

Etude empirique de l'usage de la multimodalité avec un ordinateur de poche

*Guillaume Calvet,
Julien Kahn,
Moustapha Zouinar,
Pascal Salembier,
Jean-Christophe Briois*

GRIC-IRIT
Université Paul Sabatier,
118, rte de Narbonne,
31062 Toulouse Cedex 7
calvet@irit.fr

*Laurence Nigay,
Gaëtan Rey*

CLIPS-IMAG
Université Joseph Fourier
B.P. 53
38041 Grenoble cedex 9
Laurence.Nigay@imag.fr
Gaetan.Rey@imag.fr

Laurence Pasqualetti

FT R&D-DIH/UCE
38-40 rue Gal Leclerc
92794 Issy-les-Moulineaux
laurence.pasqualetti@
francetelecom.fr

RESUME

Cet article présente une étude empirique dont l'objet est l'usage indépendant ou corrélé de différentes modalités d'interaction implémentées ou simulées sur un ordinateur de poche. L'étude concerne les modalités de sortie, comme la visualisation ou la parole ainsi que les modalités d'entrée comme le geste, la parole ou encore des modalités basées sur la manipulation "naturelle" de l'or-

dinateur lui-même ("embedded physicality" [4] ou "embodied user interface" [3]).

MOTS CLES : Multimodalité, Magicien d'Oz, Ordinateur de Poche.

NOM ET VERSION DU LOGICIEL UTILISE POUR ECRIRE LA SOUMISSION :

Word 98 MacIntosh

FORME DE PARTICIPATION :

x	Cocher la forme de participation
	Papier long
X	Poster/papier court
	Rencontres doctorales
	Ateliers
	Cours
	Tables rondes/réunions débat
	Expérience interactive
	Vidéos
	Panorama des laboratoires et organismes
	Journée « Entreprises »

THEMES DE LA SOUMISSION :

x	Cocher les thèmes choisis (vous pouvez en cocher plusieurs)
	- Fondements théoriques des IHM
	- Modèles de l'utilisateur, modèles cognitifs
	- Analyses et modèles de tâches
	- Formalismes et IHM
	- Modèles de développement des systèmes interactifs
	- Modèles distribués de coopération homme-machine
	- Modélisation des organisations humaines
	- Dispositifs, techniques d'interaction

	- Morphologie de l'interaction
X	- Nouvelles techniques d'interaction et de présentation
	- Nouvelles interfaces pour la supervision
	- Interfaces haptiques, interaction au stylo
X	- Interaction multimodale, gestuelle ou en 3 dimensions
	- Intégration de la vision dans l'interface
	- Interaction vocale
	- Interfaces écologiques
	- Coopération entre modalités/média
	- Analyse, conception et évaluation
	- Architecture des systèmes interactifs
	- Processus de conception et ergonomie
	- Méthodes de conception et d'évaluation
	- Evaluation ergonomique et validation
	- Méthodes formelles dans la conception ou l'évaluation
	- Assistance à l'utilisateur
	- Ergonomie du web
	- Systèmes, outils
	- Outils de prototypage
	- Outils de construction
	- Environnements de développement
	- Programmation visuelle
	- Programmation par démonstration
	- Communication médiatisée par l'ordinateur
	- Interaction asynchrone : téléenseignement, télé médecine, etc.
	- Interaction synchrone
	- Conception coopérative
	- Édition coopérative
	- Interaction et coopération
	- Visualisation et représentation de l'information
	- Visualisation de grands corpus de données
	- Recherche dans des documents complexes
	- Aide interactive
	- Stratégies multi-écrans
	- Interaction dans les mondes 3D
	- Interaction et informatique graphique
	- "Visualisation" pour les handicapés
	- Nouveaux champs d'application
	- Réalité virtuelle
	- Réalité augmentée
	- Collecticiels
	- Interfaces intelligentes
	- Interfaces pour les systèmes critiques
	- Informatique mobile et informatique disséminée
	- Multimédia, hypermédia, et interaction sur le WEB
	- IHM et publics adaptés
	- Interaction pour les personnes âgées et handicapées
	- Interfaces adaptées aux besoins du grand public
	- IHM et Société
	- Applications et expériences
	- Commerce électronique
	- Jeux et loisirs
	- Santé
	- Banque
	- Transport
	- Bâtiment

	- Enseignement de l'IHM
	- Autre :

Etude empirique de l'usage de la multimodalité avec un ordinateur de poche

*Guillaume Calvet,
Julien Kahn,
Moustapha Zouinar,
Pascal Salembier,
Jean-Christophe Briois*

GRIC-IRIT
Universite Paul Sabatier,
118, rte de Narbonne,
31062 Toulouse Cedex7
calvet@irit.fr

*Laurence Nigay,
Gaëtan Rey*

CLIPS-IMAG
Université Joseph Fourier
B.P. 53
38041 Grenoble cedex 9
Laurence.Nigay@imag.fr
Gaetan.Rey@imag.fr

Laurence Pasqualetti

FT R&D-DIH/UCE
38-40 rue Gal Leclerc
92794 Issy-les-Moulineaux
laurence.pasqualetti@
francetelecom.fr

RESUME

Cet article présente une étude empirique dont l'objet est l'usage indépendant ou corrélé de différentes modalités d'interaction implémentées ou simulées sur un ordinateur de poche. L'étude concerne les modalités de sortie, comme la visualisation ou la parole ainsi que les modalités d'entrée comme le geste, la parole ou encore des modalités basées sur la manipulation "naturelle" de l'ordinateur lui-même ("embedded physicality" [4] ou "embodied user interface" [3]).

MOTS CLES : Multimodalité, Magicien d'Oz, Ordinateur de Poche.

INTRODUCTION

Les progrès réalisés au cours de cette dernière décennie en matière de technologies informatiques portables (mémoire, rapidité de traitement, capacité à traiter des interfaces hétérogènes, réseau sans fil, etc.) ont ouvert de nouvelles perspectives dans le domaine de l'interaction homme-machine, et, permettent d'envisager enfin la conception de systèmes capables de traiter et de « comprendre » des actions ou expressions multimodales (combinaison de différentes modalités pour réaliser une même intention ou réalisation simultanée d'intentions différentes à l'aide de modalités différentes). Les intérêts attendus sont de plusieurs ordres : rendre l'interaction homme-machine plus efficace, plus robuste et plus flexible [8]. Cet aspect flexibilité est particulièrement important : la liberté laissée à l'utilisateur de choisir une modalité ou une combinaison de modalités la plus efficace et la plus adaptée à un contexte d'utilisation évolutif (propriété inhérente aux situations de mobilité) constitue une condition déterminante de l'acceptation et de l'utilisation effective du système. Un autre avantage attendu de l'utilisation de ces interfaces est qu'ils facilitent la récupération d'erreurs en permettant par exemple aux utilisateurs de changer de modalité lorsque l'usage d'une

modalité donnée génère des incompréhensions entre l'utilisateur et le système. Par ailleurs, on s'attend à ce que ces systèmes soient plus efficaces dans l'interprétation des actions de l'utilisateur grâce notamment à la « désambiguïsation mutuelle » [9] qui se réfère à la capacité du système à intégrer et confronter les informations provenant des différents modules d'interprétation des données issues des différentes sources d'entrée (modalités).

Par ailleurs le recours possible à de nouvelles modalités d'interaction (« embodied user interface » par exemple) dans le contexte de systèmes portables nécessite d'être plus précisément défini et leur apport évalué.

OBJECTIFS DE L'ETUDE

A la suite d'un certain nombre de travaux ([2], [5], [6], [7] etc.) nous considérerons une modalité comme un moyen de communication qui met en œuvre un dispositif physique et un langage d'interaction. Nous envisagerons la multimodalité du point de vue de l'utilisateur et elle sera considérée comme un moyen de production de ses intentions au travers de différentes modalités (regard, geste, parole, manipulation des artefacts, etc.). Dans cette perspective, nous considérons comme multimodaux les systèmes qui permettent des usages spécialisés (une modalité exclusivement dédiée à une commande), équivalents (une modalité peut être utilisée pour toutes les commandes), complémentaire (plusieurs modalités peuvent être combinées pour réaliser une même commande).

L'objectif de cette étude est d'explorer un certain nombre de questions qui ont été peu ou pas abordées jusqu'à présent : qu'est-ce qui guide le choix de l'usage d'une modalité ? Qu'est-ce qui guide le choix de réaliser une intention/une commande unimodalement ou multimodalement ? Quel est le rôle des propriétés des modalités et du contexte (caractéristiques de l'interface/système, de

l'utilisateur, de l'environnement physique, de la tâche, etc.) dans ces choix ? Comment se caractérise l'usage de la multimodalité dans différents contextes ? Existe-t-il des contextes plus adaptés à l'usage de la multimodalité que d'autres ? En outre, les travaux précédents n'ont abordé qu'un ensemble limité de modalités en entrée (principalement la parole, le pointage ou l'écriture). Enfin les tâches proposées aux sujets étaient le plus souvent « nouvelles ». Une limite à ce choix est qu'il laisse de côté l'étude des conséquences de la multimodalité dans la réalisation de tâches familières avec des systèmes plus classiques.

Dans cet article, nous présentons une étude dont l'objectif est justement de traiter ces questions et d'investiguer l'usage de différentes modalités en situation de réalisation d'une tâche familière, avec un système multimodal partiellement simulé. On cherche ainsi à contribuer à la définition de recommandations ou principes de conception des systèmes multimodaux sur ordinateur de poche.

METHODE

Sujets

15 sujets rémunérés ont participé à cette étude, tous disposaient d'une adresse électronique et étaient des utilisateurs réguliers d'Internet, mais aucun n'était informaticien.

Tâche

Afin de mettre les sujets en situation de réalisation d'une tâche familière il leur était demandé d'utiliser une application de consultation de courrier électronique au moyen d'un ordinateur de poche (PDA Jornada 540 series , Hewlett Packard) connecté à leur messagerie personnelle. L'application de consultation a été développée en JAVA 1.2. L'interface utilisateur reproduit grossièrement les fonctions de consultation des applications classiques de courrier électronique comme le présente la Figure 1.

Quatre modalités d'interaction avec le système ont été retenues : tactile (stylet), verbale, gestuelle et « embodied ». La modalité embodied consiste à associer des commandes à certaines manipulations de l'artefact (changements d'orientation, mouvements divers, etc.). Les Figures 2 et 3 donnent deux exemples de spécification de la commande de défilement vers le bas de la liste des courriers, respectivement en manipulant l'ordinateur et en faisant un geste reconnu en trois dimensions par la caméra.

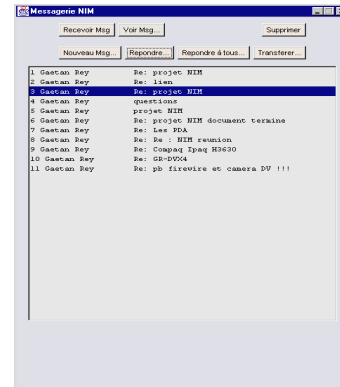


Figure 1 : interface utilisateur



Figure 2 : Exemple de commande embodied



Figure 3 : Exemple de commande gestuelle

Nous nous sommes attachés à étudier sept commandes de consultation et de navigation : recevoir, sélectionner, ouvrir, lire, fermer et supprimer un message ; ainsi que le défilement avant et arrière au sein d'une fenêtre ou d'un message. Pour chaque commande l'utilisateur avait le choix des quatre modalités décrites ci-dessus. L'équivalence entre les quatre modalités pour les sept commandes étudiées est donc totale [7]. Si les commandes tactiles étaient exécutées automatiquement par l'application, pour les autres modalités nous avons utilisé la méthode du magicien d'Oz. La mise en œuvre de cette méthodologie était motivée par le souhait d'utiliser un large éventail de modalités et de combinaisons de modalités en minimisant les contraintes et les coûts de développement. Les retours du compère étaient graphiques

(sélection, ouverture, défilement, message d'erreur, etc.) ou verbaux (lecture du contenu du courrier électronique ou des messages d'erreur).

Plate-forme expérimentale

Le sujet et le compère étaient dans des pièces séparées. Chaque pièce était équipée d'une caméra numérique afin d'enregistrer les comportements et les verbalisations des sujets et du compère durant l'interaction. De plus, la caméra située dans la pièce des sujets permettait au compère d'observer et d'entendre au moyen d'un moniteur, les commandes gestuelles, embodied ou verbales du sujet [1]. La Figure 4 schématise la plate-forme expérimentale.

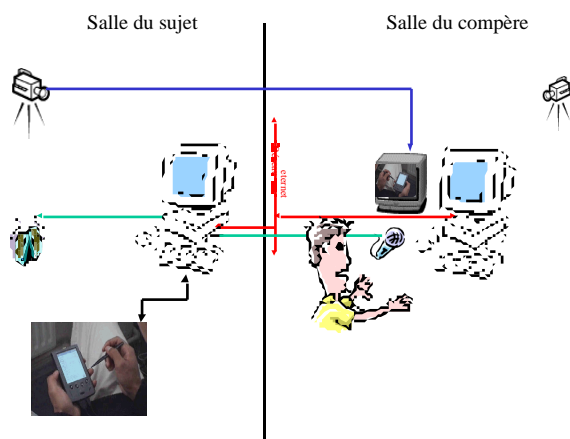


Figure 4 : Plate-forme expérimentale

Procédure expérimentale

Les sujets utilisaient le système durant 7 sessions d'interaction espacées entre elles de 4 à 10 jours. Les sessions d'interaction n'excédaient pas 15 mn. Immédiatement après l'interaction, une autoconfrontation était menée par l'expérimentateur. Rappelons sommairement, que l'autoconfrontation consiste à provoquer les verbalisations du sujet alors qu'il visionne l'enregistrement vidéo de son interaction. Dans notre étude, il était demandé au sujet de commenter divers aspects de son interaction tels que le choix des modalités, les problèmes rencontrés, etc. Cette méthode permet d'appréhender les attentes, les intentions et les stratégies des sujets.

Notons que la première session était précédée d'un apprentissage constitué d'une présentation du système et des modalités, puis d'une série d'exercices visant à faire utiliser par le sujet chaque commande au moyen de chacune de ces modalités. Lors des sessions suivantes, l'expérimentateur s'assurait de la bonne connaissance de l'usage des modalités par le sujet.

Données recueillies

Différentes données ont été recueillies au cours de l'étude. Durant l'interaction, les commandes et les fenêtres utilisées par les sujets étaient indexées temporellement et archivées automatiquement. Les actions des sujets, les modalités utilisées, la durée et le contenu des commandes étaient complétés à posteriori à partir du film vidéo de l'interaction. Enfin, l'enregistrement des autoconfrontations a été réalisé afin de permettre des analyses qualitatives à partir de l'identification des attentes, des intentions et des stratégies des sujets quant à l'usage des modalités.

Premiers résultats

Les premiers résultats obtenus mettent en évidence un certain nombre de tendances :

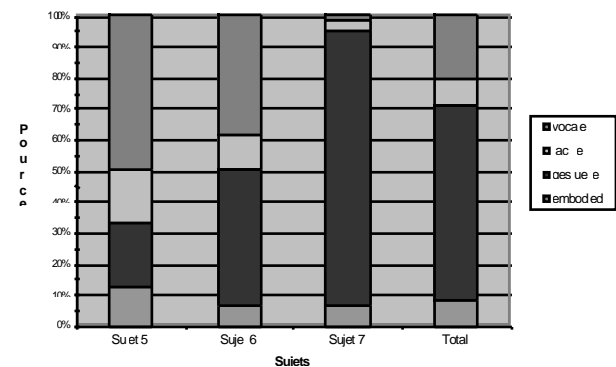


Figure 5 : Usage des modalités toutes sessions confondues pour 3 sujets

Dans la figure ci-dessus on remarque que l'ensemble des modalités a été utilisé par les sujets au cours des différentes sessions ; ce résultat pour trivial qu'il puisse paraître n'allait pas forcément de soi : on se reportera aux nombreux exemples de non-utilisation des propriétés multimodales des systèmes interactifs [8].

Par ailleurs, on note une variabilité inter-individuelle dans l'usage de la multimodalité par les sujets (cf. Figure 5 : Usage des modalités toutes sessions confondues pour 3 sujets) ; cette tendance déjà observée dans des études précédentes [5] met en évidence l'existence de stratégies différentes d'usage de la multimodalité selon les utilisateurs.

De plus, comme le montre la Figure 6, cette variabilité inter-individuelle s'accompagne d'une variabilité intra-individuelle. Ce résultat a déjà été observé dans d'autres études [5] par exemple.

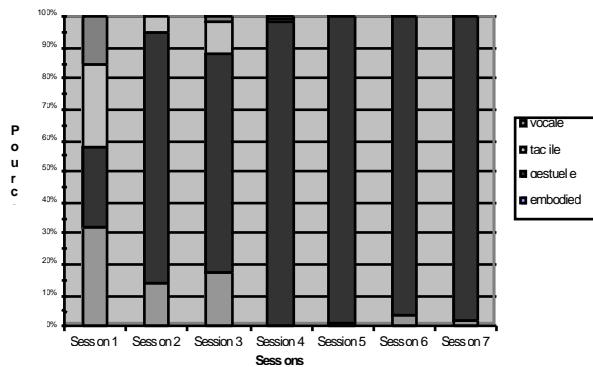


Figure 6 : Usage des modalités pour un sujet au cours des sessions

Notons que ces premiers résultats ont été obtenus sur la base d'une analyse de nature quantitative. Si elle permet de dégager sous une forme synthétique un certain nombre de tendances dans l'usage du dispositif multimodal, elles ne permettent pas de répondre de manière informative aux questions évoqués plus haut. Elles ne permettent pas non plus de distinguer les effets « évaluation de l'intérêt fonctionnel » des modalités des aspects « évaluation de l'implémentation » de ces modalités. Il s'avère donc nécessaire de passer d'un niveau d'analyse descriptif factuel à un niveau explicatif qui permette d'interpréter certains aspects clés de l'utilisation de la multimodalité par les sujets ; ceci passe par une analyse qui aille au-delà du simple relevé des enchainements de commandes sur le système comme traces de l'activité des sujets, par la prise en compte des déterminants psychologiques (intentions, attentes, etc.) de celle-ci. D'où l'intérêt de recourir à des techniques d'identification de la signification des actions des sujets basées sur l'autoconfrontation.

CONCLUSION

Nous avons présenté les premiers résultats d'une étude empirique de l'usage de la multimodalité sur un ordinateur de poche. L'analyse de nature quantitative des données recueillies n'est pas finie. La technique du magicien d'Oz, la nature de la tâche et le nombre de sessions ont permis de mettre les sujets en situation réaliste. La plateforme magicien d'Oz développée et l'approche adoptée nous semblent très prometteuses pour explorer des nouvelles formes de modalités comme les modalités basées sur la manipulation de l'ordinateur lui-même ("embodied user interface"). En effet, ces modalités reposent sur des gestes "naturels" or chacune nécessite un effort conséquent de programmation et d'intégration des capteurs requis. Or le développement et la mise au point logiciels sortent du cadre d'une boîte à outils classique. L'intérêt de notre étude réside donc dans l'obtention de données expérimentales sur l'usage sans avoir à réaliser les phases particulièrement coûteuses de développement logiciel. Enfin nous soulignons l'importance de compléter l'ana-

lyse quantitative classiquement mise en œuvre jusqu'à présent par la prise en compte du niveau intentionnel grâce à l'auto confrontation des sujets. Ceci constitue nos travaux dans un futur immédiat après l'achèvement de l'analyse de nature quantitative.

REMERCIEMENTS

Cette recherche a été réalisée dans le cadre d'un contrat de recherche FT-R&D/IMAG/IRIT (Projet HOURIA lot 1).

BIBLIOGRAPHIE

1. Calvet, G., Kahn, J. , Zouinar, M. & Briois, J.C. *Rapport pré-expérimentation HOURIA 1*, Contrat de recherche FT-R&D/IMAG/IRIT, 2000.
2. DeAngeli, A., Wolff, F., Lopez, P. & Romary, R., Relevance and perceptual constraints in multimodal referring actions. In *Proceedings of the Workshop on Deixis, Demonstration and Deictic Belief, Eleventh European Summer School in Logic, Language and Information" (ESSLI-99)*, August 9-20, Utrecht, The Netherlands, 1999.
3. Fishkin, K. P., Moran, T. P. & Harrison, B. L. Embodied User Interfaces: Towards Invisible User Interfaces, In *Proceedings of EHCI'98*, Kluwer Academic Publishers, 1998, pp 1-16.
4. Harrison, B.L., Fishkin, K.P., Want, R. Gujar, A., & Mochon, C.. "Squeeze Me, Hold Me, Tilt Me! An Exploration of Manipulative User Interfaces". In *Proceedings of SIGCHI 98*, 1998, pp. 17-24.
5. Mignot, C & Carbonell, N. Commande orale et gestuelle : étude empirique. *Technique et Science Informatiques* 15, 1996, pp. 1399-1428.
6. Nigay, L.. Thèse de doctorat, Université Joseph Fourier, Grenoble-1, 1994.
7. Nigay, L. & Coutaz J. Espaces conceptuels pour l'interaction multimédia et multimodale, *Technique et Science Informatiques*, spécial Multimédia et Collecticiel, AFCET & Hermes Publ., Vol 15(9), 1996, pp. 1195-1225.
8. Oviatt, S.L. Ten myths of multimodal interactions, *Communications of the ACM*, Vol. 42, N°11, November, 1999, pp. 74-81.
9. Oviatt, S.L. Mutual disambiguation of recognition errors in a multimodal architecture. In *Proceedings of Conference on Human Factors in Computing Systems CHI'99*, ACM Press : New-York, N.Y., 1999: pp. 576-583.