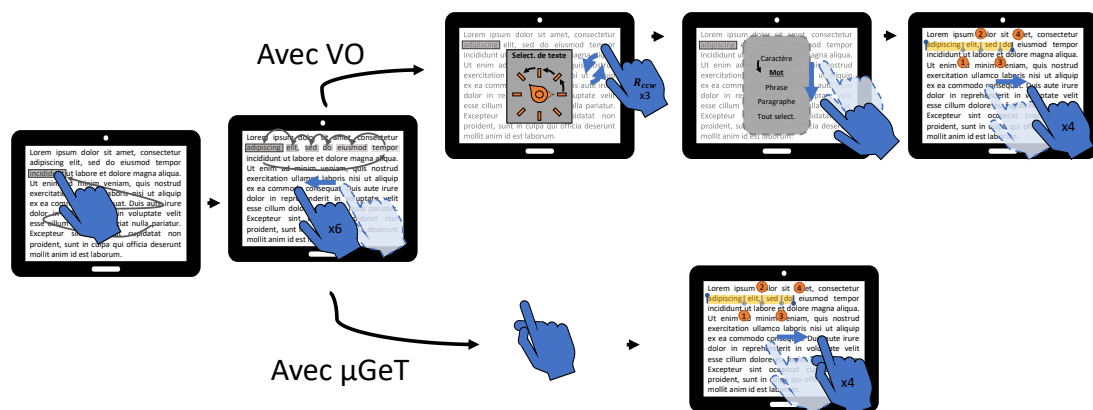


FIGURE 3.8 – Scénario de sélection de texte, de gauche à droite. Les deux premières étapes sont communes aux deux techniques : pointage direct sur un mot, puis ajustement précis de la sélection. Les étapes de la ligne supérieure concernent la technique VO : positionnement du rotor sur "Sélection de texte", puis position du sous-menu sur "Mot", puis déplacement de la poignée de sélection à la position souhaitée à l'aide SWIPE horizontaux. Les étapes de la ligne inférieure concernent la technique μ GeT : contact du majeur avec le pouce pour former un μ Geste TTF, puis déplacement de la poignée de Fin à la position souhaitée à l'aide de SWIPE horizontaux.

Nous proposons une technique d'interaction alternative, μ Get, combinant l'écran tactile et 2 μ Gestes TTF issus de l'ensemble de 8 identifiés dans le chapitre précédent : un TAP maintenu sur l'index et un TAP maintenu sur le majeur. Aucun autre μ Geste TTF n'est reconnu. Métaphoriquement, l'index et le majeur représentent les poignées de sélection de texte (appelées ancre et poignée de sélection dans la technique VO décrite ci-dessus). Lorsque les utilisateurs touchent et maintiennent leur pouce appuyé sur le doigt approprié, ils "saisissent" la poignée de sélection la plus à gauche ou la plus à droite (appelées respectivement poignée de **Début** et de **Fin**, en bleu dans l'exemple précédent $\downarrow\uparrow$). Ces μ Gestes TTF de saisie (appelés μ TAP ci-après) déclenchent un quasi-mode (c'est-à-dire un mode actif uniquement tant que l'action de déclenchement est maintenue) dans lequel Abigaïl peut modifier la position de la poignée sélectionnée (du Début avec l'index, de Fin avec le majeur). Une fois que l'une des poignées est saisie, elle peut la déplacer avec trois niveaux de granularité : 1) caractère par caractère (CbC), à l'aide de courts balayages horizontaux effectués sur l'écran (SWIPE) dont la longueur est inférieure à 150 pixels ; 2) mot par mot (WbW), à l'aide de longs SWIPE horizontaux dont la

3.2. ÉTUDES EXPÉRIMENTALES

longueur est supérieure à 150 pixels ; 3) phrase par phrase (SbS), à l'aide de SWIPE verticaux.

De manière similaire à VO, les utilisateurs effectuent le placement initial des poignées en mettant en surbrillance une entité en la pointant sur l'écran, plaçant ainsi les deux poignées au début du mot survolé. Avant toute sélection (c'est-à-dire lorsque aucune poignée n'a encore été saisie et déplacée), des SWIPE horizontaux (de longueur indifférenciée) peuvent être utilisés pour ajuster le placement initial des poignées avec une granularité unique "WbW". Si une sélection a déjà été démarrée, le pointage et les SWIPE sur l'écran (si aucune poignée n'est sélectionnée à l'aide de μ Geste TTF) n'ont aucun effet.

Tout comme avec VO, chaque changement (quasi-mode, sélection mise à jour) déclenche un retour audio, et un DOUBLE TAP sur l'écran permet d'écouter la sélection actuelle, le cas échéant, ou d'arrêter la rétroaction audio en cours.

Comparaison théorique

L'hypothèse principale motivant la conception de la technique multimodale μ Get est que la navigation dans les menus à l'aide de VO allonge le temps de réalisation, complique la trajectoire d'interaction et plus généralement, interrompt la tâche en cours. Pour rappel, une "trajectoire d'interaction" se réfère à la succession d'actions accomplies par l'utilisateur lorsqu'il interagit avec un dispositif ou une interface pour parvenir à un état donné du système (voir Chapitre 1.6.1). Nous avons spécifiquement conçu μ GeT pour éviter les interruptions dues à la navigation dans les menus en ayant tous les outils de sélection disponibles en accès direct à l'aide des μ Geste TTF.

L'accès à des participants PDV étant limité (voir chapitre 3 et [39]), nous avons souhaité dans un premier temps valider théoriquement notre hypothèse avant de la tester empiriquement. Nous avons modélisé les tâches à l'aide de l'outil de modélisation empirique Storyboard Empirical Modeling (StEM) [81], une extension du Finger-Level Model (FLM) [123] lui-même basé sur le Keystroke-Level Model (KLM) [45]. StEM nous permet de prédire les temps de réa-

3.2. ÉTUDES EXPÉRIMENTALES

lisation d'une tâche : nous l'utilisons pour comparer les trajectoires d'interaction optimales en utilisant les deux techniques de sélection de texte (μ GeT et VO).

Dans μ GeT, nous avons choisi de remplacer les modes par des quasi-modes (c'est-à-dire mode maintenu en place par une action constante de l'utilisateur, ici en maintenant un μ Tap sur l'index ou le majeur), et d'interpréter tout geste tactile sur la tablette comme manipulation explicite de la sélection, éliminant ainsi le besoin de mémoriser les états et les chemins de navigation dans les menus.

Tâche	Cible & point de départ des marqueurs ↓↑	μ GeT	VO (actions) de navigation dans le menu	Delta (μ GeT - VO)	VO : Ratio menu/total	μ GeT temps (s)	VO temps (s)	Delta temps (s)
milieu-de-mot	μ GeT: ↓↑ Douze VO: D↓↑ ouze	$\mu T_{mid} F_{long}^{\rightarrow} F_{short}^{\leftarrow}$ $\mu T_{ind} F_{short}^{\leftarrow}$	(avec rotor en position #3) $R_{ccw} R_{ccw} R_{ccw} F^{\downarrow} F^{\rightarrow} F^{\rightarrow} F^{\rightarrow}$	8 - 14 - 6 actions	4/14 29%	4,2	6,2	-2,0
(a) 2-mots-et-demi	↓↑ Six Sept Huit	μT_{mid} $F_{long}^{\rightarrow} F_{long}^{\rightarrow} F_{long}^{\rightarrow} F_{short}^{\leftarrow}$	(avec rotor en position #4) $R_{ccw} R_{ccw} R_{ccw} R_{ccw}$ $F^{\rightarrow} F^{\rightarrow} F^{\rightarrow} F^{\downarrow} F^{\leftarrow}$	9 - 18 - 9 actions	5/18 28%	4,9	8,0	-3,1
(b) 2-mots-et-demi	μ GeT: ↓↑ Trois Quatre Cinq VO: Tr↓↑ ois Quatre Cinq	$\mu T_{mid} F_{long}^{\rightarrow} F_{long}^{\rightarrow} F_{long}^{\rightarrow}$ $\mu T_{ind} F_{short}^{\leftarrow} F_{short}^{\leftarrow}$	(avec rotor en position #3) $R_{ccw} R_{ccw} R_{ccw} F^{\rightarrow} F^{\rightarrow} F^{\rightarrow}$	12 - 12 +0 actions	3/12 25%	6,4	5,3	+1,1
4-words	↓↑ Neuf Dix Onze Douze	$\mu T_{mid} F_{long}^{\rightarrow} F_{long}^{\rightarrow}$ $F_{long}^{\rightarrow} F_{long}^{\rightarrow}$	(avec rotor en position #4) $R_{ccw} R_{ccw} R_{ccw} R_{ccw}$ $F^{\rightarrow} F^{\rightarrow} F^{\rightarrow} F^{\rightarrow}$	9 - 16 - 9 actions	4/16 25%	5,1	7,1	-2,0
phrase	↓↑ Un Deux [...] Trente.	$\mu T_{mid} F^{\downarrow}$	(avec rotor en position #4) $R_{ccw} R_{ccw} R_{ccw} R_{ccw} F^{\downarrow} F^{\rightarrow}$	3 - 12 - 9 actions	5/12 42%	1,6	5,3	-3,8
(a) 2-mi-phrases	↓↑ trois Vingt quatre [...] Trente. Trente et un Trente deux	$\mu T_{mid} F^{\downarrow}$ $F_{long}^{\rightarrow} F_{long}^{\rightarrow} F_{long}^{\rightarrow} F_{long}^{\rightarrow}$	(avec rotor en position #4) $R_{ccw} R_{ccw} R_{ccw} R_{ccw}$ $F^{\downarrow} F^{\downarrow} F^{\downarrow} F^{\downarrow} F^{\rightarrow} F^{\rightarrow} F^{\rightarrow} F^{\rightarrow}$	11 - 24 - 13 actions	7/24 29%	6,0	10,6	-4,7
(b) 2-mi-phrases	↓↑ huit Vingt neuf Trente. Trente et un [...] Trente sept	$\mu T_{mid} F^{\downarrow} (F_{long}^{\rightarrow} \times 13)$	(avec rotor en position #4) $R_{ccw} R_{ccw} R_{ccw} R_{ccw}$ $F^{\downarrow} F^{\rightarrow} F^{\downarrow} (F^{\rightarrow} \times 13)$	29 - 40 - 11 actions	6/40 15%	15,9	17,7	-1,8

FIGURE 3.9 – Modélisation StEM des trajectoires d'interaction optimales utilisant GeT et VO pour chaque tâche, ainsi que le temps prédit correspondant. Les entrées utilisées pour naviguer dans les menus à l'aide de VO sont indiquées en orange.

La figure 3.9 montre une trajectoire d'interaction optimale pour chaque tâche. Nous considérons comme optimale la trajectoire d'interaction utilisant le moins d'entrées possible pour un point de départ donné. Il existe plusieurs optima : en effet un optimum est calculé pour un point de départ donné. Dans la figure 3.9, pour μ GeT, toutes les poignées sont placées au début du premier mot contenant la cible, et pour VO, toutes les poignées sont placées au début de la cible elle-même (y compris à l'intérieur d'un mot, car VO permet un placement initial précis des poignées).

3.2. ÉTUDES EXPÉRIMENTALES

Pour chaque tâche modélisée, nous utilisons l'ensemble des opérateurs StEM suivants (les temps indiqués sont fournis par l'outil [81]) :

- T (tapping), appuyer sur une cible à l'écran depuis les airs ;
- R (rotation), geste de rotation à l'écran avec deux doigts ;
- F (flicking), un mouvement linéaire balistique dans l'une des directions cardinales (haut, bas, gauche, droite). Cet opérateur correspond aux gestes SWIPE décrits dans ce document.

Cependant, afin de comparer les deux techniques, nous avons fait trois hypothèses.

Premièrement, l'outil StEM s'applique uniquement à l'interaction tactile, nous avons donc ajouté un quatrième opérateur (μ T) qui représente un Tap du pouce sur l'index ou le majeur. Nous avons supposé que ce geste était similaire à l'opérateur StEM Tap classique (T) en terme de temps. Par ailleurs, nous avons modélisé les μ T en séquence par rapport aux autres opérateurs, alors qu'ils pourraient être effectués en parallèle.

Deuxièmement, le geste pour naviguer dans le menu rotor à l'aide de VO est différent du geste de rotation (R) décrit dans le StEM. Le geste de rotation dans StEM est une entrée précise destinée à contrôler un paramètre (par exemple, l'angle de rotation d'un objet), tandis que dans notre situation, il s'agit d'une "rotation balistique" qui n'a pas besoin d'être finement contrôlée ou ciblée. Nous avons donc considéré ce geste comme équivalent à un geste de "flick" (F) en terme de temps.

Troisièmement, μ GeT implique deux types de "SWIPES" horizontaux : courts et longs. Un mouvement linéaire d'une longueur de moins de 150 pixels est considéré comme un "SWIPE" court. Comme ni le StEM ni le FLM ne font la distinction entre les deux, nous avons compté le temps des "SWIPES" longs comme étant deux fois celui des "SWIPES" réguliers. Pour résumer, nous avons utilisé les temps d'opérateur suivants :

3.2. ÉTUDES EXPÉRIMENTALES

- 669 ms pour l'opérateur T (puisque aucune localisation précise à l'écran n'est nécessaire, nous avons utilisé l'amplitude maximale pour dériver le temps de l'opérateur) ;
- 669 ms pour l'opérateur μT ;
- 216 ms pour les opérateurs F et F_{court} ;
- 432 ms pour l'opérateur F_{long} ;
- 216 ms pour l'opérateur R.

Puisque VO et μGeT utilisent tous deux des séquences de gestes balistiques de "flick" et "rotation", qui commencent et terminent leur mouvement dans l'air, ces deux opérateurs devraient normalement être modélisés par la séquence d'opérateurs TR (Tapping puis Rotation) ou TF (Tapping puis Flicking). Afin de simplifier la lecture de la figure 3.9 toutefois, nous utilisons simplement R pour TR et F pour TF, et incluons le temps de l'opérateur T dans cette notation.

Les résultats théoriques suggèrent que μGeT est plus rapide (à l'exception de la tâche (*b*) *2-mots-et-demi*) et permet des trajectoires plus courtes sur l'ensemble des tâches de sélection 3.9. Afin de confirmer ces résultats théoriques, nous avons mené une étude expérimentale.

Protocole expérimental : Comparaison expérimentale de la technique conçue avec des microgestes à la technique actuelle avec un lecteur d'écran

L'étude examine, pour une tâche de sélection de texte en situation de non-retour visuel, si μGeT est plus rapide, nécessite moins d'entrées et obtient une meilleure évaluation subjective par rapport à VO, la technique de référence.

Neuf volontaires PDV et huit volontaires voyants (dont 6 femmes : 4 PDV, 2 voyantes), âgés de 24 à 58 ans (moyenne de 33 ans, écart type de 9,1), ont participé à cette expérience. Dans les paragraphes suivants, nous nous référerons à ces participants de P1 à P9 pour les PDV et de P10 à P17 pour les participants voyants.

Tous les participants étaient droitiers et portaient un bandeau sur les yeux pendant l'expé-

3.2. ÉTUDES EXPÉRIMENTALES

rience. Aucun des participants n'a signalé d'autres handicaps mentaux ou perceptuels que la déficience visuelle. Tous utilisent quotidiennement un appareil tactile ; 8 des 9 PDV possèdent un appareil iOS et ont déjà utilisé VO. Aucun des participants voyant n'avait jamais utilisé VO. Aucun participant n'avait d'expérience avec les μ Gestes TTF.

Nous avons manipulé les mains des participants pour leur montrer comment effectuer les gestes nécessaires à l'expérience, puis les participants ont suivi une phase d'entraînement. Lorsqu'ils se sont sentis prêts, nous leur avons donné pour instruction de sélectionner une section de texte aussi rapidement et précisément que possible dans une vue de la taille d'une tablette affichant trois phrases de pseudo-texte sur une plage de treize lignes. Pour évaluer uniquement le processus de sélection et non la mémorisation du texte, nous avons utilisé des chiffres de un à soixante-deux, écrits dans l'ordre en lettres complètes (c'est-à-dire "Un Deux Trois [...] Soixante Soixante et un Soixante deux."). Chaque mot est compté séparément (c'est-à-dire "vingt et un" compte comme trois mots). Le pseudo-texte résultant a été divisé en trois phrases : "Un" à "Trente", "Trente et un" à "Cinquante quatre", "Cinquante cinq" à "Soixante deux").

Chaque participant a accompli un bloc de cinq tâches de sélection de texte par technique, soit deux blocs au total. L'ordre des techniques a été contre-balançé entre les participants. Chaque tâche impliquait un type de cible différent. Entre les blocs de la première et de la deuxième technique, les types de cibles étaient les mêmes mais à des emplacements différents dans le texte, de sorte que deux tâches ne soient jamais identiques pour un participant donné. Ces types de cibles étaient inspirés d'une étude précédente sur la sélection de texte [82] et étaient destinés à couvrir diverses situations. Les types de cibles étaient les suivants :

- 4-mots : sélectionner quatre mots consécutifs ;
- phrase : sélectionner une phrase complète ;
- 2-mots-et-demi : sélectionner deux mots et les trois caractères suivants, ou sélectionner les trois derniers caractères d'un mot et les deux mots suivants ;
- 2-mi-phrases : sélectionner les quatre derniers mots d'une phrase puis treize mots, ou

3.2. ÉTUDES EXPÉRIMENTALES

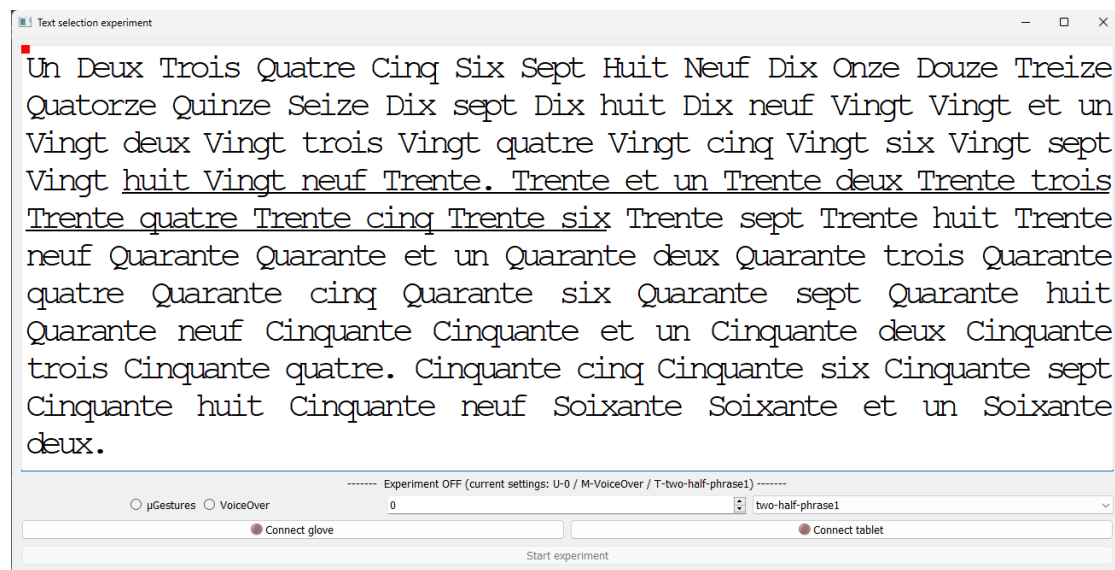


FIGURE 3.10 – Capture d'écran de l'interface à destination de l'expérimentateur. Le point rouge (en haut à gauche) indique en temps réel sur la tablette l'endroit pointé par le participant. Le texte souligné indique la cible. L'interface indique aussi en temps réel le texte mis en surbrillance ou sélectionné (non présent sur cette capture d'écran).

sélectionner les treize derniers mots d'une phrase puis les quatre premiers mots de la suivante ;

— milieu-de-mot : sélectionner 3 caractères au milieu d'un mot de 5 ou 6 caractères.

L'objectif de cette série de tâches basées sur des types différents de cibles de sélection était d'évaluer l'utilisabilité des deux techniques dans des contextes plus ou moins exigeants. L'ordre des types de cibles était toujours le même : de "facile" à "difficile" (le même ordre dans lequel les types de cibles de sélection sont listés ci-dessus). Nous avons délibérément choisi d'augmenter progressivement la difficulté pour faciliter l'apprentissage de nos participants, avec l'objectif de maintenir la durée d'une session inférieure à une heure - la disponibilité des participants PDV étant limitée, chaque créneau avait été planifié et ne pouvait dépasser au delà du temps imparti. Toutefois, certains participants PDV ont nécessité d'avantage de temps d'installation, d'explication et/ou d'entraînement, ce qui a empiété sur le temps d'expérimentation. De fait, certains PDV ont pu faire toutes les tâches, d'autres non. Pour pallier à ce problème, nous n'avons utilisé dans les résultats que les 3 premières tâches de chacun des deux blocs (μ Get et VO) : 4-mots, phrase et 2-mots-et-demi.

3.2. ÉTUDES EXPÉRIMENTALES

Pour chaque tâche, la cible était lue à haute voix par l'expérimentateur au début de chaque essai. Les participants pouvaient demander à l'expérimentateur de répéter la cible et la structure du texte (début et fin des phrases) autant de fois que nécessaire pendant l'essai. Un essai se terminait lorsque le participant spécifiait à l'expérimentateur qu'il avait terminé la sélection. Si la cible n'était pas sélectionnée, était partiellement sélectionnée, ou si du texte n'appartenant pas à la cible était sélectionné, l'essai était compté comme une erreur. Aucun retour sur le succès ou non de la tâche n'était donné. L'expérimentateur commençait le prochain essai à la demande du participant. Les participants pouvaient faire une pause quand ils le souhaitaient entre les essais.

Une fois que le premier bloc est terminé, les participants ont répondu à un questionnaire Raw Nasa-TLX (afin d'évaluer la charge de travail perçue) et à un questionnaire UMUX-LITE (afin d'évaluer l'utilisabilité générale) sur la technique utilisée dans le bloc (questionnaires disponibles en annexe A). Une fois que le deuxième bloc est terminé et que les questionnaires sont complétés, nous avons demandé aux participants quelle technique ils préféraient et pourquoi. L'expérimentateur a conclu l'expérience en recueillant des commentaires spontanés et en répondant aux questions.

Implémentation

L'expérience a été menée sur une tablette de 10,1 pouces (Huawei MediaPad T5 Lite, 24*15,5 cm), posée à plat sur une table devant les participants (figure 3.11). La tablette exécutait une page web vide en plein écran qui capturait et reliait tous les événements tactiles, en WiFi et via des websockets, à un ordinateur portable. Nous avons utilisé une correspondance absolue entre l'écran de la tablette et la zone de texte sur laquelle les tâches étaient effectuées. Dans notre expérience, nous avons mis en œuvre μ GeT ainsi qu'une réimplémentation de la technique VO, permettant un meilleur contrôle des données enregistrées.

Pour capturer les μ Gestes TTF effectués, nous avons utilisé le concept de gant modulaire expliqué au paragraphe 3.1.1. Le gant est composé de deux bagues en tissu élastique sur

3.2. ÉTUDES EXPÉRIMENTALES



FIGURE 3.11 – Photographie illustrant l'utilisation de la tablette avec la technique VoiceOver.

lesquelles sont brodés un seul connecteur en fil et en tissu conducteur (figure 3.12), chacun reliés à une entrée de la carte capacitive. Chaque surface conductrice mesure 2 cm sur 2 cm pour couvrir la phalange médiane de l'index et du majeur. Ces manchons permettent de détecter les 2 μ Gestes TTF nécessaires à l'expérimentation.

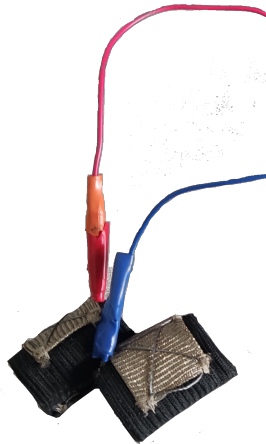


FIGURE 3.12 – Photographie des manchons utilisés dans le troisième prototype (seconde itération du gant modulaire, voir 3.1.1)

À chaque événement, le programme de la carte Arduino envoie à l'ordinateur une liste de booléens (0 et 1), chacun correspondant à l'état d'un connecteur (absence ou non de contact). La carte Arduino relaye les contacts de pression et de relâchement de chaque capteur à l'ordinateur via une connexion série filaire. Sur l'ordinateur, un logiciel écrit en Python gère l'expé-

3.2. ÉTUDES EXPÉRIMENTALES

rience, reçoit et traite les données à la fois des bagues et de la tablette, et affiche une fenêtre sur un écran, permettant à l'expérimentateur de surveiller la progression.

Résultats

Pour comparer l'efficacité de VO et de μ GeT, nous nous concentrons principalement sur le temps de sélection, qui s'étend de la fin de la phase de placement des poignées de sélection de texte à la fin de l'essai. De plus, nous comparons également le temps de pré-sélection (qui s'étend du premier contact sur l'écran ou du premier μ Geste TTF enregistrée jusqu'au début de la phase de sélection), le nombre d'actions par essai, les erreurs, la charge cognitive, l'expérience utilisateur et la préférence de l'utilisateur.

Étant donné que les participants PDV ont accompli 2 tâches de moins que les participants voyants, nous présentons leurs résultats séparément. Nous rapportons les médianes et les plages interquartiles (IQR, InterQuartile Range) pour les données continues (c'est-à-dire le temps de sélection, le temps de pré-sélection, le nombre d'actions), et utilisons des tests de Wilcoxon signés car les données collectées ne suivent pas une distribution normale, et rapportons la taille de l'effet de corrélation bisériale de rang (r). Pour les données non continues (c'est-à-dire la charge cognitive et l'expérience utilisateur), nous rapportons les moyennes et les écarts-types (SD, Standard Deviation). Nous utilisons des tests U de Mann-Whitney et rapportons la taille de l'effet de corrélation bisériale de rang (r). Nous utilisons des tests Z à deux proportions pour les données binaires (c'est-à-dire le nombre d'erreurs). Nous avons également étudié en détail toutes les trajectoires d'interaction. Nous rapportons un résumé des stratégies que nous avons observées ainsi que des explications potentielles pour les comportements qui se démarquent. Nous utilisons des tests Z à deux proportions à cette fin.

Au total, $9 \text{ PDV} \times 3 \text{ tâches} = 27 \text{ essais}$, et $8 \text{ voyants} \times 5 \text{ tâches} = 40 \text{ essais}$ ont été effectués par technique d'interaction. Nous avons exclu un essai des résultats des voyants (P11, condition VO, tâche "phrase") car le participant pensait avoir terminé l'essai, mais n'avait en réalité rien sélectionné.

3.2. ÉTUDES EXPÉRIMENTALES

Donnée	Statut	Technique	Médiane	IQR	Test statistique	Taille de l'effet
Temps de sélection (s)	PDV	μ GeT	11,0	22,5	Wilcoxon signé $p = 0,014^*$	$r = 0,534$
		VO	21,6	22,7		
	Participants voyants	μ GeT	21,7	22,3	Wilcoxon signé $p = 0,548$	$r = 0,113$
		VO	20,5	15,2		
Temps de pré-sélection (s)	PDV	μ GeT	29,2	34,6	Wilcoxon signé $p = 0,162$	$r = 0,312$
		VO	42,0	58,7		
	Participants voyants	μ GeT	21,0	18,5	Wilcoxon signé $p = 0,239$	$r = 0,265$
		VO	31,9	20,0		
# d'actions par essai	PDV	μ GeT	8	6,5	Wilcoxon signé $p = 0,012^*$	$r = 0,555$
		VO	14	10		
	Participants voyants	μ GeT	11	12	Wilcoxon signé $p = 0,24$	$r = 0,265$
		VO	12	4,8		
# d'erreurs	PDV	μ GeT	0 (total)	/	Z-Test à 2 proportions $p = 0,004^*$	/
		VO	7 (total)	/		
	Participants voyants	μ GeT	1 (total)	/	Z-Test à 2 proportions $p = 0,545$	/
		VO	2 (total)	/		

FIGURE 3.13 – Tableau synthétisant les résultats pour le temps de sélection, le temps de pré-sélection, le nombre d'action par essai et le nombre d'erreur totales, pour les deux populations de participants (PDV et voyants).

Analyse quantitative Le tableau 3.13 résume l'analyse du temps de sélection et du temps de pré-sélection. Comme il s'agit de notre principal objectif, nous avons d'abord vérifié quelle technique d'interaction avait un temps de sélection plus rapide.

Pour les PDV, la médiane du temps de sélection est plus courte pour μ GeT (11 secondes, IQR 22,5) que pour VO (21,6s, IQR 22,7). Un test de Wilcoxon montre une différence significative entre les deux techniques pour les PDV ($p = 0,014$, $r=0,534$). Les temps de sélection varient considérablement selon le type de tâche (figure 3.14). Pour les PDV, un test de Wilcoxon montre des différences significatives de temps de sélection entre les deux techniques pour les tâches "4-mots" ($p=0,027$, $r=0,822$) et "phrase" ($p=0,012$, $r=0,911$) en faveur de μ GeT (figure 3.14). La médiane du temps de pré-sélection est plus courte pour μ GeT que pour VO (PDV : 29,2s μ GeT contre 42s VO - Voyants : 21 s μ GeT contre 31,9s VO). Cette observation est valable lorsque l'on décompose les données par type de tâche (figure 3.14). Cependant, les tests de Wilcoxon ne montrent pas de différence significative.

3.2. ÉTUDES EXPÉRIMENTALES

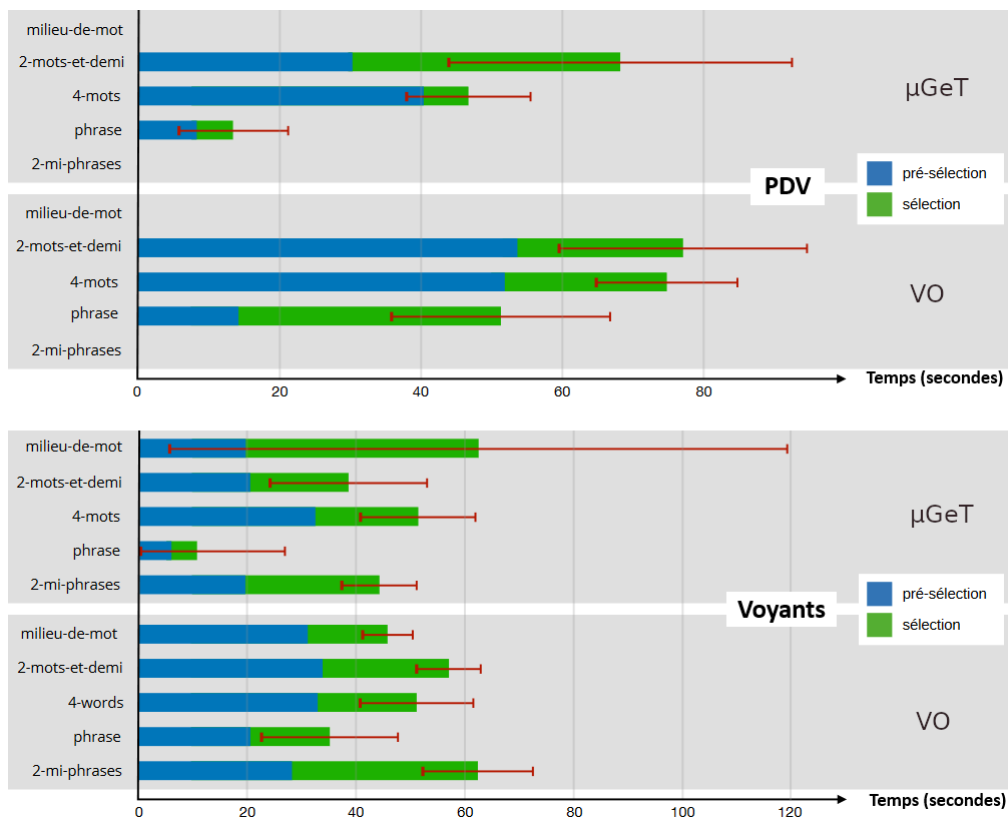


FIGURE 3.14 – Graphique montrant le temps de sélection (en vert) et de pré-sélection (en bleu) avec l'IQR (en rouge) pour chaque technique (μ GeT et VO), organisé type de tâche et par population de participants (PDV et voyants).

Le nombre médian d'entrées est plus faible pour μ GeT que pour VO (8 μ GeT contre 14 VO) (tableau résultats 01). Un test de Wilcoxon montre une différence significative entre les deux techniques ($p=0,012$, $r=0,555$). Lorsqu'on les décompose par type de tâche, les tests montrent une différence significative pour les tâches "4-mots" ($p=0,024$, $r=0,866$) et "phrase" ($p=0,028$, $r=0,844$) en faveur de μ GeT.

En ce qui concerne le nombre d'erreurs, seuls 7 des 54 essais des PDV n'ont pas abouti à une sélection de texte correspondant à la cible, toutes ces erreurs ont été faites avec VO. Un test Z à deux proportions montre une différence significative ($p=0,004$) entre les deux techniques.

3.2. ÉTUDES EXPÉRIMENTALES

Nous avons mesuré la charge cognitive perçue et l'expérience utilisateur pour chaque technique. Le tableau 3.15 montre les scores subjectifs médians par technique et l'IQR pour chacun des 6 éléments du questionnaire Nasa-TLX (c'est-à-dire la demande mentale, la demande physique, la demande temporelle, la performance globale, l'effort global, le niveau de frustration) et les 2 éléments du questionnaire UMUX-Lite (c'est-à-dire l'utilité et la convivialité). Pour chaque élément des deux questionnaires, les tests U de Mann-Whitney ne montrent pas de différence significative entre les deux techniques. Enfin, la préférence utilisateur n'a révélé aucune différence, les deux techniques ayant été appréciées au même niveau (figure 3.15).

Donnée	Statut	Technique	Médiane	IQR	Test statistique	Taille de l'effet
Exigence mentale	PVI	μ GeT	3	2	Mann-Whitney U p= 0,653	r=0,116
		VO	3	3		
	Participants voyants people	μ GeT	6	1,5	Mann-Whitney U p= 0,825	r=0,078
VO		6	1,25			
Exigence physique	PVI	μ GeT	1	1	Mann-Whitney U p= 0,083	r=0,457
		VO	2	0		
	Participants voyants people	μ GeT	2	2,5	Mann-Whitney U p= 0,870	r=0,062
VO		2	2,25			
Exigence temporelle	PVI	μ GeT	2	1	Mann-Whitney U p= 0,778	r=0,086
		VO	2	2		
	Participants voyants people	μ GeT	3,5	1,5	Mann-Whitney U p= 0,703	r=0,125
VO		3,5	1,25			
Performance	PVI	μ GeT	5	2	Mann-Whitney U p= 0,745	r=0,099
		VO	5	3		
	Participants voyants people	μ GeT	6	1,25	Mann-Whitney U p= 0,867	r=0,062
VO		6	0,5			
Effort	PVI	μ GeT	2	2	Mann-Whitney U p= 0,855	r=0,062
		VO	2	1		
	Participants voyants	μ GeT	5	2	Mann-Whitney U p= 1	r=0,016
VO		5	1,25			
Frustration	PVI	μ GeT	1	2	Mann-Whitney U p= 0,660	r=0,124
		VO	1	1		
	Participants voyants	μ GeT	4	3	Mann-Whitney U p= 0,704	r=0,126
VO		4	3,25			
Utilité	PVI	μ GeT	7	1	Mann-Whitney U p= 0,164	r=0,370
		VO	6	2		
	Participants voyants	μ GeT	5,5	1	Mann-Whitney U p= 0,234	r=0,344
VO		5	0,5			
Utilisabilité	PVI	μ GeT	6	2	Mann-Whitney U p= 0,238	r=0,321
		VO	7	1		
	Participants voyants	μ GeT	6	2	Mann-Whitney U p= 0,357	r=0,281
VO		5,5	1,25			
Technique préférée	PVI	μ GeT / VO	5 (56%) / 4 (44%) (N=9)		/	/
	Participants voyants	μ GeT / VO	4 (50%) / 4 (50%) (N=8)		/	/

FIGURE 3.15 – Tableau synthétisant les résultats des questionnaire Raw Nasa-TLX et UMUX-Lite, pour les deux populations de participants (PDV et voyants).

Pour les participants voyants, la médiane du temps de sélection est plus courte pour VO (20s, IQR 15,1) que pour μ GeT (21,7s, IQR 22,3) Un test de Wilcoxon ne montre pas de différence significative entre les deux techniques pour les participants voyants (p=0,548). Un test de

3.2. ÉTUDES EXPÉRIMENTALES

Wilcoxon montre des différences significatives de temps de sélection entre les deux techniques pour les tâches "milieu-de-mot" ($p=0,008$, $r=1$), en faveur de VO (figure 3.14). La médiane du temps de déplacement est plus courte pour μ GeT que pour VO (21s μ GeT contre 31,9s VO). Cette observation est valable lorsque l'on décompose les données par type de tâche (figure 3.14). Cependant, les tests de Wilcoxon ne montrent pas de différence significative.

Le nombre médian d'entrées est plus faible pour μ GeT que pour VO (11 μ GeT contre 12 VO) (Tableau 3.13). Un test de Wilcoxon ne montre pas de différence significative entre les deux techniques ($p=0,582$). Lorsqu'on les décompose par type de tâche, les tests montrent une différence significative pour la tâche "milieu-de-mot" ($p=0,022$, $r=1$) en faveur de VO. En ce qui concerne le nombre d'erreurs, 3 des 79 essais étaient erronés : 2 avec VO et 1 avec μ GeT. Un test Z à deux proportions ne montre pas de différence significative ($p=0,545$) entre les deux techniques.

Analyse des trajectoires d'interaction Pour mieux comprendre les différences de performance entre les deux techniques, nous avons utilisé une analyse thématique pour dériver les types de comportements de l'ensemble des trajectoires d'interaction qui ont conduit les participants à s'écarter des trajectoires optimales. Nous rappelons qu'une trajectoire optimale consiste à utiliser le moins d'entrées possible pour un point de départ donné, comme défini dans la Section 3.2.3. Il existe 3 principaux types : **désorientation**, **inertie** et **faute**. La figure 3.16 résume la classification de tous les essais.

Désorientation regroupe les comportements qui retardent la tâche de sélection car les utilisateurs prennent des étapes supplémentaires pour comprendre l'état du système. Les comportements observés sont :

- Les utilisateurs déclenchent des commentaires audio supplémentaires et prennent du temps pour écouter la sélection de texte actuelle.
- Les utilisateurs font des pauses ou hésitent (de 3 à 15 secondes) entre deux actions.
- Les utilisateurs déplacent à plusieurs reprises une poignée de sélection en arrière et en avant pour comprendre le texte localement.

3.2. ÉTUDES EXPÉRIMENTALES

- Les utilisateurs tentent de faire traverser une poignée de sélection avec l'autre poignée, ce qui n'est pas possible avec μ GeT (par exemple, faire dépasser par la gauche la poignée de fin au delà de l'emplacement de la poignée de début). Comme il n'y a pas de commentaire audio pour ce cas spécifique, les utilisateurs peuvent passer du temps à essayer de comprendre activement ce qui se passe avant de continuer la tâche.

Inertie regroupe les comportements qui retardent la tâche de sélection car les utilisateurs adoptent une stratégie sous-optimale. Les comportements observés sont :

- Les utilisateurs naviguent dans les menus VO, qui bouclent sur eux-même, dans la direction "incorrecte" (par exemple, passant du 1^{er} au 8^{ème} élément de rotor dans le sens des aiguilles d'une montre (7 actions), au lieu du sens inverse des aiguilles d'une montre (1 action)).
- Les utilisateurs effectuent une séquence valide de sélection mais de manière sous-optimale (par exemple, sélectionner une phrase mot par mot au lieu d'utiliser le mode de sélection "phrase").

Faute regroupe les comportements qui entravent largement la tâche de sélection et imposent des étapes de correction supplémentaires pour accomplir la tâche. Les comportements observés sont :

- Les utilisateurs gâchent leur sélection actuelle en (dé)sélectionnant une grande partie du texte, les obligeant à refaire des sélections de granularité plus fine.
- Les utilisateurs confondent les commandes (par exemple, effectuant un geste de rotation sur l'écran au lieu d'un SWIPE, déplaçant le menu rotor au lieu du sous-menu).
- Les utilisateurs dépassent leur sélection ou le menu d'éléments.

Ces erreurs ne sont pas mutuellement exclusives ; plusieurs d'entre elles peuvent être commises au cours d'un essai, parfois simultanément (par exemple, lors de l'utilisation de μ GeT, dépasser d'un caractère lors de la sélection puis revenir en arrière mais avec une granularité plus importante, désélectionnant alors tout le mot).

Dans ce qui suit, nous discutons des résultats pour les participants PDV et les personnes voyantes séparément.

3.2. ÉTUDES EXPÉRIMENTALES

	Statut	Technique	Compte	% d'occurrence	Test
Trajectoires optimales	PDV	μGeT	12	44%	Z-Test à 2 proportions p= 0,048
		VO	5	19%	
	Participants voyants	μGeT	10	25%	Z-Test à 2 proportions p= 0,080
		VO	4	10%	
# de déviations	PDV	μGeT	42	155%	/
		VO	51	188%	
	Participants voyants	μGeT	104	260%	/
		VO	58	145%	
Désorientation : Retours audio supplémentaires	PDV	μGeT	3	11%	Z-Test à 2 proportions p=0,329
		VO	1	4%	
	Participants voyants	μGeT	19	48%	Z-Test à 2 proportions p= 0,003
		VO	3	8%	
Désorientation : Pauses (> 3s)	PDV	μGeT	16	59%	Z-Test à 2 proportions p= 0,398
		VO	19	70%	
	Participants voyants	μGeT	34	85%	Z-Test à 2 proportions p= 0,020
		VO	24	62%	
Désorientation : Va et viens dans le texte	PDV	μGeT	6	22%	Z-Test à 2 proportions p= 0,049
		VO	1	4%	
	Participants voyants	μGeT	10	25%	Z-Test à 2 proportions p= 0,013
		VO	2	5%	
Désorientation : Tente de dépasser l'autre poignée (μGet)	PDV	μGeT	0	0%	/
		VO	/	/	
	Participants voyants	μGeT	4	10%	/
		VO	/	/	
Inertie : Navigation dans le mauvais sens du menu (VO)	PDV	μGeT	/	/	/
		VO	14	52%	
	Participants voyants	μGeT	/	/	/
		VO	17	44%	
Inertie : Sélectionner correctement mais avec une granularité plus faible	PDV	μGeT	1	4%	Z-Test à 2 proportions p= 0,049
		VO	6	22%	
	Participants voyants	μGeT	2	5%	Z-Test à 2 proportions p= 0,398
		VO	4	10%	
Faute : Gâche la sélection en cours	PDV	μGeT	4	15%	Z-Test à 2 proportions p= 0,036
		VO	0	0%	
	Participants voyants	μGeT	8	20%	Z-Test à 2 proportions p= 0,003
		VO	0	0%	
Faute : Se trompe de commande	PDV	μGeT	9	33% / 7%	Z-Test à 2 proportions p= 0,121
		VO	4	15%	
	Participants voyants	μGeT	15	38%	Z-Test à 2 proportions p= 0,048
		VO	7	18%	
Faute : Dépasse la cible (sélection ou menu)	PDV	μGeT	3 (sélection seulement)	11%	Z-Test à 2 proportions p= 0,276
		VO	6 (menu seulement)	22%	
	Participants voyants	μGet	12 (sélection seulement)	30%	Z-Test à 2 proportions p= 0,001 (sélection)
		VO	1 (sélection) / 3 (menu)	3% / 8%	

FIGURE 3.16 – Tableau synthétisant les résultats de l'analyse des trajectoires pour les deux populations de participants (PDV et voyants).

Participants PDV - performances globales. Dans l'ensemble, les PDV ont obtenu de meilleurs résultats avec μGeT qu'avec VO (Tableau 3.13). Lorsqu'on examine les trajectoires d'interaction, 12 sur 27 (44%) sont optimales avec μGeT, tandis qu'avec VO, seules 5 sur 27 (19%) le sont (le test Z à 2 proportions révèle une différence significative, p=0,048). Cela signifie que plus de la moitié du temps avec μGeT, et 4 fois sur 5 avec VoiceOver, les PDV s'écartent

3.2. ÉTUDES EXPÉRIMENTALES

de la trajectoire optimale.

Avec μGeT , les trajectoires sous-optimales sont principalement dues à des fautes (utilisation de la mauvaise commande - 9/27, 33%) et à la désorientation (aller-retour - 6/27, 22%). Avec VO, elles sont principalement dues à des fautes (dépassement de menu - 6/27, 22%) et à l'inertie (utilisation de niveaux de granularité sous-optimaux - 6/27, 22%). Les pauses sont très courantes dans les deux techniques (μGeT : 16/27, 59% ; VO : 19/27, 70%). Nos résultats montrent également que les PDV effectuent la sélection de texte avec une granularité sous-optimale 22% du temps avec VO (contre seulement 4% avec μGeT) (Z-test $p=0,049$), ce qui a considérablement allongé le temps de sélection (par exemple, P2 a mis 19,5 secondes pour sélectionner une phrase avec une granularité WbW).

Les PDV semblent parvenir à effectuer des trajectoires optimales plus souvent avec μGeT qu'avec VO (12/27 avec μGeT , 5/27 avec VO, Z-test $p=0,048$). Ils ont fait 42 "déviation" avec μGeT (soit 1,55 erreur par essai en moyenne) les ayant fait sortir de la trajectoire optimale et 51 avec VO (soit 1,88 erreur par essai).

Participants voyants - performances globales. Dans l'ensemble, les participants voyants ont obtenu des performances similaires avec les deux techniques (3.13). Lorsqu'on examine les trajectoires d'interaction, 10 sur 40 (25%) sont optimales avec μGeT , et seulement 4 sur 39 (10%) avec VO (Z-test $p=0,08$).

Avec μGeT , les trajectoires sous-optimales sont principalement dues à la désorientation (commentaires audio supplémentaires - 19/40, 48%, allers-retours - 10/40, 25%), et à des fautes (utilisation d'une granularité plus petite - 15/40, 38%, dépassement de la sélection - 12/40, 20%). Avec VO, elles sont principalement dues à des fautes (utilisation d'une granularité plus petite - 7/39, 18%). Aucun autre comportement n'a eu lieu plus de 10% du temps avec VO. Les pauses sont très courantes dans les deux techniques, mais plus fréquentes avec μGeT (μGeT : 34/40, 85% ; VO : 24/39, 62% ; Z-test $p=0,02$).

3.2. ÉTUDES EXPÉRIMENTALES

Les participants voyants ont commis peu d'erreurs avec le menu VO, à part en choisissant le chemin le plus long dans le menu - ce qui n'augmente pas substantiellement le temps de sélection et résulte très probablement d'un choix aléatoire. Les participants voyants semblent parvenir à effectuer des trajectoires optimales plus souvent avec μ Get qu'avec VO (10/40 avec μ Get, 4/40 avec VO, Z-test $p=0,08$). Ils ont commis 104 "déviations" avec μ GeT (soit 2,6 erreurs par essai en moyenne) les ayant fait sortir de la trajectoire optimale et 58 avec VO (soit 1,45 erreur par essai). Bien que proportionnellement, il semble que les participants voyants commettent beaucoup moins d'actions sous-optimales avec VO qu'avec μ GeT, ils obtiennent des performances équivalentes avec les deux techniques en termes de temps de sélection.

Performances par tâche. Les PDV ont obtenu de meilleurs résultats avec μ GeT dans les tâches "4-mots" et "phrase", et des résultats équivalents dans les tâches "2-mots-et-demi", tandis que les participants voyants ont obtenu des résultats similaires avec les deux techniques pour toutes les tâches, sauf les tâches "milieu-de-mot", où ils ont obtenu de meilleurs résultats avec VO.

Par ailleurs, le temps moyen par action est significativement plus court (test de Wilcoxon $p=0,020$, $r=0,867$) pour les tâches "les plus faciles" ("4-mots", 0,95 s, écart interquartile de 0,79) que pour les tâches "les plus difficiles" ("2-mots-et-demi", 2,13 s, écart interquartile de 0,62) avec μ GeT. Ce n'est pas le cas avec VO, où les trois types de tâches ont un temps moyen par action similaire.

Différences entre PDV et personnes voyantes. Avec μ GeT, les participants voyant ont demandé significativement plus de retour audio supplémentaire que les PDV (19/40 essais, 48% pour les participants voyants contre 3/27 essais, 11% pour les PDV, $p=0,0016$). Les participants voyants ont également fait significativement plus de pauses avec μ GeT qu'avec VO. Ils ont également dépassé leur sélection plus souvent avec μ GeT (12/40 essais, 30% pour μ GeT contre 1/40, 3% pour VO, Z-test $p=0,001$). Les participants voyants ont fait plus de déviations dans μ GeT par rapport aux PDV (2,6 actions sub-optimales par essai en moyenne contre 1,55 pour les PDV), mais moins dans VO (1,45 par essai contre 1,88 pour les PDV).

Discussion des résultats. Les PDV ont fait un nombre de déviation sensiblement similaire avec les deux techniques, mais ont été plus rapide avec μ GeT. Les participants voyants ont fait plus de déviations avec μ GeT qu'avec VO, mais ont été sensiblement aussi rapide avec les deux techniques. Cela pourrait suggérer que les actions sous-optimales sont moins coûteuses avec μ GeT qu'avec VO car elles sont plus faciles à corriger, et donc, qu'avec un nombre similaire de déviation, μ GeT est plus rapide que VO.

À l'exception de la tâche "phrase", nos observations diffèrent des prédictions théoriques attendues. Le modèle StEM ne tient compte que des actions motrices, ce qui signifie qu'il modélise des utilisateurs "experts", ce qui n'est pas le cas dans notre expérience. Par conséquent, les différences que nous observons pourraient révéler la différence de travail cognitif que les participants doivent effectuer pour accomplir leurs tâches. Si nous examinons les pauses et le nombre d'actions, les résultats révèlent une différence majeure entre les deux techniques. VO génère plus d'action mais semble nécessiter moins de temps de réflexion, contrairement à μ GeT. En supposant que μ GeT est théoriquement plus rapide que VO dans une trajectoire optimale pour une tâche de sélection de texte (Tableau 3.9), la différence peut s'expliquer en partie par deux facteurs :

1) la difficulté et le niveau écologique de la tâche. Les tâches "4-mots" et "phrase" peuvent être considérées comme écologiques dans le sens où elles nécessitent de sélectionner des parties de texte qui imitent une véritable sélection de texte (c'est-à-dire quelques mots, une phrase). De plus, elles peuvent être considérées comme plus faciles car elles nécessitent uniquement de réfléchir au niveau du "mot". En revanche, les tâches "2-mots-et-demi" nécessitent de sélectionner une partie d'un mot, ce qui se produit plus rarement lors de l'édition d'un texte. Elles peuvent être considérées comme plus difficiles car elles nécessitent des connaissances en orthographe pour réfléchir à deux niveaux de granularités, "mot" et "caractère". Nos résultats montrent que 8 sur 9 (89%) des participants PDV ont fait au moins une pause de 3 secondes ou plus dans les tâches "2-mots-et-demi" avec μ GeT (7 sur 9 (78%) avec VO). Seulement 4 sur 9 (44%) des participants PDV ont effectué des pauses dans chacune des deux autres tâches.

3.2. ÉTUDES EXPÉRIMENTALES

Le test Z à 2 proportions montre une différence significative ($p=0,0488$).

2) les différences intrinsèques entre les deux techniques d'interaction. Avec VO, les utilisateurs peuvent naviguer dans un menu qui leur sert de guide, car il offre un moyen explicite et structuré d'interagir avec le système. Par exemple, dans un menu comportant plusieurs éléments qui ne peuvent être navigués que de manière linéaire avec des balayages verticaux ou horizontaux, le choix de ce qu'il convient de faire ensuite se réduit à 3 options : élément précédent, élément suivant, utilisation de l'élément actuel. Comme une barre de menu dans un paradigme d'interaction graphique classique, où l'utilisateur parcourt chaque élément jusqu'à trouver le bon, puis clique dessus. De plus, l'icône ou le texte sert de rappel quant à la commande effectuée par le bouton. En revanche, μ GeT est semblable aux raccourcis clavier. Avec cette technique d'interaction, dans notre expérience, toutes les commandes sont accessibles en permanence. Le système ne fournit pas de structure explicite ni de retour d'information pour guider les utilisateurs sur ce qu'il convient de faire ou sur ce que font les commandes μ GeT. Elles doivent être connues à l'avance et mémorisées pour être utilisées efficacement. Les utilisateurs doivent essayer la commande pour voir ce qu'elle fait : par exemple, placer le pouce sur l'index déclenchera un retour audio indiquant que la première poignée peut être déplacée, mais ne dit pas quoi faire par la suite pour sélectionner un caractère (à savoir, un balayage horizontal court) ou une phrase (un balayage vertical). Cette différence n'est pas problématique pour les tâches "faciles", mais elle est exacerbée lorsque la complexité augmente. Par exemple, lorsque la granularité du mode de sélection doit être modifiée pendant la sélection, ou lorsque la tâche est moins familière.

Du reste, les participants voyants ont demandé significativement plus de retour audio supplémentaire que les PDV et ont globalement fait plus de déviations et montré des comportements indiquant de la désorientation. Cela indique que les participants voyants sont moins familiers avec les interfaces sans écran avec seulement un retour audio. Ils ont besoin de faire plus de pauses et de recevoir plus de retours audio pour former un modèle mental suffisant du système, en particulier de l'état du texte, et ce d'autant plus pour les tâches de "milieu-de-mot". C'est particulièrement le cas pour μ GeT. Parce que VO est intrinsèquement guidant, les par-

ticipants voyants n'ont pas besoin de se concentrer sur le fonctionnement du système, mais seulement sur la tâche de sélection de texte. En revanche, μ GeT est moins directif, ce qui fait que l'attention est partagée entre la compréhension du fonctionnement du système (c'est-à-dire se souvenir des commandes et de quand les utiliser) et la tâche en elle-même.

3.3 Conclusion du chapitre

Dans ce chapitre nous avons présenté trois applications à destination des PDV : une exploration de carte tactile, une tâche de couper/coller et une tâche de sélection de texte. Nous avons mené trois études expérimentales de ces applications pour évaluer l'utilité et l'utilisabilité de trois techniques d'interaction combinant l'écran tactile avec les μ Gestes TTF.

La première étude est basée sur l'exploration d'une carte audio-tactile de la Moldavie : les participants doivent explorer le document et utiliser les μ Gestes TTF pour déclencher des commandes et des retours d'informations afin de répondre à des questions. Tous les participants ont été intéressés par les possibilités d'interaction qu'offrent les μ Gestes TTF car ils trouvaient cette technique d'interaction intuitive, confortable et facile à utiliser.

La seconde étude est basée sur une tâche de réorganisation d'une grille d'icônes : les participants doivent déplacer les icônes à l'aide de commandes de couper/coller accessibles via des μ Gestes TTF. Tous les participants ont qualifié la technique d'interaction comme étant utilisable et utile, et d'ajouter qu'elle leur permettait de faire des tâches qu'ils ne faisaient pas dans la vie courante du fait de leur complexité (en l'occurrence ici, le couper/coller).

La troisième étude est basée sur une tâche de sélection de texte : les participants doivent sélectionner différentes cibles au sein d'un texte à l'aide de deux techniques d'interaction, l'une utilisant seulement l'écran tactile (VO), l'autre utilisant aussi les μ Gestes TTF (μ GeT). Cette étude montre que dans ce contexte, μ GeT surpasse VO en termes de vitesse de sélection et de nombre d'entrées dans une trajectoire d'interaction optimale. Cependant, les participants

3.3. CONCLUSION DU CHAPITRE

s'écartent souvent de ces trajectoires optimales, notamment avec μ GeT. Mais les erreurs sont rapidement corrigées avec cette technique d'interaction. En revanche, VO est moins susceptible de produire des actions sub-optimales, toutefois elle suscite plus d'inertie - avec des conséquences plus coûteuses.

En résumé, ces trois études montrent que les techniques d'interaction utilisant les μ Gestes TTF avec une interaction sur une surface tactile sont perçues de façon très positive par les PDV, car elles permettent de faciliter et de simplifier des interactions actuellement complexes avec les lecteurs d'écran

Conclusion

Dans ce chapitre de conclusion, nous souhaitons rappeler les contributions de notre recherche, discuter des limitations identifiées et esquisser des perspectives de recherche à court, moyen et à long terme.

Résumé des contributions

Ces travaux de thèse traitent la question d'une approche multimodale basée sur la combinaison des paradigmes tactile et micro-gestuel dans le but d'améliorer l'accessibilité et l'utilisabilité des dispositifs à écran tactile grand public dans un contexte sans retour visuel.

Dans un premier temps, nous avons exposé les défis inhérents aux handicaps rencontrés par un PDV lorsque celui-ci interagit avec un dispositif à écran tactile tel qu'un téléphone portable ou une tablette. Sans la vue, les PDV utilisent des logiciels dit "lecteurs d'écran" tels que VoiceOver d'iOS, qui présentent de manière audible le contenu visuel 2D de l'écran sous forme d'une liste auditive 1D d'éléments. Bien que cela permette aux PDV d'accéder au contenu numérique et d'interagir avec, chaque élément doit être traité un par un, augmentant le temps et les interactions nécessaires pour décrypter le contenu dans son entièreté. En outre, sans vision, il est difficile de cibler les éléments d'une interface graphique. Dans une telle situation, l'interaction se limite à des gestes simples et sans but, pouvant être effectués sans retour visuel, tels que des taps et des balayages directionnels - réduisant ainsi considérablement l'expressivité. Effectuer des actions "complexes" comme le copier-coller devient encore plus chronophage car les utilisateurs doivent naviguer étape par étape à travers les menus pour trouver la fonction

CONCLUSION

souhaitée.

Face à ce constat, nous avons donc exploré la possibilité de suppléer les écrans tactiles avec les microgestes (μ G) Thumb-To-Finger (TTF), des gestes effectués avec le pouce sur les doigts de la même main utilisables sans retour visuel et ayant un faible impact sur la charge cognitive.

Pour traiter le sujet, notre démarche de recherche s'articule en deux étapes. Une première étape est dédiée à la faisabilité des microgestes en étant en contact avec une surface tactile. Cette étape nous a permis d'identifier l'espace des possibles en termes de microgestes tout en interagissant avec une surface tactile. La deuxième étape a consisté à concevoir des techniques d'interaction combinant interaction tactile et microgestes pour des tâches utilisateur variées et à tester leur utilisabilité. Cette démarche de recherche en deux étapes a donné lieu à quatre études expérimentales.

La première étape s'adresse logiquement à la faisabilité des microgestes : peut-on utiliser les μ Gestes TTF en même temps que l'écran tactile ? Dans cette étude, nous présentons une expérimentation portant sur 33 μ Gestes TTF parmi les plus courants de la littérature afin d'évaluer à quel point ils sont réalisables et confortables lorsqu'ils sont effectués dans la plus contrainte des situations : alors que l'index est en contact avec une surface tactile, et ce dans une situation sans retour visuel. Nos résultats montrent que les μ Gestes TTF sont faisables dans cette situation, bien que sous certaines conditions. Notamment, les μ Gestes TTF doivent être simplifiés afin qu'il ne soit pas nécessaire de viser précisément une zone du doigt. Au final, nous identifions un ensemble de 8 μ Gestes TTF pouvant être utilisés simultanément à l'écran tactile afin d'en compléter l'usage en situation de non-retour visuel, permettant ainsi d'augmenter la bande-passante et l'expressivité de l'interaction purement tactile.

En nous reposant sur l'ensemble des μ Gestes TTF faisables, la deuxième étape a consisté à concevoir et tester des techniques d'interaction combinant interaction tactile et microgestes. Chaque technique conçue traite d'une tâche utilisateur différente : explorer un graphique 2D,

CONCLUSION

copier/coller d'un élément, sélection d'un texte. Nous avons mené trois études expérimentales des techniques conçues pour traiter de leur utilisabilité. Les questions traitées par ces études expérimentales sont les suivantes : Les μ Gestes TTF sont-ils perçus comme faciles, pratiques, utiles lorsque utilisés conjointement à l'écran tactile ? Ces études se présentent sous la forme de scénarios et placent les participants dans des situations concrètes d'utilisation d'un dispositif tactile sans retour visuel.

Le premier scénario est basé sur l'exploration d'un document audio-tactile [37, 38] composé d'une tablette et d'une surcouche en relief perceptible de façon tactile et vise à illustrer l'utilisation conjointe de l'écran tactile et des μ Gestes TTF à destination de PDV, et ce afin d'obtenir leur retour qualitatif. Les résultats de cette expérimentation montrent l'intérêt et l'utilité perçue de notre approche : tous les participants ont trouvé les μ Gestes TTF utiles, faciles à utiliser, intuitifs et confortables lorsqu'ils sont utilisés pour explorer un document audio-tactile.

Le deuxième scénario est basé sur la simplification d'un processus extrêmement courant ("copier/coller") dans une tâche de réorganisation d'icônes disposés sur une grille. Dans ce scénario l'objectif est d'utiliser les μ Gestes TTF comme une commande supplémentaire afin d'éviter des étapes de navigation dans un menu. Les résultats de cette étude indiquent que les μ Gestes TTF sont perçus comme faciles, rapides à utiliser et permettent d'effectuer des commandes courantes que la plupart des PDV n'utilisent ou ne connaissent pas au quotidien du fait de leur complexité (notamment le copier/coller).

Le troisième scénario est basé sur une tâche de sélection de texte. Comme dans le deuxième scénario, les μ Gestes TTF servent ici à éviter l'utilisation des menus hiérarchiques qui allongent et complexifient les trajectoires d'interaction. Nous avons montré théoriquement que μ GeT, une technique d'interaction utilisant conjointement les μ Gestes TTF et l'écran tactile, surpasse la technique d'interaction traditionnelle proposée par VoiceOver (VO) en termes de vitesse de sélection et de nombre d'actions dans une trajectoire d'interaction optimale. Cependant, les participants s'écartent souvent de ces trajectoires optimales : les résultats empiriques montrent alors que les PDV sélectionnent le texte plus rapidement avec μ GeT, bien qu'ils soient plus sus-

ceptibles de produire des trajectoires sous-optimales. Cependant, ces actions sous-optimales ont un impact limité sur le temps de sélection et peuvent être rapidement corrigées. En revanche, VO est moins susceptible de produire des actions sous-optimales dans leur ensemble, mais celles-ci ont des conséquences plus coûteuses en nombres d'actions et donc en temps de réalisation.

Enfin, nous avons réalisé plusieurs prototypes de gants servant à capter les μ Gestes TTF nécessaires à l'interaction. Le premier prototype devait être en mesure de capturer les 33 μ Gestes TTF de la première expérimentation et embarquait pour ce faire de nombreux capteurs. Les expérimentations suivantes n'utilisaient qu'un sous-ensemble réduit de μ Gestes TTF parmi les 8 identifiés dans notre première étude. Aussi nous avons donc construit deux prototypes de gants modulaires, permettant de raccorder un ou plusieurs manchons en tissus à placer sur les doigts et comportant seulement les capteurs nécessaires à l'interaction - diminuant ainsi l'encombrement au strict nécessaire et améliorant l'ergonomie globale.

Limitations et Perspectives

Nos travaux ont une dimension très exploratoire : en l'absence de connaissance antérieure sur l'utilisation conjointe d'un écran tactile et des μ Gestes TTF dans une situation sans retour visuel, nous avons employé une démarche incrémentale et exploratoire (comme considérer plusieurs types de tâches) avec validation d'une étape avant de passer à la suivante de notre parcours. Nous avons ainsi commencé par nous poser la question de la faisabilité d'une telle combinaison. Puis, après avoir validé que les μ Gestes TTF - ou du moins certains d'entre eux - étaient réalisables dans les conditions évoquées, nous avons étudié la question de l'utilité et de l'utilisabilité de cette combinaison en considérant des tâches utilisateur variées que nous avons implémentées au sein de scénarios d'usage que nous avons soumis à des PDV.

Néanmoins, ces études ont été menées en laboratoire, dans des conditions contrôlées. L'utilité et l'utilisabilité perçue par les participants ont certainement été impactées par de nombreux biais cognitifs, bien connus des expérimentateurs :

CONCLUSION

- l'effet de nouveauté⁷ et l'effet "waouh"⁸, induits par l'aspect novateur de la technique d'interaction (utilisant un gant de captation) ;
- l'effet Hawthorne⁹, le biais de désirabilité sociale¹⁰ et le biais de courtoisie¹¹, qui motivent et poussent le participant à répondre plus favorablement aux questions.

Pour répondre à cette limitation, une première solution serait d'approfondir nos travaux avec une étude plus longitudinale et se déroulant en milieu écologique. Nous pourrions développer une fonctionnalité permettant au dispositif tactile (téléphone portable, tablette) de reconnaître les μ Gestes TTF au niveau de l'OS, de sorte à ce qu'ils soient utilisables avec toutes les applications. Tout un chacun pourrait alors utiliser les μ Gestes TTF à sa convenance, sur son téléphone portable ou sur tablette, chez soi comme en mobilité.

En récoltant les données d'utilisation de façon transparente, nous pourrions nous affranchir en partie des biais évoqués plus haut. En interrogeant régulièrement les utilisateurs, nous pourrions aussi obtenir des retours qualitatifs plus pertinents sur l'utilité et l'utilisabilité réelle des μ Gestes TTF appliqués à un dispositif tactile en situation d'usage quotidien.

Par ailleurs, cette étude longitudinale nous permettrait de faire monter les utilisateurs en compétence sur l'utilisation des μ Gestes TTF. Nous pourrions ainsi compléter notre quatrième étude portant sur la comparaison entre VoiceOver (VO) et μ GeT (voir Chapitre 3), une technique d'entrée multimodale utilisant l'écran tactile et les μ Gestes TTF. Des utilisateurs experts / non-néophytes se rapprocheraient davantage de la réalité ainsi que des trajectoires d'interaction optimales modélisées à l'aide de l'outil théorique STEM, nous permettant de confirmer ou préciser nos résultats.

Réaliser cette étude en milieu écologique ne pourrait toutefois pas se faire avec le prototype de gant actuel. Bien qu'il ait été amélioré et miniaturisé au cours de nos travaux, le dispositif de

7. Appeal to novelty - Wikipedia : https://en.wikipedia.org/wiki/Appeal_to_novelty/

8. Blog UX - Qu'est-ce que " l'effet waouh " ? : <https://blog-ux.com/quest-ce-que-leffet-waouh/>

9. Effet Hawthorne - Wikipedia : https://fr.wikipedia.org/wiki/Effet_Hawthorne

10. Social desirability bias - Wikipedia : https://en.wikipedia.org/wiki/Social-desirability_bias

11. Courtesy bias - Wikipedia : https://en.wikipedia.org/wiki/Response_bias#Courtesy_bias

CONCLUSION

captation est encore trop rudimentaire et peu discret : les capteurs sont reliés à la carte Arduino avec des fils, la carte elle-même étant portée sur le poignet et devant être reliée à un ordinateur ou un téléphone portable / tablette à l'aide d'un câble. Afin d'encourager son usage, nous pourrions développer une fonctionnalité permettant au dispositif tactile (téléphone portable, tablette) de reconnaître les μ Gestes TTF au niveau de l'OS, de sorte qu'ils soient utilisables avec toutes les applications. Tout un chacun pourrait alors utiliser les μ Gestes TTF à sa convenance, sur son téléphone portable ou sur tablette, chez soi comme en mobilité. Qui plus est, le dispositif devrait être sobre et facile à utiliser, à l'instar du dispositif présenté par Boldu et al. en 2018 [31] - voire d'un anneau connecté grand public tel que l'ArcX. Un tel dispositif limiterait cependant la localisation μ Gestes TTF et le nombre de commandes possibles - à moins de disposer de plusieurs bagues.

D'ailleurs, équiper plusieurs bagues nous permet d'envisager un usage bimanuel des μ Gestes TTF. Utiliser les deux mains pour effectuer plusieurs μ Gestes TTF - de façon successives ou simultanées - et conjointement à l'écran tactile offre un espace de conception extrêmement large, et permettrait de potentiellement décupler l'expressivité de façon exponentielle. Nous pouvons imaginer par exemple une suite à notre scénario #1 se déroulant sur une carte audio-tactile et dans laquelle il serait possible calculer une distance sur une carte, ou dans un programme de géométrie, voire opérer des actions directionnelles, comme par exemple "Copier ce que je pointe avec la main gauche vers l'endroit que je pointe de la main droite". Ou encore, de demander à l'aide des μ Gestes TTF un itinéraire entre deux points d'intérêt pointés par les index des deux mains : par exemple, un premier μ Gestes TTF avec la main droite servirait à indiquer le point de départ et un second μ Gestes TTF avec la main gauche servirait à indiquer le point d'arrivée.

En résumé, les perspectives à court et moyen termes concernant l'étude des μ Gestes TTF comme suppléant de l'interaction tactile sont vastes et variées, qu'il s'agisse d'une réflexion sur les usages ou sur l'innovation technologique nécessaire à l'adoption d'une telle façon d'interagir.

Toutefois, il est possible d'envisager l'avenir des μ Gestes TTF à plus long terme. Nos tra-

CONCLUSION

vaux se sont concentrés exclusivement sur un usage multimodal combinant l'écran tactile et les μ Gestes TTF. Comme présenté dans l'introduction de ce mémoire, nous avons fait le choix de travailler avec l'écran tactile car nous souhaitons nous conformer aux usages et demandes actuels : les PDV veulent utiliser un dispositif grand public, et notamment un téléphone portable, quant bien même celui-ci n'est pas suffisamment accessible. Se pose alors la question : avec l'essor de l'informatique ubiquitaire et l'évolution de la technologie qui tendent vers des systèmes de plus en plus portables et/ou intégrés à l'environnement (comme les montres et bagues connectées, les lunettes de réalité augmentée et virtuelle, la domotique etc.), le futur des dispositifs grand public pourrait bien se passer totalement d'écran - voire de surface d'interaction tactile, comme avec des projections volumétriques dans les airs (hologrammes) via des lunettes de réalité augmentée par exemple. Dans cette hypothèse, les téléphones portables sont-ils voués à disparaître des usages grand public ? Le cas échéant, les PDV voudront-ils utiliser ces nouveaux dispositifs d'interaction, quitte à sacrifier davantage d'accessibilité puisque ces dispositifs seront alors quasi-exclusivement basés sur le canal visuel ? Quid de nos travaux, deviendront-ils alors obsolètes ?

Plusieurs pistes de réflexion s'offrent à nous. Premièrement, il est peu probable que les interactions utilisant une surface tactile finissent par disparaître totalement. Qui plus est, de plus en plus d'objets physiques du quotidien, tels que les frigidaires, les thermostats et les micro-ondes, sont dotés d'écrans tactiles, avec lesquels les μ Gestes TTF pourraient être utilisés.

Deuxièmement, les μ Gestes TTF peuvent être utilisés sans écran tactile. D'ailleurs, durant nos expérimentations, plusieurs PDV nous ont fait la remarque que les μ Gestes TTF seraient pratiques même utilisés sans le téléphone portable ou la tablette, étayant leur argument par de nombreux exemples (par exemple, celui de la télécommande pour activer les feux aux passages piétons ; comme manette de jeu ; ou encore de remplacer complètement les interactions sur l'écran tactile du téléphone portable par des μ Gestes TTF).

Une première perspective serait d'étudier des microgestes sans toucher l'écran tactile par exemple avec la main non dominante tandis que la main dominante interagit avec l'écran tactile. Une deuxième perspective à long terme de notre travail serait d'étudier les μ Gestes TTF comme

BIBLIOGRAPHIE

modalité utilisée seule à destination des PDV. En effet, si les μ Gestes TTF (et le paradigme gestuel en général) ont beaucoup été étudiés (en l'air ; sur une surface ; avec la main libre ou avec un objet tenu ; en tant que modalité d'interaction principale ou secondaire), à notre connaissance, ils ne l'ont pas beaucoup été pour les PDV. Il serait donc intéressant d'explorer cette piste complémentaire à nos travaux, pour toutes les raisons évoquées dans ce mémoire.

Bibliographie

- [1] 1998. Assistive Technology Act. <https://www.congress.gov/105/plaws/publ394/PLAW-105publ394.pdf> Consulté le 10 Octobre 2023..
- [2] 2005. Déficience Visuelle - Etudes et Résultats.
- [3] 2020. GlobalWebIndex's flagship report on device ownership and usage. <https://www.gwi.com/reports/trends-2020> Consulté le 10 October 2023.
- [4] Ali Abdolrahmani, Ravi Kuber, and Stacy M. Branham. 2018. "Siri Talks at You" : An Empirical Investigation of Voice-Activated Personal Assistant (VAPA) Usage by Individuals Who Are Blind. (2018). <https://doi.org/10.1145/3234695.3236344>
- [5] Carl Halladay Abraham, Bert Boadi-Kusi, Enyam Komla Amewuho Morny, and Prince Agyekum. 2021. Smartphone usage among people living with severe visual impairment and blindness. *Assistive Technology* (2021). <https://doi.org/10.1080/10400435.2021.1907485>
- [6] Jurine Adolf. 1989. An Expert System for the Selection of Input Devices. *Proceedings of the Human Factors Society Annual Meeting* 33, 5 (1989), 335–339. <https://doi.org/10.1177/154193128903300522>
- [7] Sonu Agarwal and Sanjay Ghosh. 2017. Evaluation of Microgesture Recognition Using a Smartwatch. (2017). <https://doi.org/10.1109/ICMLA.2017.00-24>
- [8] Faisal Ahmed, Muhammad Islam, Yevgen Borodin, and Iv Ramakrishnan. 2010. Assistive

BIBLIOGRAPHIE

- web browsing with touch interfaces. 235–236. <https://doi.org/10.1145/1878803.1878848>
- [9] Roland Aigner, Daniel Wigdor, Hrvoje Benko, Michael Haller, David Lindlbauer, Alexandra Ion, Shengdong Zhao, and Jeffrey Tzu Kwan Valino Koh. 2012. Understanding Mid-Air Hand Gestures : A Study of Human Preferences in Usage of Gesture Types for HCI. (2012). <https://www.microsoft.com/en-us/research/publication/understanding-mid-air-hand-gestures-a-study-of-human-preferences-in-usage-of-gesture-types>
- [10] Jason Alexander, Teng Han, William Judd, Pourang Irani, and Sriram Subramanian. 2012. Putting your best foot forward : Investigating real-world mappings for foot-based gestures. (2012). <https://doi.org/10.1145/2207676.2208575>
- [11] Fábio M. Alfieri André T. Sugawara, Vinícius D. Ramos and Linamara R. Battistella. 2018. Abandonment of assistive products : assessing abandonment levels and factors that impact on it. *Disability and Rehabilitation : Assistive Technology* 13, 7 (2018), 716–723. <https://doi.org/10.1080/17483107.2018.1425748> PMID : 29334475.
- [12] Daniel Ashbrook. 2010. *Enabling Mobile Microinteractions*. Ph. D. Dissertation. USA.
- [13] Daniel Ashbrook, Carlos Tejada, Dhwanit Mehta, Anthony Jiminez, Goudam Muralitharam, Sangeeta Gajendra, and Ross Tallents. 2016. Bitey : An Exploration of Tooth Click Gestures for Hands-Free User Interface Control. (2016). <https://doi.org/10.1145/2935334.2935389>
- [14] Vikas Ashok, Yevgen Borodin, Svetlana Stoyanchev, Yuri Puzis, and I. V. Ramakrishnan. 2014. Wizard-of-Oz Evaluation of Speech-Driven Web Browsing Interface for People with Vision Impairments. In *Proceedings of the 11th Web for All Conference (Seoul, Korea) (W4A '14)*. Association for Computing Machinery, New York, NY, USA, Article 12, 9 pages. <https://doi.org/10.1145/2596695.2596699>
- [15] Mauro Avila Soto and Markus Funk. 2018. Look, a Guidance Drone ! Assessing the Social Acceptability of Companion Drones for Blind Travelers in Public Spaces. In *Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility*

- (Galway, Ireland) (*ASSETS '18*). Association for Computing Machinery, New York, NY, USA, 417–419.
- [16] Shiri Azenkot. 2014. *Eyes-Free input on mobile devices*. Ph. D. Dissertation.
- [17] Shiri Azenkot and Nicole B. Lee. 2013. Exploring the Use of Speech Input by Blind People on Mobile Devices. In *Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility (Bellevue, Washington) (ASSETS '13)*. Association for Computing Machinery, New York, NY, USA, Article 11, 8 pages. <https://doi.org/10.1145/2513383.2513440>
- [18] Shiri Azenkot and Shumin Zhai. 2012. Touch Behavior with Different Postures on Soft Smartphone Keyboards. In *Proceedings of the 14th International Conference on Human-Computer Interaction with Mobile Devices and Services (San Francisco, California, USA) (MobileHCI '12)*. Association for Computing Machinery, New York, NY, USA, 251–260. <https://doi.org/10.1145/2371574.2371612>
- [19] Mark S. Baldwin, Gillian R. Hayes, Oliver L. Haimson, Jennifer Mankoff, and Scott E. Hudson. 2017. The Tangible Desktop : A Multimodal Approach to Nonvisual Computing. *ACM Trans. Access. Comput.* 10, 3, Article 9 (2017), 28 pages.
- [20] Mark S. Baldwin, Rushil Khurana, Duncan McIsaac, Yuqian Sun, Tracy Tran, Xiaoyi Zhang, James Fogarty, Gillian R. Hayes, and Jennifer Mankoff. 2019. Tangible Interfaces. In *Web Accessibility*, Yeliz Yesilada and Simon Harper (Eds.). Springer London, London, 715–735. https://doi.org/10.1007/978-1-4471-7440-0_36 Series Title : Human–Computer Interaction Series.
- [21] Sandra Bardot, Marcos Serrano, and Christophe Jouffrais. 2016. From Tactile to Virtual : Using a Smartwatch to Improve Spatial Map Exploration for Visually Impaired Users. In *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services (Florence, Italy) (MobileHCI '16)*. Association for Computing Machinery, New York, NY, USA, 100–111. <https://doi.org/10.1145/2935334.2935342>

- [22] Sandra Bardot, Marcos Serrano, Bernard Oriola, and Christophe Jouffrais. 2017. Identifying how visually impaired people explore raised-line diagrams to improve the design of touch interfaces. In *Conference on Human Factors in Computing Systems - Proceedings*, Vol. 2017-May. ACM Press, New York, New York, USA, 550–555. <https://doi.org/10.1145/3025453.3025582>
- [23] Thomas Baudel and Michel Beaudouin-Lafon. 1993. Charade : Remote Control of Objects Using Free-Hand Gestures. *Commun. ACM* 36, 7 (1993), 28–35. <https://doi.org/10.1145/159544.159562>
- [24] Michel Beaudouin-Lafon. 2000. Instrumental Interaction : An Interaction Model for Designing Post-WIMP User Interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (The Hague, The Netherlands) (CHI '00)*. Association for Computing Machinery, New York, NY, USA, 446–453. <https://doi.org/10.1145/332040.332473>
- [25] Michel Beaudouin-Lafon. 2004. Designing Interaction, Not Interfaces. In *Proceedings of the Working Conference on Advanced Visual Interfaces (Gallipoli, Italy) (AVI '04)*. Association for Computing Machinery, New York, NY, USA, 15–22. <https://doi.org/10.1145/989863.989865>
- [26] Yacine Bellik. 1995. *Interfaces multimodales : concepts, modeles et architectures*. Ph. D. Dissertation. <http://www.theses.fr/1995PA112178> Thèse de doctorat dirigée par LIENARD, J.-S. Sciences appliquées Paris 11 1995.
- [27] Hemant Bhaskar Surale, Fabrice Matulic, and Daniel Vogel. 2017. Experimental Analysis of Mode Switching Techniques in Touch-based User Interfaces multi-touch ; touch input ; mode switching. (2017). <https://doi.org/10.1145/3025453.3025865>
- [28] Jeffrey P. Bigham, Craig M. Prince, and Richard E. Ladner. 2008. WebAnywhere : A Screen Reader on-the-Go. In *Proceedings of the 2008 International Cross-Disciplinary Conference on Web Accessibility (W4A) (Beijing, China) (W4A '08)*. Association for Computing Machinery, New York, NY, USA, 73–82. <https://doi.org/10.1145/1368044.1368060>

- [29] Alison M. Binns, Catey Bunce, Chris Dickinson, Robert Harper, Rhiannon Tudor-Edwards, Margaret Woodhouse, Pat Linck, Alan Suttie, Jonathan Jackson, Jennifer Lindsay, James Wolffsohn, Lindsey Hughes, and Tom H. Margrain. 2012. How Effective is Low Vision Service Provision ? A Systematic Review. *Survey of Ophthalmology* 57, 1 (2012), 34–65. Publisher : Elsevier.
- [30] Toine Bogers, Ammar Al-Basri, Claes Rytlig, Mads Møller, Mette Rasmussen, Nikita Michelsen, and Sara Jørgensen. 2019. *A Study of Usage and Usability of Intelligent Personal Assistants in Denmark : 14th International Conference, iConference 2019, Washington, DC, USA, March 31–April 3, 2019, Proceedings*. 79–90. https://doi.org/10.1007/978-3-030-15742-5_7
- [31] Roger Boldu, Alexandru Dancu, Denys J.C. C Matthies, Pablo Gallego Cascón, Shanaka Ransir, and Suranga Nanayakkara. 2018. Thumb-In-Motion : Evaluating Thumb-to-Ring Microgestures for Athletic Activity. In *Proceedings of the Symposium on Spatial User Interaction (SUI '18)*. Association for Computing Machinery, New York, NY, USA, 150–157. <https://doi.org/10.1145/3267782.3267796>
- [32] Roger Boldu, Denys J.C. Matthies, Haimo Zhang, and Suranga Nanayakkara. 2020. AiSee : An Assistive Wearable Device to Support Visually Impaired Grocery Shoppers. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 4, 4, Article 119 (2020), 25 pages. <https://doi.org/10.1145/3432196>
- [33] Matthew Bonner, Jeremy Brudvik, Gregory Abowd, and W. Edwards. 2010. No-Look Notes : Accessible Eyes-Free Multi-touch Text Entry, Vol. 6030. 409–426. https://doi.org/10.1007/978-3-642-12654-3_24
- [34] Jens Bornschein, Denise Bornschein, and Gerhard Weber. 2018. Comparing Computer-Based Drawing Methods for Blind People with Real-Time Tactile Feedback. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (Montreal QC, Canada) (CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–13. <https://doi.org/10.1145/3173574.3173689>

BIBLIOGRAPHIE

- [35] Robin Brewer, Mark Cartwright, Aaron Karp, Bryan Pardo, and Anne Marie Piper. 2016. An Approach to Audio-Only Editing for Visually Impaired Seniors. 307–308. <https://doi.org/10.1145/2982142.2982196>
- [36] Anke Brock. 2017. Tangible Interaction for Visually Impaired People : why and how. In *World Haptics Conference - Workshop on Haptic Interfaces for Accessibility*. Fuerstenfeldbruck, Germany, 3. <https://inria.hal.science/hal-01523745>
- [37] Anke M. Brock, Philippe Truillet, Bernard Oriola, Delphine Picard, and Christophe Jouffrais. 2015. Interactivity Improves Usability of Geographic Maps for Visually Impaired People. *Human-Computer Interaction* 30, 2 (2015), 156–194. <https://doi.org/10.1080/07370024.2014.924412>
- [38] Emeline Brule, Gilles Bailly, Anke Brock, Frédéric Valentin, Grégoire Denis, and Christophe Jouffrais. 2016. MapSense : Multi-sensory interactive maps for children living with visual impairments. In *Conference on Human Factors in Computing Systems - Proceedings*. Association for Computing Machinery, 445–457. <https://doi.org/10.1145/2858036.2858375>
- [39] Emeline Brulé, Brianna J. Tomlinson, Oussama Metatla, Christophe Jouffrais, and Marcos Serrano. 2020. *Review of Quantitative Empirical Evaluations of Technology for People with Visual Impairments*. Association for Computing Machinery, New York, NY, USA, 1–14.
- [40] William Buxton. 1983. Lexical and Pragmatic Considerations of Input Structures. *SIG-GRAPH Comput. Graph.* 17, 1 (1983), 31–37. <https://doi.org/10.1145/988584.988586>
- [41] William A.S. Buxton. 1990. A three-state model of graphical input. In *IFIP TC13 International Conference on Human-Computer Interaction*.
- [42] Maria Claudia Buzzi, Marina Buzzi, Barbara Leporini, and Amaury Trujillo. 2017. Analyzing visually impaired people's touch gestures on smartphones. *Multimedia Tools and Applications* 76, 4 (2017), 5141–5169. <https://doi.org/10.1007/s11042-016-3594-9>

- [43] Claude Cadoz. 1994. Le geste canal de communication homme/machine : la communication "instrumentale". *Technique Et Science Informatiques* 13 (1994), 31–61. <https://api.semanticscholar.org/CorpusID:59781972>
- [44] Stuart K. Card, Jock D. Mackinlay, and George G. Robertson. 1990. The Design Space of Input Devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Seattle, Washington, USA) (*CHI '90*). Association for Computing Machinery, New York, NY, USA, 117–124. <https://doi.org/10.1145/97243.97263>
- [45] Stuart K. Card, Thomas P. Moran, and Allen Newell. 1980. The Keystroke-Level Model for User Performance Time with Interactive Systems. *Commun. ACM* 23, 7 (1980), 396–410. <https://doi.org/10.1145/358886.358895>
- [46] Stuart K. Card, Thomas P. Moran, and Allen Newell. 1983. *The Psychology of Human-Computer Interaction* (1 ed.). CRC Press. <https://doi.org/10.1201/9780203736166>
- [47] Zaira Cattaneo and Tomaso Vecchi. 2011. *Blind Vision : The Neuroscience of Visual Impairment*. The MIT Press. <https://doi.org/10.7551/mitpress/9780262015035.001.0001>
- [48] Adrien Chaffangeon Caillet, Alix Goguey, and Laurence Nigay. 2023. μ Glyph : A Microgesture Notation. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) (*CHI '23*). Association for Computing Machinery, New York, NY, USA, Article 269, 28 pages. <https://doi.org/10.1145/3544548.3580693>
- [49] Edwin Chan, Teddy Seyed, Wolfgang Stuerzlinger, Xing-Dong Dong Yang, and Frank Maurer. 2016. User Elicitation on Single-Hand Microgestures. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (*CHI '16*). Association for Computing Machinery, New York, NY, USA, 3403–3414. <https://doi.org/10.1145/2858036.2858589>
- [50] Mei-Ling Chen and Hao-Chuan Wang. 2018. How Personal Experience and Technical Knowledge Affect Using Conversational Agents. In *Proceedings of the 23rd International*

- Conference on Intelligent User Interfaces Companion (Tokyo, Japan) (IUI '18 Companion)*. Association for Computing Machinery, New York, NY, USA, Article 53, 2 pages. <https://doi.org/10.1145/3180308.3180362>
- [51] Victor Chen, Xuhai Xu, Richard Li, Yuanchun Shi, Shwetak Patel, and Yuntao Wang. 2021. Understanding the Design Space of Mouth Microgestures. (2021). arXiv :2106.00931 <https://arxiv.org/abs/2106.00931v1>
- [52] Peggy Pei-Chia Chiang, Patricia Mary O'Connor, Richard Thomas Le Mesurier, and Jill Elizabeth Keefe. 2011. A Global Survey of Low Vision Service Provision. *Ophthalmic Epidemiology* 18, 3 (2011), 109–121. <https://doi.org/10.3109/09286586.2011.560745>
- [53] Daniel Costa and Carlos Duarte. 2020. Alternative Modalities for Visually Impaired Users to Control Smart TVs. *Multimedia Tools Appl.* 79, 43–44 (nov 2020), 31931–31955. <https://doi.org/10.1007/s11042-020-09656-1>
- [54] Benjamin R. Cowan, Nadia Pantidi, David Coyle, Kellie Morrissey, Peter Clarke, Sara Al-Shehri, David Earley, and Natasha Bandeira. 2017. "What Can i Help You with?" : Infrequent Users' Experiences of Intelligent Personal Assistants. In *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services (Vienna, Austria) (MobileHCI '17)*. Association for Computing Machinery, New York, NY, USA, Article 43, 12 pages. <https://doi.org/10.1145/3098279.3098539>
- [55] Andrew Crossan, Stephen Brewster, and Alexander Ng. 2010. Foot Tapping for Mobile Interaction. (2010). <https://doi.org/10.5555/2146303.2146366>
- [56] Michael D. Crossland, Rui S. Silva, and Antonio F. Macedo. 2014. Smartphone, tablet computer and e-reader use by people with vision impairment. *Ophthalmic and Physiological Optics* 34, 5 (2014), 552–557.
- [57] Rafael Jeferson Pezzuto Damaceno, Juliana Cristina Braga, and Jesús Pascual Mena-Chalco. 2018. Mobile device accessibility for the visually impaired : problems mapping

- and recommendations. *Universal Access in the Information Society* 17, 2 (2018), 421–435. <https://doi.org/10.1007/s10209-017-0540-1>
- [58] Alireza Darvishy, Hans-Peter Hutter, Edin Beljulji, and Zeno Heeb. 2023. A new conversational interaction concept for document creation and editing on mobile devices for visually impaired users. arXiv :2301.02546
- [59] James W. Davis and Serge Vaks. 2001. A perceptual user interface for recognizing head gesture acknowledgements. *ACM International Conference Proceeding Series* 15-16-Nove (2001), 21. <https://doi.org/10.1145/971478.971504>
- [60] Jessyca L. Derby, Emily A. Rickel, Kelly June Harris, Jade Lovell, and Barbara S. Chaparro. 2020. “We Didn’t Catch That!” Using Voice Text Input on a Mixed Reality Headset in Noisy Environments. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 64 (2020), 2102 – 2106. <https://api.semanticscholar.org/CorpusID:231875948>
- [61] Christina Dicke and Jörg Müller. 2015. Evaluating Mid-Air List Interaction for Spatial Audio Interfaces. In *Proceedings of the 3rd ACM Symposium on Spatial User Interaction (Los Angeles, California, USA) (SUI '15)*. Association for Computing Machinery, New York, NY, USA, 24–33. <https://doi.org/10.1145/2788940.2788945>
- [62] Pierre Dragicevic. 2004. *Un modèle d'interaction en entrée pour des systèmes interactifs multi-dispositifs hautement configurables*. Ph. D. Dissertation. <http://www.theses.fr/2004NANT2053> Thèse de doctorat dirigée par Hégron, Gérard et Fekete, Jean-Daniel Informatique Nantes 2004.
- [63] Pierre Dragicevic. 2016. Fair Statistical Communication in HCI. Springer, Cham, 291–330. https://doi.org/10.1007/978-3-319-26633-6_13
- [64] Carlos Duarte, Simon Desart, David Costa, and Bruno Dumas. 2017. Designing Multi-modal Mobile Interaction for a Text Messaging Application for Visually Impaired Users. *Frontiers in ICT* 4 (12 2017). <https://doi.org/10.3389/fict.2017.00026>

- [65] Julie Ducasse, Anke M. Brock, and Christophe Jouffrais. 2018. Accessible Interactive Maps for Visually Impaired Users. In *Mobility of Visually Impaired People*, Edwige Pissaloux and Ramiro Velazquez (Eds.). Springer International Publishing, Cham, 537–584. https://doi.org/10.1007/978-3-319-54446-5_17
- [66] Aarthi Easwara Moorthy and Kim-Phuong Vu. 2014. Privacy Concerns for Use of Voice Activated Personal Assistant in the Public Space. *International Journal of Human-Computer Interaction* 31 (04 2014), 307–335. <https://doi.org/10.1080/10447318.2014.986642>
- [67] Darren Edge and Alan Blackwell. 2006. Correlates of the cognitive dimensions for tangible user interface. *Journal of Visual Languages & Computing* 17 (08 2006), 366–394. <https://doi.org/10.1016/j.jvlc.2006.04.005>
- [68] Christos Efthymiou and Martin Halvey. 2016. Evaluating the Social Acceptability of Voice Based Smartwatch Search. 267–278. https://doi.org/10.1007/978-3-319-48051-0_20
- [69] Yasmine N. El-Glaly, Francis Quek, Tonya Smith-Jackson, and Gurjot Dhillon. 2013. Touch-Screens Are Not Tangible : Fusing Tangible Interaction with Touch Glass in Readers for the Blind. In *Proceedings of the 7th International Conference on Tangible, Embedded and Embodied Interaction (Barcelona, Spain) (TEI '13)*. Association for Computing Machinery, New York, NY, USA, 245–253. <https://doi.org/10.1145/2460625.2460665>
- [70] Christoph Endres, Tim Schwartz, and Christian A. Müller. 2011. Geremin" : 2D Microgestures for Drivers Based on Electric Field Sensing. In *Proceedings of the 16th International Conference on Intelligent User Interfaces (Palo Alto, CA, USA) (IUI '11)*. Association for Computing Machinery, New York, NY, USA, 327–330. <https://doi.org/10.1145/1943403.1943457>
- [71] Gauthier Robert Jean Faisandaz, Alix Goguey, Christophe Jouffrais, and Laurence Nigay. 2022. Keep in Touch : Combining Touch Interaction with Thumb-to-Finger μ Gestures for People with Visual Impairment. In *Proceedings of the 2022 International Conference on*

- Multimodal Interaction* (Bengaluru, India) (*ICMI '22*). Association for Computing Machinery, New York, NY, USA, 105–116. <https://doi.org/10.1145/3536221.3556589>
- [72] Gauthier Robert Jean Faisandaz, Alix Goguey, Christophe Jouffrais, and Laurence Nigay. 2023. μ GeT : Multimodal Eyes-Free Text Selection Technique Combining Touch Interaction and Microgestures. In *INTERNATIONAL CONFERENCE ON MULTIMODAL INTERACTION* (Paris, France) (*ICMI '23*). Association for Computing Machinery, New York, NY, USA, 594–603. <https://doi.org/10.1145/3577190.3614131>
- [73] Catherine Feng. 2016. Designing wearable mobile device controllers for blind people : A co-design approach. In *ASSETS 2016 - Proceedings of the 18th International ACM SIGACCESS Conference on Computers and Accessibility*. Association for Computing Machinery, Inc, 341–342. <https://doi.org/10.1145/2982142.2982144>
- [74] George W. Fitzmaurice and William Buxton. 1997. An Empirical Evaluation of Graspable User Interfaces : Towards Specialized, Space-Multiplexed Input. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems* (Atlanta, Georgia, USA) (*CHI '97*). Association for Computing Machinery, New York, NY, USA, 43–50. <https://doi.org/10.1145/258549.258578>
- [75] James D. Foley, Victor L. Wallace, and Peggy Chan. 1984. The human factors of computer graphics interaction techniques. *IEEE Computer Graphics and Applications* 4, 11 (1984), 13–48. <https://doi.org/10.1109/MCG.1984.6429355>
- [76] WHO Programme for the Prevention of Blindness and Deafness. 2003. Consultation on development of standards for characterization of vision loss and visual functioning : Geneva, 4-5 September 2003. , WHO/PBL/03.91 pages. <https://iris.who.int/handle/10665/68601> Consulté le 14 octobre 2023.
- [77] Marcus Fuhrer. 2001. Assistive technology outcomes research : challenges met and yet unmet. *American journal of physical medicine and rehabilitation / Association of Academic Physiatrists* 80 (2001), 528–35.

- [78] Vinitha Gadiraju, Jérémie Garcia, Shaun Kane, and Anke M. Brock. 2021. “It is Fascinating to Make These Beasts Fly” : Understanding Visually Impaired People’s Motivations and Needs for Drone Piloting. In *Proceedings of the 23rd International ACM SIGACCESS Conference on Computers and Accessibility (Virtual Event, USA) (ASSETS '21)*. Association for Computing Machinery, New York, NY, USA, Article 42, 12 pages. <https://doi.org/10.1145/3441852.3471219>
- [79] Pablo Gallego Cascón, Denys J.C. Matthies, Sachith Muthukumarana, and Suranga Nanayakkara. 2019. Chewlt. An Intraoral Interface for Discreet Interactions. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (Glasgow, Scotland Uk) (CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–13. <https://doi.org/10.1145/3290605.3300556>
- [80] Debjyoti Ghosh, Pin Sym Foong, Shengdong Zhao, Di Chen, and Morten Fjeld. 2018. EDITalk : Towards Designing Eyes-Free Interactions for Mobile Word Processing. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (Montreal QC, Canada) (CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–10.
- [81] Alix Goguey, Géry Casiez, Andy Cockburn, and Carl Gutwin. 2018. Storyboard-Based Empirical Modeling of Touch Interface Performance. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (Montreal QC, Canada) (CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–12. <https://doi.org/10.1145/3173574.3174019>
- [82] Alix Goguey, Sylvain Malacria, and Carl Gutwin. 2018. Improving Discoverability and Expert Performance in Force-Sensitive Text Selection for Touch Devices with Mode Gauges. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (2018)*. <https://doi.org/10.1145/3173574>
- [83] Nora Griffin-Shirley, Devender R Banda, Paul M Ajuwon, Jongpil Cheon, Jaehoon Lee, Hye Ran Park, and Sanpalei Nylla Lyngdoh. 2017. A Survey on the Use of Mobile Appli-

- cations for People who Are Visually Impaired. *Journal of Visual Impairment and Blindness* 111 (2017), 307–323.
- [84] William Grussenmeyer and eelke folmer. 2016. AudioDraw : user preferences in non-visual diagram drawing for touchscreens. 1–8. <https://doi.org/10.1145/2899475.2899483>
- [85] Tiago Guerreiro, Joaquim Jorge, and Daniel Gonçalves. 2012. Exploring the Non-Visual Acquisition of Targets on Touch Phones and Tablets.
- [86] Tiago Guerreiro, Kyle Montague, João Guerreiro, Rafael Nunes, Hugo Nicolau, and Daniel Gonçalves. 2015. Blind People Interacting with Large Touch Surfaces : Strategies for One-handed and Two-handed Exploration. <https://doi.org/10.1145/2817721.2817743>
- [87] Sean G. Gustafson, Bernhard Rabe, and Patrick M. Baudisch. 2013. Understanding Palm-Based Imaginary Interfaces : The Role of Visual and Tactile Cues When Browsing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Paris, France) (*CHI '13*). Association for Computing Machinery, New York, NY, USA, 889–898. <https://doi.org/10.1145/2470654.2466114>
- [88] Yvette Hatwell, Arlette Streri, and Edouard Gentaz (Eds.). 2003. *Touching for Knowing : Cognitive psychology of haptic manual perception*. Advances in Consciousness Research, Vol. 53. John Benjamins Publishing Company, Amsterdam. <https://doi.org/10.1075/aicr.53>
- [89] Renate Häuslschmid, Benjamin Menrad, and Andreas Martin Butz. 2015. Freehand vs. micro gestures in the car : Driving performance and user experience. *2015 IEEE Symposium on 3D User Interfaces (3DUI)* (2015), 159–160. <https://api.semanticscholar.org/CorpusID:9537323>
- [90] Juan David Hincapié-Ramos, Xiang Guo, Paymahn Moghadasian, and Pourang Irani. 2014. Consumed Endurance : A Metric to Quantify Arm Fatigue of Mid-Air Interactions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (To-

BIBLIOGRAPHIE

- ronto, Ontario, Canada) (*CHI '14*). Association for Computing Machinery, New York, NY, USA, 1063–1072. <https://doi.org/10.1145/2556288.2557130>
- [91] S Hocherman. 1993. Proprioceptive guidance and motor planning of reaching movements to unseen targets. *Experimental brain research* 95, 2 (1993), 349–358. <https://doi.org/10.1007/BF00229793>
- [92] Taslima Hossain, Yeasmin Akter, and Md Rahman. 2020. Voice mail application for visually impaired persons. *Recent Research in Science and Technology* (01 2020), 15–18. <https://doi.org/10.25081/rrst.2020.12.6086>
- [93] Yi-Ta Hsieh, Antti Jylhä, Valeria Orso, Luciano Gamberini, and Giulio Jacucci. 2016. Designing a Willing-to-Use-in-Public Hand Gestural Interaction Technique for Smart Glasses. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) (*CHI '16*). Association for Computing Machinery, New York, NY, USA, 4203–4215. <https://doi.org/10.1145/2858036.2858436>
- [94] Da-Yuan Huang, Liwei Chan, Shuo Yang, Fan Wang, Rong-Hao Liang, De-Nian Yang, Yi-Ping Hung, and Bing-Yu Chen. 2016. DigitSpace : Designing Thumb-to-Fingers Touch Interfaces for One-Handed and Eyes-Free Interactions. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) (*CHI '16*). Association for Computing Machinery, New York, NY, USA, 1526–1537. <https://doi.org/10.1145/2858036.2858483>
- [95] Amy Hurst and Jasmine Tobias. 2011. Empowering Individuals with Do-It-Yourself Assistive Technology. In *The Proceedings of the 13th International ACM SIGACCESS Conference on Computers and Accessibility* (Dundee, Scotland, UK) (*ASSETS '11*). Association for Computing Machinery, New York, NY, USA, 11–18. <https://doi.org/10.1145/2049536.2049541>
- [96] Takeo Igarashi and John F. Hughes. 2001. Voice as Sound : Using Non-Verbal Voice Input for Interactive Control. In *Proceedings of the 14th Annual ACM Symposium on User Interface Software and Technology* (Orlando, Florida) (*UIST '01*). Association for Computing Machinery, New York, NY, USA, 155–156. <https://doi.org/10.1145/502348.502372>

- [97] Danielle Irvine, Alex Zemke, Gregg Pusateri, Leah Gerlach, Rob Chun, and Walter M. Jay. 2014. Tablet and Smartphone Accessibility Features in the Low Vision Rehabilitation. *Neuro-Ophthalmology* 38, 2 (2014), 53–59.
- [98] Robert J.K. Jacob. 1990. What you look at is what you get : Eye movement-based interaction techniques. *Conference on Human Factors in Computing Systems - Proceedings* (1990), 11–18. <https://doi.org/10.1145/97243.97246>
- [99] Rabia Jafri, Shatha MUSAED M. Althbiti, Nora Abdulrazaq A. Alattas, Amal Abdullah A. Albraiki, and Sarah Hassan A. Almuhawwis. 2022. Tac-Trace : A Tangible User Interface-Based Solution for Teaching Shape Concepts to Visually Impaired Children. *IEEE Access* 10 (2022), 131153–131165.
- [100] Jean-François Jégo. 2013. *Interaction basée sur des gestes définis par l'utilisateur : Application à la réalité virtuelle*. Ph. D. Dissertation. <http://www.theses.fr/2013ENMP0033/document> Thèse de doctorat dirigée par Fuchs, Philippe Informatique temps réel, robotique et automatique Paris, ENMP 2013.
- [101] Haiyan Jiang, Dongdong Weng, Zhenliang Zhang, and Feng Chen. 2019. HiFinger : One-Handed Text Entry Technique for Virtual Environments Based on Touches between Fingers. *Sensors (Switzerland)* 19, 14 (2019), 1–24. <https://doi.org/10.3390/s19143063>
- [102] Hwayeon Joh, Yun Jung Lee, and Uran Oh. 2022. Understanding the Touchscreen-Based Nonvisual Target Acquisition Task Performance of Screen Reader Users. In *Proceedings of the 19th International Web for All Conference (Lyon, France) (W4A '22)*. Association for Computing Machinery, New York, NY, USA, Article 11, 10 pages. <https://doi.org/10.1145/3493612.3520454>
- [103] Jingun Jung, Sunmin Son, Sangyoon Lee, Yeonsu Kim, and Geehyuk Lee. 2021. ThroughHand : 2D Tactile Interaction to Simultaneously Recognize and Touch Multiple Objects. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (Yokohama, Japan) (CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 29, 13 pages. <https://doi.org/10.1145/3411764.3445530>

BIBLIOGRAPHIE

- [104] Jeffrey W. Jutai, J. Graham Strong, and Elizabeth Russell-Minda. 2009. Effectiveness of Assistive Technologies for Low Vision Rehabilitation : A Systematic Review. *Journal of Visual Impairment & Blindness* 103, 4 (2009), 210–222.
- [105] Raine A. Kajastila and Tapio Lokki. 2009. A gesture-based and eyes-free control method for mobile devices. *CHI '09 Extended Abstracts on Human Factors in Computing Systems* (2009).
- [106] Shaun Kane, Jeffrey Bigham, and Jacob Wobbrock. 2008. Slide Rule : Making Mobile Touch Screens Accessible to Blind people using Multi-Touch Interaction Techniques. *ASSETS'08 : The 10th International ACM SIGACCESS Conference on Computers and Accessibility*, 73–80. <https://doi.org/10.1145/1414471.1414487>
- [107] Shaun K. Kane, Chandrika Jayant, Jacob O. Wobbrock, and Richard E. Ladner. 2009. Freedom to Roam : A Study of Mobile Device Adoption and Accessibility for People with Visual and Motor Disabilities. In *Proceedings of the 11th International ACM SIGACCESS Conference on Computers and Accessibility* (Pittsburgh, Pennsylvania, USA) (*Assets '09*). Association for Computing Machinery, New York, NY, USA, 115–122. <https://doi.org/10.1145/1639642.1639663>
- [108] Shaun K Kane, Meredith Ringel Morris, and Jacob O Wobbrock. 2013. Touchplates : Low-Cost Tactile Overlays for Visually Impaired Touch Screen Users. *Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility* (2013). <https://doi.org/10.1145/2513383>
- [109] Shaun K Kane and Jacob O Wobbrock. 2011. Usable Gestures for Blind People : Understanding Preference and Performance. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2011). <https://doi.org/10.1145/1978942>
- [110] Hsin Liu Kao, Artem Dementyev, Joseph A. Paradiso, and Chris Schmandt. 2015. NailO : Fingernails as an input surface. *Conference on Human Factors in Computing Systems - Proceedings 2015-April*, April (2015), 3015–3018. <https://doi.org/10.1145/2702123.2702572>

- [111] Maria Karam and m.c Schraefel. 2005. A Taxonomy of Gestures in Human Computer Interactions. *Electronics and Computer Science* (01 2005).
- [112] Keiko Katsuragawa, Ju Wang, Ziyang Shan, Ningshan Ouyang, Omid Abari, and Daniel Vogel. 2019. Tip-Tap : Battery-Free Discrete 2D Fingertip Input. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology* (New Orleans, LA, USA) (*UIST '19*). Association for Computing Machinery, New York, NY, USA, 1045–1057. <https://doi.org/10.1145/3332165.3347907>
- [113] Akif Khan and Shah Khusro. 2021. An insight into smartphone-based assistive solutions for visually impaired and blind people : issues, challenges and opportunities. *Universal Access in the Information Society* 20, 2 (2021), 265–298. <https://doi.org/10.1007/s10209-020-00733-8>
- [114] Nem Khan Dim, Chaklam Silpasuwanchai, Sayan Sarcar, and Xiangshi Ren. 2016. Designing Mid-Air TV Gestures for Blind People Using User- and Choice-Based Elicitation Approaches. 204–214. <https://doi.org/10.1145/2901790.2901834>
- [115] Shah Khusro, Badam Niazi, Akif Khan, and Iftikhar Alam. 2019. Evaluating Smartphone Screen Divisions for Designing Blind-Friendly Touch-Based Interfaces. *2019 International Conference on Frontiers of Information Technology (FIT)* (2019), 328–3285.
- [116] Elmar Krajnc, Mathias Knoll, Johannes Feiner, and Mario Traar. 2011. A Touch Sensitive User Interface Approach on Smartphones for Visually Impaired and Blind Persons. In *Symposium of the Workgroup Human-Computer Interaction and Usability Engineering of the Austrian Computer Society*.
- [117] Yuki Kubo, Yuto Koguchi, Buntarou Shizuki, Shin Takahashi, and Otmar Hilliges. 2019. AudioTouch : Minimally Invasive Sensing of Micro-Gestures via Active Bio-Acoustic Sensing. In *Proceedings of the 21st International Conference on Human-Computer Interaction with Mobile Devices and Services* (Taipei, Taiwan) (*MobileHCI '19*). Association for Computing Machinery, New York, NY, USA, Article 36, 13 pages. <https://doi.org/10.1145/3338286.3340147>

- [118] Li Chieh Kuo, Haw Yen Chiu, Cheung Wen Chang, Hsiu Yun Hsu, and Yun Nien Sun. 2009. Functional workspace for precision manipulation between thumb and fingers in normal hands. *Journal of Electromyography and Kinesiology* 19, 5 (oct 2009), 829–839. <https://doi.org/10.1016/J.JELEKIN.2008.07.008>
- [119] Christophe Lachenal. 2004. *Modèle et infrastructure logicielle pour l'interaction multi-instrument multisurface*. Thèses et habilitations. 200 pages..
- [120] R. Lachman, J. L. Lachman, and E. C. Butterfield. 2015. *Cognitive Psychology and Information Processing* (0 ed.). Psychology Press. <https://doi.org/10.4324/9781315798844>
- [121] Ecosse L. Lamoureux, Julie F. Pallant, Konrad Pesudovs, Gwyn Rees, Jennifer B. Hassell, and Jill E. Keeffe. 2007. The Impact of Vision Impairment Questionnaire : An Assessment of Its Domain Structure Using Confirmatory Factor Analysis and Rasch Analysis. *Investigative Ophthalmology and Visual Science* 48, 3 (2007), 1001–1006. <https://doi.org/10.1167/iovs.06-0361> arXiv :https://arvojournals.org/arvo/content_public/journal/iovs/932941/z7g00307001001.pdf
- [122] Florian Lang, Verena Pues, Albrecht Schmidt, and Tonja-Katrin Machulla. 2023. Braille-Buddy : A Tangible User Interface to Support Children with Visual Impairment in Learning Braille. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) (*CHI '23*). Association for Computing Machinery, New York, NY, USA, Article 218, 14 pages.
- [123] Ahreum Lee, Kiburm Song, Hokyoung Blake Ryu, Jieun Kim, and Gyuhyun Kwon. 2015. Fingerstroke time estimates for touchscreen-based mobile gaming interaction. *Human movement science* 44 (2015), 211–224.
- [124] Do Young Lee, Soo Hwan Lee, and Ian Oakley. 2020. Nailz : Sensing Hand Input with Touch Sensitive Nails. *Conference on Human Factors in Computing Systems - Proceedings* (2020). <https://doi.org/10.1145/3313831.3376778>
- [125] Garam Lee, Luis Cavazos Quero, Jing Yang, Hyunhee Jung, Jooyoung Son, and Jundong

- Cho. 2017. Slate Master : A Tangible Braille Slate Tutor for Mobile Devices. In *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services* (Vienna, Austria) (*MobileHCI '17*). Association for Computing Machinery, New York, NY, USA, Article 108, 6 pages. <https://doi.org/10.1145/3098279.3122151>
- [126] Johnny Chung Lee. 2010. In Search of a Natural Gesture. *XRDS* 16, 4 (jun 2010), 9–12. <https://doi.org/10.1145/1764848.1764853>
- [127] Lik Hang Lee, Kit Yung Lam, Tong Li, Tristan Braud, Xiang Su, and Pan Hui. 2019. Quad-metric Optimized Thumb-to-Finger Interaction for Force Assisted One-Handed Text Entry on Mobile Headsets. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 3, 3, Article 94 (2019), 27 pages. <https://doi.org/10.1145/3351252>
- [128] Frank Chun Yat Li, David Dearman, and Khai N. Truong. 2009. Virtual Shelves : Interactions with Orientation Aware Devices. In *Proceedings of the 22nd Annual ACM Symposium on User Interface Software and Technology* (Victoria, BC, Canada) (*UIST '09*). Association for Computing Machinery, New York, NY, USA, 125–128. <https://doi.org/10.1145/1622176.1622200>
- [129] Peter H. Lindsay and Donald A. Norman. 1972. *Human information processing : An introduction to psychology*. Academic Press, San Diego, CA, US. Pages : xxvii, 737.
- [130] Roman Lissermann, Jochen Huber, Aristotelis Hadjakos, and Max Mühlhäuser. 2013. EarPut : Augmenting behind-the-Ear Devices for Ear-Based Interaction. In *CHI '13 Extended Abstracts on Human Factors in Computing Systems* (Paris, France) (*CHI EA '13*). Association for Computing Machinery, New York, NY, USA, 1323–1328. <https://doi.org/10.1145/2468356.2468592>
- [131] Gonçalo Ferreira Lobo, David Gonçalves, Pedro Pais, Tiago Guerreiro, and André Rodrigues. 2021. *Using Games to Practice Screen Reader Gestures*. Association for Computing Machinery, New York, NY, USA.
- [132] Jack M. Loomis, Roberta L. Klatzky, and Susan J. Lederman. 1991. Similarity of Tactual

- and Visual Picture Recognition with Limited Field of View. *Perception* 20 (1991), 167 – 177.
- [133] Jack M. Loomis and Susan J. Lederman. 1986. Tactual perception. *Handbook of perception and human performance, Vol. 2 : Cognitive processes and performance*. (1986), 1–41. Place : Oxford, England Publisher : John Wiley & Sons.
- [134] Irene Lopatovska, Katrina Rink, Ian Knight, Kieran Raines, Kevin Cosenza, Harriet Williams, Perachya Sorsche, David Hirsch, Qi Li, and Adrianna Martinez. 2018. Talk to me : Exploring user interactions with the Amazon Alexa. *Journal of Librarianship and Information Science* 51 (03 2018), 096100061875941. <https://doi.org/10.1177/0961000618759414>
- [135] Zhihan Lu, Muhammad Sikandar Lal Khan, and Shafiq Ur Réhman. 2013. Hand and Foot Gesture Interaction for Handheld Devices. In *Proceedings of the 21st ACM International Conference on Multimedia (Barcelona, Spain) (MM '13)*. Association for Computing Machinery, New York, NY, USA, 621–624. <https://doi.org/10.1145/2502081.2502163>
- [136] Ewa Luger and Abigail Sellen. 2016. "Like Having a Really Bad PA" : The Gulf between User Expectation and Experience of Conversational Agents. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (San Jose, California, USA) (CHI '16)*. Association for Computing Machinery, New York, NY, USA, 5286–5297. <https://doi.org/10.1145/2858036.2858288>
- [137] Vikas Luthra and Sanjay Ghosh. 2015. Understanding, Evaluating and Analyzing Touch Screen Gestures for Visually Impaired Users in Mobile Environment. https://doi.org/10.1007/978-3-319-20681-3_3
- [138] K. G. Maheswari, R. K. Meenakshi, Ganapathi Nalinipriya, K Anandasayanam, B Hariram, and G Maheswara Pandian. 2022. Dynamic AI based Email Voice Assistant for Web Services. *2022 International Conference on Smart Technologies and Systems for Next Generation Computing (ICSTSN) (2022)*, 1–4. <https://api.semanticscholar.org/CorpusID:248406769>

- [139] Sebastián Marichal, Andrea Rosales, Gustavo Sansone, Ana Cristina Pires, Ewelina Bakala, Fernando Gonzalez Perilli, Bruno Fleischer, and Josep Blat. 2018. LETSmath. In *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct (Barcelona, Spain) (MobileHCI '18)*. Association for Computing Machinery, New York, NY, USA, 313–320. <https://doi.org/10.1145/3236112.3236157>
- [140] Natalina Martiniello, Werner Eisenbarth, Christine M. Lehane, Aaron P Johnson, and Walter Wittich. 2019. Exploring the use of smartphones and tablets among people with visual impairments : Are mainstream devices replacing the use of traditional visual aids ? *Assistive Technology* 34 (2019), 34 – 45.
- [141] Denys J. C. Matthies, Bodo Urban, Katrin Wolf, and Albrecht Schmidt. 2019. Reflexive Interaction : Extending the Concept of Peripheral Interaction. In *Proceedings of the 31st Australian Conference on Human-Computer-Interaction (Fremantle, WA, Australia) (OZCHI'19)*. Association for Computing Machinery, New York, NY, USA, 266–278. <https://doi.org/10.1145/3369457.3369478>
- [142] David McGookin, Stephen Brewster, and WeiWei Jiang. 2008. Investigating Touchscreen Accessibility for People with Visual Impairments. In *Proceedings of the 5th Nordic Conference on Human-Computer Interaction : Building Bridges (Lund, Sweden) (NordiCHI '08)*. Association for Computing Machinery, New York, NY, USA, 298–307. <https://doi.org/10.1145/1463160.1463193>
- [143] David McGookin, Euan Robertson, and Stephen Brewster. 2010. Clutching at Straws : Using Tangible Interaction to Provide Non-Visual Access to Graphs. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (Atlanta, Georgia, USA) (CHI '10)*. Association for Computing Machinery, New York, NY, USA, 1715–1724. <https://doi.org/10.1145/1753326.1753583>
- [144] Giuseppe Melfi, Karin Müller, Thorsten Schwarz, Gerhard Jaworek, and Rainer Stiefelhagen. 2020. Understanding What You Feel : A Mobile Audio-Tactile System for Graphics Used at Schools with Students with Visual Impairment. In *Proceedings of the*

- 2020 CHI Conference on Human Factors in Computing Systems (Honolulu, HI, USA) (CHI '20). Association for Computing Machinery, New York, NY, USA, 1–12. <https://doi.org/10.1145/3313831.3376508>
- [145] Oussama Metatla, Fiore Martin, Tony Stockman, and Nick Bryan-Kinns. 2014. "Non-Visual Menu Navigation : the Effect of an Audio-Tactile Display". 213–217. <https://doi.org/10.14236/ewic/hci2014.25>
- [146] Michael Minge, Manfred Thuring, Ingmar Wagner, and Carina Kuhr. 2017. The me-CUE Questionnaire : A Modular Tool for Measuring User Experience, Vol. 486. 115–128. https://doi.org/10.1007/978-3-319-41685-4_11
- [147] J. Morris and J. Mueller. 2014. Blind and Deaf Consumer Preferences for Android and iOS Smartphones. In *Inclusive Designing*, P. M. Langdon, J. Lazar, A. Heylighen, and H. Dong (Eds.). Springer International Publishing, Cham, 69–79.
- [148] Meredith Ringel Morris, Jacob O. Wobbrock, and Andrew D. Wilson. 2010. Understanding Users' Preferences for Surface Gestures. In *Proceedings of Graphics Interface 2010 (Ottawa, Ontario, Canada) (GI '10)*. Canadian Information Processing Society, CAN, 261–268.
- [149] Brad A. Myers. 1998. A Brief History of Human-Computer Interaction Technology. *Interactions* 5, 2 (1998), 44–54. <https://doi.org/10.1145/274430.274436>
- [150] Suranga Nanayakkara, Roy Shilkrot, and Pattie Maes. 2012. EyeRing : A Finger-Worn Assistant. In *CHI '12 Extended Abstracts on Human Factors in Computing Systems (Austin, Texas, USA) (CHI EA '12)*. Association for Computing Machinery, New York, NY, USA, 1961–1966. <https://doi.org/10.1145/2212776.2223736>
- [151] Tao Ni, Doug Bowman, and Chris North. 2011. AirStroke : Bringing unistroke text entry to freehand gesture interfaces. 2473–2476. <https://doi.org/10.1145/1978942.1979303>
- [152] Tao Ni, Doug A. Bowman, Chris North, and Ryan P. McMahan. 2011. Design and Evaluation of Freehand Menu Selection Interfaces Using Tilt and Pinch Gestures. *Int. J. Hum.-*

BIBLIOGRAPHIE

- Comput. Stud.* 69, 9 (2011), 551–562. <https://doi.org/10.1016/j.ijhcs.2011.05.001>
- [153] Jakob Nielsen. 1994. *Heuristic Evaluation*. John Wiley amp ; Sons, Inc., USA, 25–62.
- [154] J. Nielsen. 1994. *Usability Engineering*. Elsevier Science. <https://books.google.fr/books?id=95As20F67f0C>
- [155] Michael Nielsen, Moritz Stoerring, Thomas Moeslund, and Erik Granum. 2003. A Procedure for Developing Intuitive and Ergonomic Gesture Interfaces for HCI. *LNCS (LNAI)*, 409–420. https://doi.org/10.1007/978-3-540-24598-8_38
- [156] Laurence Nigay and Joëlle Coutaz. 1996. Espaces conceptuels pour l'interaction multi-média et multimodale. *Technique Et Science Informatiques - TSI* 15 (1996).
- [157] Sherly Noel. 2020. Human computer interaction(HCI) based Smart Voice Email (Vmail) Application - Assistant for Visually Impaired Users (VIU). In *2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT)*. 895–900.
- [158] Donald A. Norman. 2007. The next UI Breakthrough, Part 2 : Physicality. *Interactions* 14, 4 (jul 2007), 46–47. <https://doi.org/10.1145/1273961.1273986>
- [159] David J. Osborne. 1985. *Computers at Work : A Behavioural Approach*. <https://api.semanticscholar.org/CorpusID:60065854>
- [160] João Oliveira, Tiago Guerreiro, Hugo Nicolau, Joaquim Jorge, and Daniel Gonçalves. 2011. Blind People and Mobile Touch-Based Text-Entry : Acknowledging the Need for Different Flavors. In *The Proceedings of the 13th International ACM SIGACCESS Conference on Computers and Accessibility (Dundee, Scotland, UK) (ASSETS '11)*. Association for Computing Machinery, New York, NY, USA, 179–186. <https://doi.org/10.1145/2049536.2049569>
- [161] M Sile O'Modhrain, Nicholas A Giudice, John A Gardner, and Gordon E Legge. 2015. Designing Media for Visually-Impaired Users of Refreshable Touch Displays : Possibilities and Pitfalls. *IEEE Transactions on Haptics* 8 (2015), 248–257.

BIBLIOGRAPHIE

- [162] APA Committee on Disability Issues in Psychology. 1992. Guidelines for Nonhandicapping Language in APA Journals. <http://www.apastyle.org/manual/related/nonhandicapping-language.aspx> Consulté le 22 Mai 2023..
- [163] World Health Organization. [n. d.]. International Classification of Diseases, 11th revision (ICD(11)). <https://www.who.int/classifications/icd/en/> Consulté le 22 mai 2023.
- [164] World Health Organization. [n. d.]. International classification of functioning, disability and health : ICF. <https://icd.who.int/dev11/1-icf/en> Consulté le 22 mai 2023.
- [165] World Health Organization. 2020. *Rapport mondial sur la vision*. Organisation mondiale de la Santé. xviii, 170 p. pages.
- [166] Godwin O. Ovenseri-Ogbomo, Harriette Osafo-Agyei, Ralph E.U. Akpalaba, James Addy, and Elizabeth O. Ovenseri. 2016. Impact of low vision services on the quality of life of low vision patients in Ghana. *African Vision and Eye Health* 75, 1 (2016), 5 pages. <https://doi.org/10.4102/aveh.v75i1.19>
- [167] Joyojeet Pal, Anandhi Viswanathan, Priyank Chandra, Anisha Nazareth, Vaishnav Kameswaran, Hariharan Subramonyam, Aditya Johri, Mark S. Ackerman, and Sile O'Modhrain. 2017. Agency in Assistive Technology Adoption : Visual Impairment and Smartphone Use in Bangalore. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) (CHI '17). Association for Computing Machinery, New York, NY, USA, 5929–5940.
- [168] Hari prasath Palani, Paul Fink, and Nicholas Giudice. 2020. Design Guidelines for Schematizing and Rendering Haptically Perceivable Graphical Elements on Touchscreen Devices. *International Journal of Human-Computer Interaction* 36 (2020), 1–22. <https://doi.org/10.1080/10447318.2020.1752464>
- [169] Hemanth Kumar Panikala. 2021. Voice Based Email For Visually Challenged People. *SSRN Electronic Journal* (2021). <https://doi.org/10.2139/ssrn.4132651>
- [170] Chris Perkins and Ann Gardiner. 2003. Real world map reading strategies. *Cartographic Journal* 40, 3 (2003), 265–268. <https://doi.org/10.1179/000870403225012970>

BIBLIOGRAPHIE

- [171] Simon Perrault, Eric Lecolinet, James Eagan, and Yves Guiard. 2013. WatchIt : Simple gestures and eyes-free interaction for wristwatches and bracelets. 1451–1460. <https://doi.org/10.1145/2470654.2466192>
- [172] Madeline Phillips and Michael J. Proulx. 2018. Social Interaction Without Vision : An Assessment of Assistive Technology for the Visually Impaired. *Technology and Innovation* (2018). <https://api.semanticscholar.org/CorpusID:149653942>
- [173] Munyaporn Pooripanyakun, Andrew Wodehouse, and Jorn Mehnen. 2022. The Design and Evaluation of an Eyes-Free Touchscreen Interface. <https://doi.org/10.54941/ahfe100915>
- [174] R. L. Potter, L. J. Weldon, and B. Shneiderman. 1988. Improving the Accuracy of Touch Screens : An Experimental Evaluation of Three Strategies. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Washington, D.C., USA) (*CHI '88*). Association for Computing Machinery, New York, NY, USA, 27–32. <https://doi.org/10.1145/57167.57171>
- [175] Ivan Poupyrev, Shigeaki Maruyama, and Jun Rekimoto. 2002. Ambient Touch : Designing Tactile Interfaces for Handheld Devices. In *Proceedings of the 15th Annual ACM Symposium on User Interface Software and Technology* (Paris, France) (*UIST '02*). Association for Computing Machinery, New York, NY, USA, 51–60. <https://doi.org/10.1145/571985.571993>
- [176] Ivan Poupyrev, Tatsushi Nashida, and Makoto Okabe. 2007. Actuation and Tangible User Interfaces : The Vaucanson Duck, Robots, and Shape Displays. In *Proceedings of the 1st International Conference on Tangible and Embedded Interaction* (Baton Rouge, Louisiana) (*TEI '07*). Association for Computing Machinery, New York, NY, USA, 205–212.
- [177] Alisha Pradhan, Kanika Mehta, and Leah Findlater. 2018. "Accessibility Came by Accident" : Use of Voice-Controlled Intelligent Personal Assistants by People with Disabilities. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) (*CHI '18*). Association for Computing Machinery, New York, NY, USA, 1–13. <https://doi.org/10.1145/3173574.3174033>

- [178] Manuel Prätorius, Dimitar Valkov, Ulrich Burgbacher, and Klaus Hinrichs. 2014. DigiTap : An Eyes-Free VR/AR Symbolic Input Device. In *Proceedings of the 20th ACM Symposium on Virtual Reality Software and Technology (Edinburgh, Scotland) (VRST '14)*. Association for Computing Machinery, New York, NY, USA, 9–18. <https://doi.org/10.1145/2671015.2671029>
- [179] Halley P. Profita, James Clawson, Scott Gilliland, Clint Zeagler, Thad Starner, Jim Budd, and Ellen Yi-Luen Do. 2013. Don't Mind Me Touching My Wrist : A Case Study of Interacting with on-Body Technology in Public. In *Proceedings of the 2013 International Symposium on Wearable Computers (Zurich, Switzerland) (ISWC '13)*. Association for Computing Machinery, New York, NY, USA, 89–96. <https://doi.org/10.1145/2493988.2494331>
- [180] Mahfuz Rahman, Sean Gustafson, Pourang Irani, and Sriram Subramanian. 2009. Tilt Techniques : Investigating the Dexterity of Wrist-Based Input. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (Boston, MA, USA) (CHI '09)*. Association for Computing Machinery, New York, NY, USA, 1943–1952. <https://doi.org/10.1145/1518701.1518997>
- [181] H. Rudy Ramsey and Michael E. Atwood. 1979. Human Factors in Computer Systems : A Review of the Literature. <https://api.semanticscholar.org/CorpusID:60622169>
- [182] Yosra Rekik, Radu-Daniel Vatavu, and Laurent Grisoni. 2014. Understanding Users' Perceived Difficulty of Multi-Touch Gesture Articulation. In *ICMI'14, the 16th ACM International Conference on Multimodal Interaction*. Istanbul, Turkey. <https://doi.org/10.1145/2663204.2663273>
- [183] Matt Rice, R. Daniel Jacobson, Reginald G. Golledge, and David Jones. 2005. Design Considerations for Haptic and Auditory Map Interfaces. *Cartography and Geographic Information Science* 32, 4 (2005), 381–391.
- [184] Julie Rico and Stephen Brewster. 2009. Gestures All around Us : User Differences in Social Acceptability Perceptions of Gesture Based Interfaces. In *Proceedings of the 11th*

- International Conference on Human-Computer Interaction with Mobile Devices and Services* (Bonn, Germany) (*MobileHCI '09*). Association for Computing Machinery, New York, NY, USA, Article 64, 2 pages. <https://doi.org/10.1145/1613858.1613936>
- [185] Julie Rico and Stephen Brewster. 2010. Usable Gestures for Mobile Interfaces : Evaluating Social Acceptability. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Atlanta, Georgia, USA) (*CHI '10*). Association for Computing Machinery, New York, NY, USA, 887–896. <https://doi.org/10.1145/1753326.1753458>
- [186] André Rodrigues, Kyle Montague, Hugo Nicolau, and Tiago Guerreiro. 2015. Getting Smartphones to Talkback : Understanding the Smartphone Adoption Process of Blind Users. In *Proceedings of the 17th International ACM SIGACCESS Conference on Computers and Accessibility* (*ASSETS '15*). Association for Computing Machinery, New York, NY, USA, 23–32. <https://doi.org/10.1145/2700648.2809842>
- [187] M. C. Rodriguez-Sanchez, M. A. Moreno-Alvarez, E. Martin, S. Borromeo, and J. A. Hernandez-Tamames. 2014. Accessible smartphones for blind users : A case study for a wayfinding system. *Expert Systems with Applications* 41, 16 (2014), 7210–7222. <https://doi.org/10.1016/j.eswa.2014.05.031>
- [188] Marco Romano, Andrea Bellucci, and Ignacio Aedo. 2015. Understanding Touch and Motion Gestures for Blind People on Mobile Devices, Vol. 9296. 38–46. https://doi.org/10.1007/978-3-319-22701-6_3
- [189] Sherry Ruan, Jacob O. Wobbrock, Kenny Liou, Andrew Ng, and James A. Landay. 2018. Comparing Speech and Keyboard Text Entry for Short Messages in Two Languages on Touchscreen Phones. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 1, 4, Article 159 (2018), 23 pages.
- [190] Jaime Ruiz, Yang Li, and Edward Lank. 2011. User-Defined Motion Gestures for Mobile Interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Vancouver, BC, Canada) (*CHI '11*). Association for Computing Machinery, New York, NY, USA, 197–206. <https://doi.org/10.1145/1978942.1978971>

BIBLIOGRAPHIE

- [191] Barbara Ryan. 2014. Models of low vision care : past, present and future. *Clinical and Experimental Optometry* 97, 3 (2014), 209–213.
- [192] G. Révész. 1950. *Psychology and art of the blind*. Longmans, Green, Oxford, England. Pages : xiv, 338.
- [193] Debasis Samanta and Tuhin Chakraborty. 2020. VectorEntry : Text Entry Mechanism Using Handheld Touch-Enabled Mobile Devices for People with Visual Impairments. *ACM Trans. Access. Comput.* 13, 3, Article 12 (aug 2020), 29 pages. <https://doi.org/10.1145/3406537>
- [194] Javier Sánchez and J. R. Togores. 2012. Designing Mobile Apps for Visually Impaired and Blind Users. In *International Conference on Advances in Computer-Human Interaction*.
- [195] Neil Sarkar, B. O’Hanlon, Geoffrey Lee, Duncan Strathearn, Mahdi Olfat, Arash Rohani, and Raafat R. Mansour. 2017. Optical MEMS index finger microgesture input sensor for mobile and wearable devices. *2017 19th International Conference on Solid-State Sensors, Actuators and Microsystems (TRANSDUCERS) (2017)*, 250–253. <https://api.semanticscholar.org/CorpusID:1522643>
- [196] Marcia J. Scherer and Rob Glueckauf. 2005. Assessing the Benefits of Assistive Technologies for Activities and Participation. *Rehabilitation Psychology* 50 (2005), 132–141. Place : US Publisher : Educational Publishing Foundation.
- [197] Oliver Schneider, Jotaro Shigeyama, Robert Kovacs, Thijs Jan Roumen, Sebastian Marwecki, Nico Boeckhoff, Daniel Amadeus Gloeckner, Jonas Bounama, and Patrick Baudisch. 2018. DualPanto : A Haptic Device That Enables Blind Users to Continuously Interact with Virtual Worlds. In *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology (Berlin, Germany) (UIST ’18)*. Association for Computing Machinery, New York, NY, USA, 877–887. <https://doi.org/10.1145/3242587.3242604>
- [198] Adriano Scoditti, Thomas Vincent, Joelle Coutaz, Renaud Blanch, and Nadine Mandran. 2011. TouchOver : Decoupling Positioning from Selection on Touch-Based Handheld

BIBLIOGRAPHIE

- Devices. In *Proceedings of the 23rd Conference on l'Interaction Homme-Machine* (Sophia Antipolis, France) (*IHM '11*). Association for Computing Machinery, New York, NY, USA, Article 6, 4 pages. <https://doi.org/10.1145/2044354.2044362>
- [199] Jeremy Scott, David Dearman, Koji Yatani, and Khai N. Truong. 2010. Sensing Foot Gestures from the Pocket. In *Proceedings of the 23rd Annual ACM Symposium on User Interface Software and Technology* (New York, New York, USA) (*UIST '10*). Association for Computing Machinery, New York, NY, USA, 199–208. <https://doi.org/10.1145/1866029.1866063>
- [200] Andrew Sears and Mark Young. 2003. Physical Disabilities and Computing Technologies : An Analysis Of Impairments. (2003), 482–503.
- [201] Marcos Serrano. 2010. *Interaction multimodale en entrée : conception et prototypage*. Ph.D. Dissertation. <http://www.theses.fr/2010GREN033> Thèse de doctorat dirigée par Nigay, Laurence Informatique Grenoble 2010.
- [202] Orit Shaer and Eva Hornecker. 2010. Tangible User Interfaces : Past, Present, and Future Directions. *Found. Trends Hum.-Comput. Interact.* 3, 1–2 (jan 2010), 1–137.
- [203] Adwait Sharma, Joan Sol Roo, and Jürgen Steimle. 2019. Grasping Microgestures : Eliciting Single-Hand Microgestures for Handheld Objects. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19, Chi)*. Association for Computing Machinery, New York, NY, USA, 1–13. <https://doi.org/10.1145/3290605.3300632>
- [204] Kristen Shinohara and Jacob O. Wobbrock. 2016. Self-Conscious or Self-Confident? A Diary Study Conceptualizing the Social Accessibility of Assistive Technology. 8, 2 (2016).
- [205] B. Shneiderman. 1987. *Direct Manipulation : A Step beyond Programming Languages*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 461–467.
- [206] B. Shneiderman. 1991. Touch screens now offer compelling uses. *IEEE Software* 8, 2 (1991), 93–94. <https://doi.org/10.1109/52.73754>

- [207] Ben Shneiderman. 1997. Direct Manipulation for Comprehensible, Predictable and Controllable User Interfaces. In *Proceedings of the 2nd International Conference on Intelligent User Interfaces (Orlando, Florida, USA) (IUI '97)*. Association for Computing Machinery, New York, NY, USA, 33–39. <https://doi.org/10.1145/238218.238281>
- [208] Mohamed Soliman, Franziska Mueller, Lena Hegemann, Joan Sol Roo, Christian Theobalt, and Jürgen Steimle. 2018. FingerInput : Capturing Expressive Single-Hand Thumb-to-Finger Microgestures. *Proceedings of the 2018 ACM International Conference on Interactive Surfaces and Spaces* 18, December (2018), 177–187. <https://doi.org/10.1145/3279778.3279799>
- [209] Caleb Southern, James Clawson, Brian Frey, Gregory Abowd, and Mario Romero. 2012. An Evaluation of BrailleTouch : Mobile Touchscreen Text Entry for the Visually Impaired. In *Proceedings of the 14th International Conference on Human-Computer Interaction with Mobile Devices and Services (San Francisco, California, USA) (MobileHCI '12)*. Association for Computing Machinery, New York, NY, USA, 317–326. <https://doi.org/10.1145/2371574.2371623>
- [210] Joan Stelmack. 2001. Quality of Life of Low-Vision Patients and Outcomes of Low-Vision Rehabilitation :. *Optometry and Vision Science* 78, 5 (2001), 335–342. <https://doi.org/10.1097/00006324-200105000-00017>
- [211] Calkin Suero Montero, Jason Alexander, Mark Marshall, and Sriram Subramanian. 2010. Would you do that? - Understanding social acceptance of gestural interfaces. *ACM International Conference Proceeding Series*, 275–278. <https://doi.org/10.1145/1851600.1851647>
- [212] Yanke Tan, Sang Ho Yoon, and Karthik Ramani. 2017. BikeGesture : User Elicitation and Performance of Micro Hand Gesture as Input for Cycling. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '17, Vol. Part F1276)*. Association for Computing Machinery, New York, NY, USA, 2147–2154. <https://doi.org/10.1145/3027063.3053075>

BIBLIOGRAPHIE

- [213] Aleksandra Taniberg, Lars Botin, and Kashmiri Stec. 2018. Context of use affects the social acceptability of gesture interaction. *NordiCHI '18 : Proceedings of the 10th Nordic Conference on Human-Computer Interaction*, 731–735. <https://doi.org/10.1145/3240167.3240250>
- [214] Emma Tebbutt, Rebecca Brodmann, Johan Borg, Malcolm MacLachlan, Chapal Khasnabis, and Robert Horvath. 2016. Assistive products and the Sustainable Development Goals (SDGs). *Globalization and Health* 12, 1 (2016), 79. <https://doi.org/10.1186/s12992-016-0220-6>
- [215] Tribhuwan Kumar Tewari, Anshul Arya, and Sameer Rastogi. 2010. Message Reading Through Eye Blinking. *International Journal of Computer Applications* 2, 6 (2010), 1–4. <https://doi.org/10.5120/676-951>
- [216] Hussain Tinwala and I. Scott MacKenzie. 2010. Eyes-Free Text Entry with Error Correction on Touchscreen Mobile Devices. In *Proceedings of the 6th Nordic Conference on Human-Computer Interaction : Extending Boundaries* (Reykjavik, Iceland) (*NordiCHI '10*). Association for Computing Machinery, New York, NY, USA, 511–520. <https://doi.org/10.1145/1868914.1868972>
- [217] Paulus A. Tiwari, Pratiksha Zodawar, Harsha P. Nimkar, Trishna Rotke, Priya G. Wanjari, and Umesh Samarth. 2020. A Review on Voice based E-Mail System for Blind. In *2020 International Conference on Inventive Computation Technologies (ICICT)*. 435–438.
- [218] Stephen Tregear. 2004. Vision Rehabilitation for Elderly Individuals with Low Vision or Blindness.
- [219] Hsin Ruey Tsai, Lee Ting Huang, Cheng Yuan Wu, and Yi Ping Hung. 2016. ThumbRing : Private interactions using one-handed thumb motion input on finger segments. *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct, MobileHCI 2016* (2016), 791–798. <https://doi.org/10.1145/2957265.2961859>
- [220] Hsin Ruey Tsai, Te Yen Wu, Min Chieh Hsiu, Jui Chun Hsiao, Da Yuan Huang, Yi Ping

- Hung, Mike Y. Chen, and Bing Yu Chen. 2017. SegTouch : Enhancing touch input while providing touch gestures on screens using thumb-to-index-finger gestures. In *Conference on Human Factors in Computing Systems - Proceedings*, Vol. Part F1276. ACM, New York, NY, USA, 2164–2171. <https://doi.org/10.1145/3027063.3053109>
- [221] Janice Y. Tsai, Tawfiq Ammari, Abraham Wallin, and Jofish Kaye. 2018. Alexa, play some music : Categorization of Alexa Commands. In *Voice-based Conversational UX Studies and Design Wokrshop at CHI*.
- [222] Gregg Vanderheiden and J. Jordan. 2012. *Design for People with Functional Limitations*. 1407–1441. <https://doi.org/10.1002/9781118131350.ch51>
- [223] Gregg C. Vanderheiden. 1996. Use of Audio-Haptic Interface Techniques to Allow Nonvisual Access to Touchscreen Appliances. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 40, 24 (1996), 1266–1266. <https://doi.org/10.1177/154193129604002430>
- [224] Tomaso Vecchi, Carla Tinti, and Cesare Cornoldi. 2004. Spatial memory and integration processes in congenital blindness. *NeuroReport : For Rapid Communication of Neuroscience Research* 15, 18 (2004), 2787–2790. Place : US Publisher : Lippincott Williams & Wilkins.
- [225] Eduardo Velloso, Jason Alexander, Andreas Bulling, and Hans Gellersen. 2015. Interactions Under the Desk : A Characterisation of Foot Movements for Input in a Seated Position, Vol. 9296. 384–401. https://doi.org/10.1007/978-3-319-22701-6_29
- [226] Aaron Visschedijk, Hyunyoung Kim, Carlos Tejada, and Daniel Ashbrook. 2022. ClipWidgets : 3D-Printed Modular Tangible UI Extensions for Smartphones. In *Sixteenth International Conference on Tangible, Embedded, and Embodied Interaction* (Daejeon, Republic of Korea) (TEI '22). Association for Computing Machinery, New York, NY, USA, Article 8, 11 pages. <https://doi.org/10.1145/3490149.3501314>
- [227] Daniel Vogel and Ravin Balakrishnan. 2005. Distant Freehand Pointing and Clicking on Very Large, High Resolution Displays. In *Proceedings of the 18th Annual ACM Sympo-*

- sium on User Interface Software and Technology* (Seattle, WA, USA) (*UIST '05*). Association for Computing Machinery, New York, NY, USA, 33–42. <https://doi.org/10.1145/1095034.1095041>
- [228] Alexandra Vtyurina, Adam Fourney, Meredith Ringel Morris, Leah Findlater, and Ryen W. White. 2019. Bridging Screen Readers and Voice Assistants for Enhanced Eyes-Free Web Search. In *The World Wide Web Conference* (San Francisco, CA, USA) (*WWW '19*). Association for Computing Machinery, New York, NY, USA, 3590–3594. <https://doi.org/10.1145/3308558.3314136>
- [229] Alexandra Vtyurina, Adam Fourney, Meredith Ringel Morris, Leah Findlater, and Ryen W. White. 2019. VERSE : Bridging Screen Readers and Voice Assistants for Enhanced Eyes-Free Web Search. In *Proceedings of the 21st International ACM SIGACCESS Conference on Computers and Accessibility* (Pittsburgh, PA, USA) (*ASSETS '19*). Association for Computing Machinery, New York, NY, USA, 414–426. <https://doi.org/10.1145/3308561.3353773>
- [230] Jérémy Wambecke, Alix Goguey, Laurence Nigay, Lauren Dargent, Daniel Hauret, Stéphanie Lafon, and Jean Samuel Louis De Visme. 2021. M[eye]cro : Eye-gaze+Microgestures for Multitasking and Interruptions. *Proceedings of the ACM on Human-Computer Interaction* 5, EICS (2021), 22. <https://doi.org/10.1145/3461732>
- [231] Cheng-Yao Wang, Min-Chieh Hsiu, Po-Tsung Chiu, Chiao-Hui Chang, Liwei Chan, Bing-Yu Chen, and Mike Y. Chen. 2015. PalmGesture : Using Palms as Gesture Interfaces for Eyes-Free Input. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services* (Copenhagen, Denmark) (*MobileHCI '15*). Association for Computing Machinery, New York, NY, USA, 217–226. <https://doi.org/10.1145/2785830.2785885>
- [232] Saiwen Wang, Jie Song, Jaime Lien, Ivan Poupyrev, and Otmar Hilliges. 2016. Interacting with Soli : Exploring Fine-Grained Dynamic Gesture Recognition in the Radio-Frequency Spectrum. In *Proceedings of the 29th Annual Symposium on User Interface Software and*

- Technology* (Tokyo, Japan) (*UIST '16*). Association for Computing Machinery, New York, NY, USA, 851–860. <https://doi.org/10.1145/2984511.2984565>
- [233] David Way. 2014. EMGRIE : Ergonomic Microgesture Recognition and Interaction Evaluation, a case study. (12 2014).
- [234] Martin Weigel, Tong Lu, Gilles Bailly, Antti Oulasvirta, Carmel Majidi, and Jürgen Steimle. 2015. ISkin : Flexible, Stretchable and Visually Customizable On-Body Touch Sensors for Mobile Computing. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (Seoul, Republic of Korea) (*CHI '15*). Association for Computing Machinery, New York, NY, USA, 2991–3000. <https://doi.org/10.1145/2702123.2702391>
- [235] Martin Weigel, Aditya Shekhar Nittala, Alex Olwal, and Jürgen Steimle. 2017. SkinMarks : Enabling Interactions on Body Landmarks Using Conformal Skin Electronics. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) (*CHI '17*). Association for Computing Machinery, New York, NY, USA, 3095–3105. <https://doi.org/10.1145/3025453.3025704>
- [236] Mark Weiser. 1993. Some Computer Science Issues in Ubiquitous Computing. *Commun. ACM* 36, 7 (jul 1993), 75–84. <https://doi.org/10.1145/159544.159617>
- [237] Eric Whitmire, Mohit Jain, Divye Jain, Greg Nelson, Ravi Karkar, Shwetak Patel, and Mayank Goel. 2017. DigiTouch : Reconfigurable Thumb-to-Finger Input and Text Entry on Head-Mounted Displays. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 1, 3, Article 113 (2017), 21 pages. <https://doi.org/10.1145/3130978>
- [238] Daniel Wigdor and Dennis Wixon. 2011. *Brave NUI World : Designing Natural User Interfaces for Touch and Gesture* (1st ed.). Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.
- [239] Jacob O. Wobbrock, Meredith Ringel Morris, and Andrew D. Wilson. 2009. User-Defined Gestures for Surface Computing. In *Proceedings of the SIGCHI Conference on Human*

- Factors in Computing Systems* (Boston, MA, USA) (*CHI '09*). Association for Computing Machinery, New York, NY, USA, 1083–1092. <https://doi.org/10.1145/1518701.1518866>
- [240] Katrin Wolf. 2016. Microgestures—Enabling Gesture Input with Busy Hands. In *Peripheral Interaction : Challenges and Opportunities for HCI in the Periphery of Attention*, Saskia Bakker, Doris Hausen, and Ted Selker (Eds.). Springer International Publishing, Cham, 95–116. https://doi.org/10.1007/978-3-319-29523-7_5
- [241] Katrin Wolf, Anja Naumann, Michael Rohs, and Jörg Müller. 2011. Taxonomy of Microinteractions : Defining Microgestures Based on Ergonomic and Scenario-Dependent Requirements. In *Proceedings of the 13th IFIP TC 13 International Conference on Human-Computer Interaction - Volume Part I (INTERACT'11, Vol. 6946 LNCS)*. Springer-Verlag, Berlin, Heidelberg, 559–575. https://doi.org/10.1007/978-3-642-23774-4_45
- [242] Pui Chung Wong, Kening Zhu, and Hongbo Fu. 2018. FingerT9 : Leveraging Thumb-to-Finger Interaction for Same-Side-Hand Text Entry on Smartwatches. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) (*CHI '18*). Association for Computing Machinery, New York, NY, USA, 1–10. <https://doi.org/10.1145/3173574.3173752>
- [243] Mike Wu and Ravin Balakrishnan. 2003. Multi-Finger and Whole Hand Gestural Interaction Techniques for Multi-User Tabletop Displays. In *Proceedings of the 16th Annual ACM Symposium on User Interface Software and Technology* (Vancouver, Canada) (*UIST '03*). Association for Computing Machinery, New York, NY, USA, 193–202. <https://doi.org/10.1145/964696.964718>
- [244] Yiqi Xiao and Renke He. 2018. The Handlebar as an Input Field : Evaluating Finger Gestures Designed for Bicycle Riders. *Advances in Intelligent Systems and Computing* (2018). <https://api.semanticscholar.org/CorpusID:67382737>
- [245] Zheer Xu, Weihao Chen, Dongyang Zhao, Jiehui Luo, Te-Yen Wu, Jun Gong, Sicheng Yin, Jialun Zhai, and Xing-Dong Yang. 2020. BiTipText : Bimanual Eyes-Free Text Entry on a Fingertip Keyboard. In *Proceedings of the 2020 CHI Conference on Human Factors in*

- Computing Systems* (Honolulu, HI, USA) (*CHI '20*). Association for Computing Machinery, New York, NY, USA, 1–13. <https://doi.org/10.1145/3313831.3376306>
- [246] Zheer Xu, Pui Chung Wong, Jun Gong, Te-Yen Wu, Aditya Shekhar Nittala, Xiaojun Bi, Jürgen Steimle, Hongbo Fu, Kening Zhu, and Xing-Dong Yang. 2019. TipText : Eyes-Free Text Entry on a Fingertip Keyboard. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology* (New Orleans, LA, USA) (*UIST '19*). Association for Computing Machinery, New York, NY, USA, 883–899. <https://doi.org/10.1145/3332165.3347865>
- [247] Linghui Ye, Frode Eika Sandnes, and I. Scott MacKenzie. 2020. QB-Gest : Qwerty Bimanual Gestural Input for Eyes-Free Smartphone Text Input. In *Universal Access in Human-Computer Interaction. Design Approaches and Supporting Technologies : 14th International Conference, UAHCI 2020, Held as Part of the 22nd HCI International Conference, HCII 2020, Copenhagen, Denmark, July 19–24, 2020, Proceedings, Part I* (Copenhagen, Denmark). Springer-Verlag, Berlin, Heidelberg, 223–242. https://doi.org/10.1007/978-3-030-49282-3_16
- [248] Cheng Zhang, Anandghan Waghmare, Pranav Kundra, Yiming Pu, Scott Gilliland, Thomas Ploetz, Thad E. Starner, Omer T. Inan, and Gregory D. Abowd. 2017. FingerSound : Recognizing Unistroke Thumb Gestures Using a Ring. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 1, 3, Article 120 (2017), 19 pages. <https://doi.org/10.1145/3130985>
- [249] Mingrui Ray Zhang, Ruolin Wang, Xuhai Xu, Qisheng Li, Ather Sharif, and Jacob O. Wobbrock. 2021. Voicemoji : Emoji Entry Using Voice for Visually Impaired People (*CHI '21*). Association for Computing Machinery, New York, NY, USA, Article 37, 18 pages. <https://doi.org/10.1145/3411764.3445338>
- [250] Xiaoyi Zhang, Tracy Tran, Yuqian Sun, Ian Culhane, Shobhit Jain, James Fogarty, and Jennifer Mankoff. 2018. Interactiles : 3D Printed Tactile Interfaces to Enhance Mobile Touchscreen Accessibility. In *Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility* (Galway, Ireland) (*ASSETS '18*). Association

BIBLIOGRAPHIE

- for Computing Machinery, New York, NY, USA, 131–142. <https://doi.org/10.1145/3234695.3236349>
- [251] Yang Zhang and Chris Harrison. 2015. Tomo : Wearable, Low-Cost Electrical Impedance Tomography for Hand Gesture Recognition. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software and Technology (Charlotte, NC, USA) (UIST '15)*. Association for Computing Machinery, New York, NY, USA, 167–173. <https://doi.org/10.1145/2807442.2807480>
- [252] Kaixing Zhao, Sandra Bardot, Marcos Serrano, Mathieu Simonnet, Bernard Oriola, and Christophe Jouffrais. 2021. Tactile Fixations : A Behavioral Marker on How People with Visual Impairments Explore Raised-Line Graphics. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (Yokohama, Japan) (CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 27, 12 pages. <https://doi.org/10.1145/3411764.3445578>
- [253] Yu Zhong, T. V. Raman, Casey Burkhardt, Fadi Biadsy, and Jeffrey P. Bigham. 2014. JustSpeak : Enabling Universal Voice Control on Android (*W4A '14*). Association for Computing Machinery, New York, NY, USA, Article 36, 4 pages. <https://doi.org/10.1145/2596695.2596720>

Annexe A

Annexe I

A.1 La déficience visuelle

En 2001, l'Organisation mondiale de la Santé (OMS) a développé la Classification Internationale du Fonctionnement, du Handicap et de la Santé (CIF) [164]. Cette classification a pour objectif d'établir un langage uniforme et normalisé ainsi qu'un cadre permettant la description des états de santé, incluant la vision, l'audition, la marche, l'apprentissage, la mémoire, et d'autres domaines connexes tels que la mobilité, l'éducation et les interactions sociales. Selon cette classification, et en accord avec la 11^{ème} édition de la Classification internationale des Maladies (CIM 11) [163], le terme "déficience" est employé de manière globale pour décrire un dysfonctionnement ou une altération de la structure corporelle associée à un problème de santé. Les déficiences visuelles sont donc définies comme une affection du système visuel altérant une ou plusieurs fonctions de la vision. Elles couvrent une large palette d'affections, tels que le daltonisme, la myopie, la presbytie, l'astigmatisme, la cécité nocturne, une sensibilité excessive à la lumière, une réduction ou des taches dans le champ visuel.

De manière générale cependant, les études basées sur la population se basent exclusivement sur l'acuité visuelle pour évaluer la déficience visuelle [165] : il s'agit d'une mesure évaluant la capacité du système visuel à distinguer deux points très contrastés dans l'espace.

A.1. LA DÉFICIENCE VISUELLE

L'acuité visuelle de loin est généralement évaluée en utilisant une échelle de vision (figure A.1) placée à une distance fixe (d'ordinaire, en France, 5 mètres). Elle est représentée par une fraction : le numérateur indique la distance à laquelle une personne peut lire la plus petite ligne de l'échelle, tandis que le dénominateur représente la distance à laquelle une personne ayant une vision considérée "normale" peut lire cette même ligne. Par exemple, une acuité visuelle de 3/10 signifie qu'au mieux, une personne peut lire une lettre issue de cette ligne à 5 mètres alors qu'une personne ayant une vision "normale" pourrait la lire à 16,5 mètres. Une acuité visuelle de 10/10 est donc considérée comme une vision "normale".

L'acuité visuelle de près est mesurée en fonction de la plus petite taille de police pouvant être discernée par une personne à une distance de test donnée. Elle est établie par la difficulté à discerner des caractères d'imprimerie de police 6 à 40 centimètres de distance.

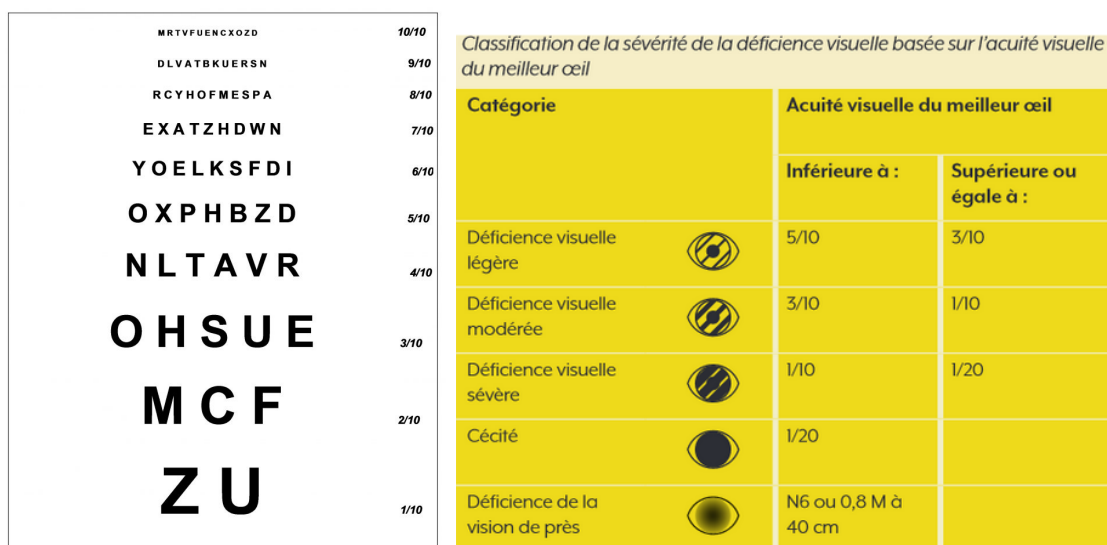


FIGURE A.1 – **Gauche** : Echelle de vision. **Droite** : Tableau de l'OMS résumant la mesure de la déficience visuelle. Illustration issue de [165].

En se basant sur cette mesure, la CIM classe la gravité du déficit visuel en quatre niveaux : 1) déficience légère (acuité visuelle <5/10) ; 2) modérée (<3/10) ; 3) sévère (<1/10) ; 4) cécité (<1/20) [165] (A.1). La déficience visuelle sévère et la cécité sont également définies par un champ visuel réduit : lorsque l'étendue de la zone dans laquelle un objet peut être détecté par

A.1. LA DÉFICIENCE VISUELLE

la vision périphérique alors que l'oeil est focalisé sur un point central est inférieur à 20° et 10° respectivement.

Dans le cadre de cette thèse, nous utilisons le terme "déficience visuelle" tel que défini par la CIM pour parler des niveaux de déficience modérés, sévères, ainsi que de la cécité. Nos travaux portant sur les IHM, nous utilisons tout au long de ce document le terme "personne ayant une déficience visuelle" (abrégé PDV) pour parler de personnes dont la déficience ne permet pas d'effectuer des tâches de nature numérique nécessitant la vision (c'est à dire, interagir avec un écran) sans une aide supplémentaire notable (par exemple, un agrandissement du texte ou une vocalisation).

Le handicap Dans le langage courant, le terme "handicap" est souvent utilisé comme synonyme de "déficience"¹. Il faut cependant distinguer la déficience du handicap. Le handicap fait référence aux limitations et restrictions qu'une personne peut rencontrer dans ses interactions avec son environnement - qu'il soit physique, social ou attitudinal.

Selon la CIF, le handicap est un concept qui englobe non seulement les déficiences physiques ou mentales d'un individu, mais également les limitations d'activité et les restrictions de participation auxquelles il peut être confronté en raison de facteurs environnementaux [164] : barrières physiques, attitudes négatives, politiques discriminatoires, difficultés d'accès aux services, aux transports, à l'éducation, à l'emploi . . .

On comprend donc que la déficience seule ne fait pas le handicap. En 2003, les chercheurs Sears et Young ont introduit en IHM le concept de "déficiences et incapacités induites par la situation" (SIID, pour Situationally-Induced Impairments and Disabilities) [200]. Selon eux, l'environnement et le contexte dans lesquels une personne évolue peuvent causer des difficultés similaires à celles associées à un handicap, même si cette personne ne présente pas initialement de déficience au sens médical. Par exemple, une personne en fauteuil roulant et un parent

1. APA - Committee on Disability Issues in Psychology. (1992). Guidelines for Nonhandicapping Language in APA Journals. Consulté le 22 Mai 2023.

avec une poussette font face à un handicap similaire lorsqu'ils essaient d'accéder à un bâtiment entouré d'escaliers.

Dans notre cas de figure, le niveau de handicap d'une PDV n'est donc pas seulement dû à la déficience visuelle elle-même, l'environnement joue aussi un rôle important : l'accès à des soins, aides techniques et services adaptés, voire l'attitude des individus qu'une personne rencontre définissent en partie son fonctionnement quotidien, et donc son niveau de handicap. Cela implique que le niveau de handicap des PDV peut être réduit : en optimisant leur fonctionnement quotidien et fournissant des ajustements pratiques pour compenser les conséquences de leur déficience [218, 191, 29] - et ce, notamment à l'aide de dispositifs technologiques - il devient possible d'améliorer significativement leur qualité de vie [210, 52, 121, 166].

A.2 Les technologies d'assistance

Tout article, équipement ou système, utilisé pour accroître, maintenir ou améliorer les capacités fonctionnelles des personnes handicapées en renforçant leur indépendance dans l'exécution d'activités de la vie quotidienne peut être considéré comme une technologie d'assistance (TA) [1]. Selon la CIF [163], il s'agit de "tout produit, instrument, équipement ou technologie adapté ou spécialement conçu pour améliorer le fonctionnement d'une personne en situation de handicap". Pour l'organisation internationale de normalisation (ISO)² "tout produit (y compris les dispositifs, équipements, instruments, technologies et logiciels) spécialement produit ou généralement disponible, pour prévenir, compenser, surveiller, soulager ou neutraliser les déficiences, les limitations d'activité et les restrictions de participation" est une technologie d'assistance.

Il existe une large gamme de TA à destination des PDV, comprenant une variété d'outils allant des dispositifs physiques aux solutions logicielles. Certaines TA sont employées par les PDV pouvant encore user de leur vue [104] : elles servent à compenser une baisse d'acuité

2. ISO. ISO 9999 Assistive products for persons with disability – Classification and terminology. Geneva : International Organization for Standardization, 2007.

A.2. LES TECHNOLOGIES D'ASSISTANCE

et/ou une pathologie entraînant des symptômes plus spécifiques (par exemple, problème de perception des couleurs). Ces TA sont principalement des dispositifs d'agrandissement tels que des loupes optiques et des loupes électroniques - qui permettent aussi de modifier la luminosité, les couleurs, le contraste etc.



FIGURE A.2 – **Gauche** : Loupe électronique (aussi appelée “loupe vidéo”). **Droite** : Clavier Perkins permettant d’écrire et de lire en Braille sur ordinateur.

D’autre part, de nombreuses TA utilisent la perception tactile ou auditive plutôt que visuelle pour permettre aux PDV ne pouvant pas recourir à leur vue d’accéder à du contenu et d’interagir avec leur environnement. Ces aides utilisent des systèmes tels que le braille, les dessins et symboles tactiles en relief, ou encore la synthèse vocale [222] pour proposer des moyens d’interactions alternatifs (figure TTA) aux interfaces graphiques et aux périphériques d’entrée/sortie traditionnels (c’est à dire, écran, clavier, souris).

Ces TA jouent un rôle essentiel dans l’accessibilité numérique aux PDV et la demande mondiale de ces aides est très élevée, avec environ 1 milliard de personnes handicapées ayant besoin d’un ou plusieurs dispositifs pour améliorer leur qualité de vie [214]. Bigham notamment, met en évidence le rôle crucial des TA dans l’accès à l’information et à la technologie pour les personnes en situation de handicap [28]. Scherer et Glueckauf ont montrés que les TA facilitent les activités quotidiennes en autonomie telles que la lecture, la cuisine et les déplace-

ments, améliorant par conséquent leur qualité de vie subjective et leur estime de soi [196]. Le cas des PDV recourant à un lecteur d'écran est un exemple concret. Grâce à cette technologie leur permettant d'utiliser un ordinateur ou un smartphone de façon autonome, elles parviennent à acquérir une meilleure intégration sociale, impactant directement sur leur position sociale et économique [107].

Malgré ces avantages, les taux d'abandon et de non-utilisation des TA traditionnels sont préoccupants. Selon certains rapports, le pourcentage d'abandon varie de 30% à 50% pour l'ensemble des dispositifs, et peut atteindre jusqu'à 75% pour certains dispositifs spécifiques [77]. En effet, ces dispositifs ne sont pas utilisables par tous en raison de divers facteurs tels que leur disponibilité, leur coût, leur acceptabilité - voire leur facilité d'usage [2, 172, 165, 11].

C'est pourquoi les appareils grand public tels que les smartphones et les tablettes - dont l'usage croît chaque année [3] - sont de plus en plus utilisés par les PDV [140], car ils sont généralement plus abordables et attirent moins l'attention sur l'utilisateur [97]. Dotés de nombreuses fonctionnalités (notamment, affichage, caméra, audio, GPS, internet), les smartphones modernes peuvent exécuter de nombreuses applications d'assistance pour les PDV, comparables à d'autres dispositifs traditionnels plus spécialisés [56, 167]. Par exemple, BlindSquare, une application de navigation conçue pour iOS (le système d'exploitation mobile exclusif d'Apple) utilisant le GPS du smartphone, coûte 39,99\$, tandis que le Victor Reader Trek, un dispositif GPS parlant dédié, coûte 975\$. Parmi les applications actuellement utilisées par les PDV : ScanLife Barcode et QR Reader permettent d'obtenir des informations sur un produit en scannant le code barre ; AiPoly et AudioLabels permettent de reconnaître des objets et identifier les couleurs ; EyeSight et Ultra Magnifier sont utilisés pour l'agrandissement ; MessagEase Keyboard, AccessNote et BrailleTouch pour l'écriture . . .

Cela permet aux PDV de n'utiliser que leur smartphone plutôt que de devoir acheter et gérer une kyrielle de TA différentes.

A.3 Accessibilité des IHM

Néanmoins, l'utilisation même de ces appareils peut être une cause de handicap pour les PDV lorsque les fonctionnalités proposées ne sont pas suffisamment adaptées [113]. On parle alors de problèmes d'accessibilité numérique.

Selon le Référentiel général d'amélioration de l'accessibilité (RGAA)³ "L'accessibilité numérique consiste à rendre les contenus et services numériques compréhensibles et utilisables par les personnes en situation de handicap".

Or, les PDV sont confrontés à des défis inhérents au fonctionnement de ces appareils grand public, ceux-ci utilisant principalement des écrans tactiles [147], dont l'usage repose essentiellement sur des interfaces visuelles [194]. La prise de conscience des enjeux sociétaux et l'utilisation croissante des smartphones par les PDV poussent les fabricants, concepteurs et développeurs à améliorer l'accessibilité de leurs interfaces et applications [137]. Cela se traduit souvent, dans le cas d'une déficience visuelle, par une altération de l'interface du site web ou de l'application : par exemple un changement de la taille et/ou de la couleur de la police d'écriture, un changement de la couleur de fond ou des images, une réorganisation du contenu (c'est à dire, liens, menus) - mais aussi de s'assurer que ce contenu peut être lu par une plage braille ou une synthèse vocale dans le cas où l'utilisateur ne pourrait pas du tout utiliser sa vue.

Il s'agit là toutefois de problèmes liés au retour d'information en sortie, celle produite par le système à destination de l'utilisateur. Sur les appareils grand public, elles se font essentiellement sous forme de rétroactions visuelles, mais puisque ceux-ci sont aussi dotés de haut-parleurs et de fonctions vibrantes, les informations peuvent être en partie adaptées lorsqu'il s'agit de rendre les fonctionnalités accessibles au PDV : sous forme de retour auditif à travers le son ou la parole ; sous forme de retour haptique par le toucher ou les vibrations.

La déficience visuelle handicape aussi l'usage des smartphones en entrée. Les entrées

3. Direction interministérielle du numérique Référentiel général d'amélioration de l'accessibilité - version 4.1.2

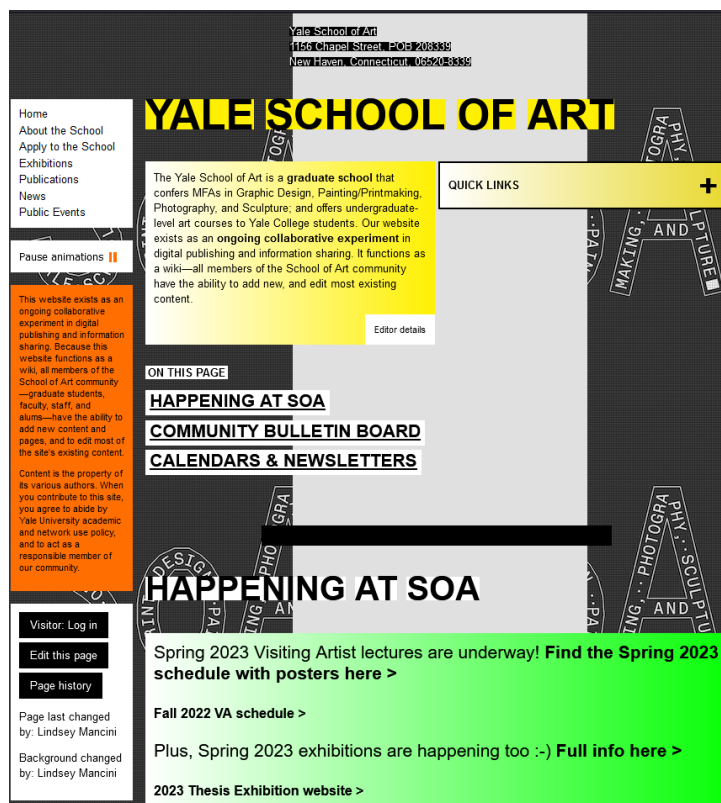


FIGURE A.3 – Exemple de mauvaise accessibilité pour PDV. Capture d'écran issue du site internet de l'université d'Art de Yale.

font référence aux moyens par lesquels les utilisateurs interagissent avec un système. Elles peuvent inclure le toucher, la voix, les gestes, le clavier, la souris ou l'utilisation d'autres dispositifs d'entrée spécialisés. Cependant, les appareils grand public sont dotés d'écrans tactiles qui requièrent principalement de manipuler directement [205, 207] des éléments appartenant à l'interface graphique (appelée GUI, pour Graphical User Interface). Cela nécessite de viser précisément les éléments interactifs qui la composent (par exemple, liens, boutons, icônes ...) ce qui est un défi majeur pour les PDV - particulièrement dans le cas des smartphones et tablettes, dotés d'écrans tactiles de taille modeste. Par ailleurs, la taille et l'aspect mobile de ces appareils restreint l'usage de dispositifs physiques supplémentaires plus accessibles aux PDV, tel qu'un clavier par exemple. Les PDV apprécient le smartphone pour sa multifonctionnalité et sa portabilité, or ajouter des dispositifs supplémentaires serait préjudiciable à la praticité perçue du dispositif.

Annexe A

Annexe II