
Rapport sur la demande de Monsieur Philippe Roussille en vue de l'obtention du Doctorat
de l'Université Toulouse III – Paul Sabatier spécialité Informatique.

Monsieur Philippe Roussille présente une thèse en informatique sous le titre « Techniques d'assistance à la saisie de texte sur périphériques mobiles dans le cas de la déficience visuelle ». Cette thèse a été préparée sous la direction de M. Mathieu Raynal.

Cette thèse s'inscrit dans le domaine de la saisie de texte. Elle vise à étudier le fonctionnement de la liste de prédiction. Les contributions de ces travaux de thèse se situent aussi bien dans le cadre de la modélisation et de l'évaluation des claviers logiciels que dans des cas concrets menant à de nouvelles solutions.

Le manuscrit comporte 145 pages ainsi qu'une bibliographie et des annexes. Il est structuré en six chapitres. Après une introduction qui fixe les objectifs de la thèse et décrit la structure du manuscrit, le premier chapitre présente un état de l'art de la saisie sur dispositifs mobiles. Les deuxième et troisième chapitres présentent la proposition, le clavier déductif DUCK ainsi que la première étude expérimentale. Les quatrième et cinquième chapitres s'attaquent à deux problématiques particulières à savoir l'interaction avec la liste de prédiction et la gestion des mots courts. Le sixième chapitre présente un usage journalier du clavier DUCK en intégrant le mode mots courts et l'interaction avec la liste de prédiction choisie. Enfin, le manuscrit se termine par une conclusion et des perspectives sur le travail mené. La structure globale du manuscrit en six chapitres apparaît claire même si le nombre de chapitres fait apparaître un déséquilibre de volume. J'aurais regroupé les chapitres afin de mieux les équilibrer.

Le premier chapitre dresse un état de l'art de la saisie de texte sur dispositifs mobiles. Pour cela l'étude est axée sur deux points, la méthodologie d'évaluation et les différents systèmes de saisie de texte.

Le premier point est présenté à travers le degré d'utilisabilité : efficacité, efficience et satisfaction. Il en découle les aspects et problèmes suivants à considérer :

- La performance ultime,
- La performance initiale et temps d'apprentissage,
- Les erreurs en tant que telles et la gestion de ces erreurs,
- L'attention visuelle et cognitive,
- L'appréciabilité.

Afin de pouvoir prendre en compte ces aspects, différentes mesures sont décrites dans la suite du chapitre. La première partie concerne la vitesse (caractère, mot) et l'erreur (omission, addition, substitution, transposition). Cette partie est pédagogique à travers différents exemples pratiques permettant de bien comprendre la technique et l'objectif de mesure. La présentation se poursuit avec un aspect important et souvent mal pris en compte, la prédiction qui est souvent couplée au système de saisie. La mesure peut se faire principalement selon le taux de succès, les frappes économisées et le temps économisé. Enfin, la partie sur les mesures se termine par la mesure de la satisfaction utilisateur à travers l'expérience utilisateur, l'utilisabilité et la mesure de la charge de la tâche. Là encore la rédaction est claire. J'aurais toutefois aimé voir apparaître la double tâche dans la mesure de la charge de la tâche et en particulier son usage dans le cadre des tâches sans regard.

Le second point contient un état de l'art des systèmes de saisie de texte sur dispositifs mobiles. En préambule j'aurais aimé disposer de plusieurs définitions :

- du mot mobile (utilisateur mobile, utilisateur en situation de mobilité, dispositif transportable, etc.) car cela induit différentes difficultés et donc des contextes d'évaluation différents.
- de la déficience visuelle considérée. Comment un conducteur ne pouvant regarder son écran est-il considéré ? Là encore cela induit des choix différents.

Les principales techniques présentes dans la littérature sont présentées : reconnaissance vocale, clavier logiciel, clavier gestuel, etc. Le lien avec la déficience visuelle n'est toutefois pas toujours facile à faire. Une classification à partir de critères spécifiques à la déficience visuelle et à la mobilité serait un plus pour cette compréhension.

Le deuxième chapitre présente la technique proposée, le clavier DUCK. La proposition est justifiée par une présentation des difficultés à utiliser un clavier logiciel oralisé que ce soit pour la saisie d'un caractère ou la gestion des erreurs. Le clavier DUCK se propose donc :

- de conserver la disposition classique AZERTY afin de faciliter l'apprentissage ;
- de s'appuyer sur une frappe précise du premier caractère du mot de manière similaire aux claviers logiciels oralisés classiques ;
- de permettre des frappes approximatives avec un retour vocal principal au niveau du mot pour les caractères suivants.

La technique proposée est décrite comme déjà utilisée dans la littérature par SHARK. Pourtant, en ne prenant pas en compte le tracé de l'utilisateur mais uniquement des frappes disjointes, cette proximité est discutable. Il conviendrait de clarifier ce point. Lorsque la saisie du mot est terminée, une double frappe active le système de déduction. Si aucun mot n'existe, un retour sonore informe l'utilisateur. Si un seul mot est proposé, il est inséré. Dans le cas contraire une liste de mots remplace le clavier et l'utilisateur doit choisir son mot. A ce stade, le manuscrit n'indique pas la technique utilisée pour choisir le mot. L'algorithme de déduction est ensuite décrit. Il s'appuie sur une présélection des mots commençant par la première lettre saisie et de même nombre de caractères. Puis un calcul de distance moyenne (distance touche contenant la lettre / position de frappe) est appliqué afin de trouver les quatre mots les plus proches et donc, a priori, les plus probables. Il est posé comme hypothèses que la saisie approximative doit permettre une augmentation de la vitesse de saisie et que le système de déduction doit faire baisser le nombre d'erreurs.

Le troisième chapitre peut être considéré comme une étude préliminaire de prise en main du clavier afin de vérifier la pertinence de la proposition. L'étude très courte comportait une seule session et douze participants. Le clavier DUCK était comparé au clavier VODKA, un clavier logiciel oralisé implémenté pour l'expérimentation. Ce dernier ne disposait pas de système de correction, de prédiction ou de déduction. Les deux claviers partageaient le même agencement AZERTY sans ponctuation et sans accent. Chaque participant a utilisé les deux claviers et a écrit 8 phrases courtes avec chaque clavier. Il n'est pas indiqué si la fréquence des lettres respecte la fréquence de la langue. Les résultats montrent que DUCK est plus rapide lorsque les mots font plus de six caractères. Mais au niveau des caractères, le taux d'erreur de DUCK est supérieur à VODKA. La validation de mots incorrects explique ce point. Le score SUS de DUCK est meilleur que VODKA mais la préférence des utilisateurs se porte majoritairement sur le clavier VODKA. A la suite de ce test préliminaire, le travail se focalise sur l'utilisation de la liste de déduction afin de minimiser le temps de validation et sur la saisie des mots courts afin d'améliorer les performances du clavier DUCK. Il est dommage que d'autres points ne soient pas abordés comme l'inexistence du mot dans la liste de déduction ou la saisie d'un mot incorrect.

Le quatrième chapitre explore l'interaction avec la liste de prédiction. Il s'agit d'améliorer la vitesse de sélection du mot dans la liste de prédiction et de réduire les erreurs de sélection du mot. La sélection d'un mot se décompose en trois phases : l'exploration, la sélection et la validation. Une première étude porte sur la présentation des mots à travers différentes dispositions : absolue (ligne, colonne et grille), relative (anneau centré sur le doigt) ou par page (un mot par page). Une évaluation a été menée avec 12 participants mais dont la vision était normale ou corrigée (yeux bandés). Il en ressort que la disposition absolue semble la plus appropriée. Une deuxième étude porte sur le retour vocal proposé à l'utilisateur. Deux retours sont évalués : « descriptif » et « prescriptif ». Une évaluation a été menée avec 12 participants mais dont la vision était normale ou corrigée (yeux bandés). Il en ressort que le retour audio « prescriptif » permet d'explorer moins d'items mais que le retour audio « descriptif » en ligne donne les meilleurs résultats. Pourtant, en cas d'absence du mot de la liste, les participants parcourent tout de même 3,28 items en moyenne. Pourquoi ? De même la vitesse de la synthèse vocale n'a pas été intégrée dans les tests, c'est dommage car il est courant d'utiliser des vitesses très élevées et cela aurait pu modifier les résultats sur le retour audio « prescriptif ». Cela aurait pu être intéressant. La dernière étude porte sur la technique de validation afin d'éviter les erreurs lors du relâchement. Quatre techniques sont étudiées : le « relâchement », le « dual-tap », le « double-tap » et la « ligne brisée ». Là encore, une évaluation a été menée avec 12 participants mais dont la vision était normale ou corrigée (yeux bandés). Les meilleurs résultats ont été obtenus avec le « relâchement » et le « double-tap ». Pour la technique du « relâchement » il n'est pas indiqué si un délai est utilisé afin de ne pas valider un élément si son activation s'est faite dans un délai très court avant le relâchement. De même les résultats auraient-ils été les mêmes avec des utilisateurs plus habitués à l'usage de la synthèse vocale ? En synthèse, il est indiqué que la présentation sous forme de liste sera utilisée avec un retour descriptif et une validation par « relâchement ».

Le cinquième chapitre explore la gestion des mots courts. Dans la langue française, 47% des mots usuels sont formés d'au plus 4 lettres. Il est proposé d'utiliser des listes statiques de 8 mots courts, une liste étant associée au premier caractère des mots courts. Il y a donc 22 listes car quatre caractères ne sont jamais en tête de mots courts. Les mots sont choisis en se basant sur la fréquence des mots. L'ordre de ces mots est toujours le même mais la manière qui a servi à ordonner ces mots n'est pas indiquée. L'accès à la liste se fait après la saisie du premier caractère. Pour différencier l'accès à la liste de la saisie d'un mot, deux techniques sont étudiées : le « double-press » et le « dual-tap ». Les techniques apparaissent compliquées car elles demandent des appuis-relâchement, pourquoi ne pas avoir simplement distingué l'accès à la liste grâce à l'appui de deux doigts et maintenir ces deux doigts pour le choix du mot dans la liste ? Une évaluation a été menée avec 12 participants mais dont la vision était normale ou corrigée (yeux bandés). Les résultats ne montrent pas de différence significative entre les deux techniques. Par contre le taux d'erreur semble élevé, de l'ordre de 12%. Ceci est dû à la sélection du mot voisin lors du relâchement. Ceci est surprenant dans la mesure où dans le chapitre précédent ce type d'erreur faisait l'objet d'une étude. Il n'est pas clairement indiqué que le mode « double-press » est choisi.

Le sixième chapitre contient une expérimentation sur la découverte et l'apprentissage de DUCK pour un usage journalier. Le clavier reprend les résultats des chapitres précédents avec en plus un mode « glisser-taper » permettant d'écrire un mot absent du dictionnaire. Une évaluation a été menée avec 4 participants non-voyants. Deux sessions de 20 minutes (ou au plus 20 phrases) ont été menées. Entre ces deux sessions, les utilisateurs pouvaient venir s'entraîner selon leur désir. Les résultats montrent une progression des résultats entre les deux

sessions, indiquant un apprentissage. Dès la première session, DUCK se révèle plus rapide que VODKA. Le taux d'erreur présenté est calculé sur les mots et non les caractères. DUCK étant basé sur un dictionnaire cela induit des résultats différents que pour VODKA, ces résultats sont difficiles à évaluer. Le taux d'usage des mots courts augmente entre les deux sessions. Concernant l'évaluation, il manque quelques informations. Par exemple le nombre de mots absents du dictionnaire, le nombre de caractères tapés, l'effort des utilisateurs, etc. Cette partie pourrait être approfondie afin de mieux comprendre les résultats.

Le manuscrit se termine par une conclusion dressant le bilan du travail accompli et des contributions de cette thèse. Les ouvertures proposées sont à plusieurs niveaux. La première concerne l'amélioration de la déduction et d'introduire la complétion. La deuxième amélioration possible concerne la correction qui se fait au niveau mot et non au niveau caractère. La troisième aborde la problématique de la ponctuation.

Le travail présenté est très intéressant car il aborde un problème difficile, la saisie de texte à destination des personnes déficientes visuelles. Partant d'une solution existante, le clavier logiciel oralisé, ce travail apporte des solutions afin d'améliorer les performances ou l'expérience utilisateur. Le manuscrit est réparti sur six chapitres mais mériterait d'être réduit à quatre chapitres. Cela permettrait d'harmoniser les contenus. De plus, il conviendrait de clarifier les participants impliqués dans les différentes évaluations, étaient-ils différents ? De même, certaines évaluations menées avec des voyants occultés peuvent amener à discuter les résultats. En effet, l'usage de la synthèse vocale n'est pas chose courante et des utilisateurs non-voyants auraient pu obtenir des résultats différents. Enfin, beaucoup de termes apparaissent au fur et à mesure du manuscrit, ceci ne facilite pas la lecture.

En résumé, les travaux de thèse de Philippe Roussille traduisent une approche conceptuelle à travers les concepts présentés et une approche pratique à travers les différentes solutions proposées. Un nombre considérable d'études expérimentales ont été menées. Certaines solutions méritent toutefois des précisions, des approfondissements et des expérimentations complémentaires.

Je donne un avis favorable à la soutenance de ce mémoire en tant que thèse d'informatique de l'Université Toulouse III – Paul Sabatier.

Fait à Metz, le 30 décembre 2016

Benoît Martin

