

# Interaction Multimodale sur des artéfacts mobiles

M. Zouinar, P. Salembier, L. Pasqualetti, L. Nigay, , G. Calvet, J. Kahn, G. Rey

France Telecom R&D, Université de Grenoble, Université de Toulouse

## Introduction

Les récents progrès accomplis dans la miniaturisation des microprocesseurs et dans les réseaux informatiques sans fil permettent d'envisager que la «boîte grise» du ordinateur personnel est «condamnée à disparaître» ou du moins à ne pas être le seul lieu d'interaction entre l'utilisateur et le monde numérique. Ceci s'inscrit dans un double mouvement issu des travaux technologiques sur le concept d'ordinateur ubiquitaire et évanescent (*Ubiquitous Computing* et *Disappearing Computer*) et de l'évolution des idées dans le domaine des modèles d'interaction. En effet on s'oriente à présent progressivement vers des modèles d'interaction dans lesquels les ressources informatiques sont distribuées dans une multitude d'objets usuels avec lesquels l'utilisateur va interagir de manière explicite ou implicite. L'idée est ici d'utiliser l'environnement comme interface, comme système de manipulation de ressources techniques fonctionnellement limitées mais contextuellement pertinentes (notion d'interface "*tangible*"). Le dispositif peut être physiquement manipulé de façon signifiante : une action sur le dispositif renvoie à une fonction prédéfinie ( ce qui nécessite la définition d'une sémantique et éventuellement d'une syntaxe de l'interaction avec le dispositif).

La notion d'*objet communicant* recouvre des réalités technologiques et conceptuelles très diverses ; parmi les propriétés spécifiques attachées généralement à ces objets on retiendra notamment :

- l'augmentation numérique plutôt que la substitution : l'idée directrice est de repartir d'objets usuels dont on s'efforce de préserver les avantages intrinsèques inhérents à leur constitutivité matérielle (notamment les propriétés d'*affordance* et de supports à l'*awareness*) tout en leur adjoignant des fonctionnalités supplémentaires (ex : le papier et le stylo numérique, le frigidaire communicant,...).
- le caractère transportable des dispositifs : il permet de bénéficier de ressources "embarquées" de traitement et de communication et favorisent ainsi la mobilité et le nomadisme intellectuel. Ces dispositifs peuvent prendre plusieurs formes (PDA, téléphone portable, vêtement communicant,...).
- la capacité à communiquer de façon autonome ou contrôlée : outre leur fonction classique de support à la communication entre utilisateurs, les dispositifs portables peuvent détecter la présence d'un dispositif de même type (ou de ressources informatiques distribuées dans l'environnement) et échanger des informations selon des règles éventuellement pré-établies mais indépendamment de toute commande donnée intentionnellement par l'utilisateur.

Dans le projet HOURIA nous étudions la manière dont certaines réalisations physiques de cette notion d'objet communicant vont pouvoir être utilisées facilement et de façon efficace par des individus mobiles, et s'intégrer sans provoquer de ruptures dans leurs activités quotidiennes et les environnements physiques et sociaux dans lesquels ils se situent à chaque instant. Nous justifions notre choix de la multimodalité par les deux constats suivants :

- 1) Ces artéfacts reposent sur des dispositifs physiques aux capacités restreintes, sortant du cadre classique (un grand écran, un clavier et une souris). Il convient donc de

concevoir des modalités d'interaction reposant sur des paradigmes autres que la manipulation directe "écran-souris" comme les interfaces tangibles ou encore les modalités incarnées "embodied user interface".

2) Le contexte d'utilisation de ces artefacts est par définition très variable. En effet les caractéristiques physiques (bruit,...) et sociales (intimité, intrusivité,...) de l'environnement déterminent un ensemble de contraintes contextuelles qui vont nécessiter des adaptations interactionnelles ; de ce point de vue la multimodalité constitue un type de réponse adapté à cette exigence d'adaptabilité et d'adaptativité. Le développement de l'informatique mobile représente donc un domaine privilégié pour l'application des techniques d'interaction multimodale.

### **Problématique et objectifs de l'étude**

La multimodalité a donné lieu à de nombreux travaux théoriques et empiriques. Les travaux théoriques ont principalement porté sur la définition des notions de modalité et de multimodalité, et sur le développement « d'espaces de conception ». Ces espaces sont des cadres conceptuels qui fournissent un ensemble de notions permettant de décrire les modalités et les possibilités de combinaison de celle-ci au regard de l'interaction utilisateur-système. Par exemple, les modèles TYCOON (Martin & Béroule, 1993) et CARE (Coutaz & Nigay, 1996) proposent un ensemble de concepts qui décrivent différents types de relation théorique de composition ou « coopération » entre modalités : assignation, complémentarité, équivalence, redondance, concurrence, etc.

Les travaux empiriques ont exploré l'usage *effectif* des modalités et l'apport *réel* (c'est-à-dire l'efficacité) de la multimodalité dans des situations d'interaction avec des systèmes multimodaux plus ou moins simulés. Du point de vue de l'usage, ces études ont permis de montrer comment les utilisateurs combinent différentes modalités pour interagir avec les systèmes et dans quels cas il utilise la multimodalité. Certaines études, telle celle d'Oviatt, DeAngeli, & Kuhn (1997) par exemple, ont ainsi montré que la multimodalité n'est pas toujours utilisée (20% environ du temps d'une session d'interaction) ; les cas d'usage de celle-ci apparaissent lorsque les utilisateurs décrivent dans leurs commandes des informations spatiales (par exemple : localisation, taille, orientation ou forme d'un objet). Par ailleurs, plusieurs types de combinaison ont été observés (Guyomard, Le Meur, Poignonnet & Siroux 1995 ; Mignot & Carbonel, 1996 ; Oviatt et al. 1997, etc.) : des combinaisons de complémentarité (par exemple, l'utilisateur complète un énoncé verbale en pointant par désignation tactile un objet visé sur l'écran) et des combinaisons de redondance (par exemple, l'utilisateur désigne oralement et explicitement un objet et, en *même* temps, le désigne tactilement sur l'interface). Du point de vue de l'efficacité, peu d'études ont systématiquement abordé ce point et les critères d'évaluation utilisés tournent la plupart autour du gain de temps qu'apporte les interfaces multimodales comparées aux interfaces monomodales (Exemple, ).

Bien qu'ils aient apporté de nombreux résultats intéressants sur l'usage de la multimodalité, ces travaux n'ont abordé qu'un ensemble limité de modalités en entrée (principalement la parole, le pointage ou l'écriture). En outre, peu de ces études ont analysé les phénomènes d'appropriation de la multimodalité au travers de nombreuses sessions d'interaction avec le système. Enfin, les tâches proposées aux sujets étaient le plus souvent « nouvelle » pour les utilisateurs. 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 ».

L'objectif général de l'étude présentée ici vise à contribuer à la conception des systèmes multimodaux et à l'analyse de leur usage en abordant ces différents points, c'est-à-dire :

- Appréhender la multimodalité dans des situations d'interaction « signifiantes » pour les utilisateurs, c'est-à-dire des contextes de tâches/activités qui leur sont familiers.
- Etudier les processus d'appropriation de la multimodalité au travers de plusieurs sessions d'interaction
- Explorer l'usage de modalités « nouvelles »/ « originales » ceci afin de voir empiriquement quelle place elles prennent dans l'interaction homme-machine, les problèmes qu'elles posent individuellement et dans leur relation avec d'autres modalités.

Par ailleurs, par cette étude nous visons également à répondre à un certain nombre de questions précises relatives à l'usage de la multimodalité : comment s'effectue et qu'est-ce qui guide le choix la multimodalité ? Comment s'effectuent le choix et les changements de modalité de l'utilisateur ? Quel est le rôle des propriétés des modalités dans ce choix ? Quels sont les critères qui amènent un utilisateur à choisir ou abandonner une modalité d'interaction au profit d'une autre ? Comment les utilisateurs combinent les modalités ?

D'un point de vue pratique, notre objectif est d'arriver à dégager à des recommandations ergonomiques aussi génériques que possible pour la conception de systèmes multimodaux.

Dans la suite de cet article, nous partirons d'un certain nombre de travaux (Coutaz & Nigay, 1996) pour considérer 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).

## **Méthode**

Afin de répondre à ces questions, nous avons opté pour le développement d'un système multimodal partiellement simulé (utilisation de la technique du Magicien d'Oz). Notre objectif était de contraindre aussi peu que possible l'interaction tout en limitant le coût de développement du système et de nous affranchir des limitations technologiques actuelles à des fins exploratoires centrées sur de nouvelles modalités.

## **Sujets**

10 sujets rémunérés ont participé à cette phase de l'expérimentation ; ils disposaient tous d'une messagerie électronique et étaient des utilisateurs réguliers d'internet mais aucun n'était informaticien.

## **Tâche et nature des modalités testées**

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) qui a été connecté pour les besoins de l'expérimentation à leur messagerie personnelle<sup>1</sup>.

---

<sup>1</sup> L'application de consultation a été développée en JAVA 1.2.

Les quatre modalités étudiées ont été implémentées :

- *la modalité vocale*, qui est présentée comme un mode d'interaction en langage naturel sans contraintes particulières (pas d'énoncés prédéfinis à utiliser par les sujets).
- *la modalité tactile*, qui nécessite un stylet pour pointer les boutons de commandes.
- *la modalité gestuelle*, qui repose sur un code prédéfini qui impose de conserver le PDA dans une main et d'effectuer le geste de la main libre devant l'appareil.
- *la modalité qualifiée improprement (et temporairement...) d'"embodied"*, qui consiste à associer des commandes à certaines manipulations prédéfinies de l'artefact (changements d'orientation, mouvements divers, etc.)

Seule la modalité tactile était réellement implémentée, les autres modalités étaient simulées par le compère grâce à une interface identique et un système de vidéo numérique.



Figure 1 : *Mise en œuvre des modalités tactile, gestuelle et "embodied"*.

Chaque modalité permettait de réaliser un même ensemble de commandes avec comme référentiel celles qui ont définies pour la modalité tactile.

#### Procédure et plateforme expérimentales

L'activité des sujets durant les sessions (d'une durée moyenne de 15 minutes) était enregistrée au moyen de différents capteurs (caméras, capteurs d'événements systèmes) (figure 2). Les sessions expérimentales étaient précédées d'une phase d'apprentissage et de rappel, et étaient espacées de 1 à plusieurs jours. Elles étaient immédiatement suivies de séances d'autoconfrontation qui faisaient l'objet d'un enregistrement vidéo.

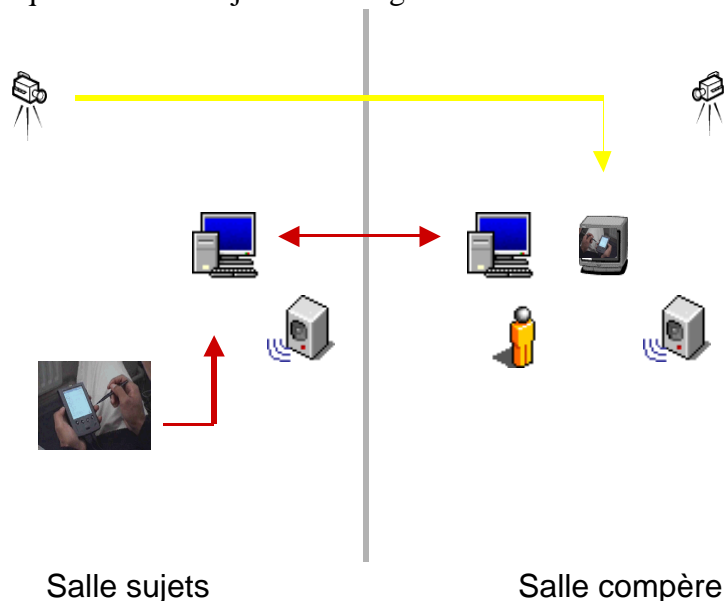


Figure 2 : *Représentation schématique de la plate-forme de simulation Magicien d'Oz.*

## **Traitement et analyse des données**

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<sup>2</sup>. 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.

## **Résultats<sup>3</sup>**

### Usage global

Un des premiers résultats obtenu sur l'usage des modalités, indique que l'ensemble des modalités a été utilisé par les sujets au cours des différentes sessions. Par ailleurs, on observe une évolution dans l'usage de la multimodalité au cours des sessions :

- Lors des premières séances on observe une répartition homogène dans l'usage des modalités qui correspond à une phase de test informel du système. Ce fait est corroboré par les données issues des autoconfrontation.
- Mais rapidement des préférences dans l'usage des modalités apparaissent dans les sessions suivantes
- Ces préférences varient en fonction des sujets.

A un niveau très global (toutes sessions et tous sujets confondus), on ne constate pas de spécialisation d'une modalité à la réalisation d'une ou de plusieurs commandes. Donc, si il y a spécialisation, elle ne peut se situer qu'à un niveau individuel, d'où la nécessité d'effectuer une analyse des données spécifique pour chaque sujet.

### Spécialisation intra-individuelle

Comme mentionné ci dessus, on observe chez de nombreux sujets une tendance à des usages préférentiels de modalités. Ces tendances sont ponctuelles (elles n'apparaissent qu'au cours d'une seule session d'interaction) ou récurrentes (elles apparaissent dans d'autres sessions) individualisées (une même modalité est utilisée pour réaliser une seule et même commande) plurielles (une même modalité est utilisée pour réaliser plusieurs commandes).

### Facteurs influençant les choix de modalité

Les graphes d'activité et les verbalisations des sujets ont permis d'identifier plusieurs facteurs qui sont susceptibles d'orienter le choix et le changement d'usage des modalités. Nous ne présentons ici que les quatre principaux :

- *Contexte de l'activité en cours (procédure opératoire récurrente) :*

Nous avons pu identifier que les tendances à l'usage préférentiel se traduisent le plus souvent par des séquences d'actions récurrentes mises en place par les sujets. Ces séquences sont fortement liées au contexte de l'activité en cours.

- *Changement de but dans l'activité :*

Les changements de modalité sont souvent associés à des phases de réorientation locale de l'activité.

- *Propriétés des modalités (implémentation) :*

---

<sup>2</sup> Le traitement des données a été effectué à l'aide du logiciel Actogram (<http://www.univ-tlse2.fr/ltc/kronos/>).

<sup>3</sup> Pour une présentation détaillée des résultats cf Calvet & al., 2001.

Lors de l'évaluation d'un dispositif multimodal, il convient de bien distinguer les caractéristiques inhérentes à une modalité donnée du mode d'implémentation de cette modalité. Le fait qu'une modalité soit peu utilisée ne signifie pas forcément que cette modalité soit inadaptée mais que sa réalisation technique peut être inadéquate ou contraignante. Ainsi au cours de l'expérimentation la modalité tactile a parfois été sous employée par les sujets du fait de la contrainte induite par la tenue du stylet. Ce dernier est perçu comme un appendice externe au PDA que l'on finit éventuellement par poser pour ne plus avoir à le tenir ce qui va contribuer à encore limiter son utilisation ou à la restreindre à des cas spécifiques (problème avec les autres modalités pour accomplir une action particulière notamment la fermeture des messages d'erreur).

- *Correction d'erreurs*

De nombreux changements de modalités sont apparus dans des situations dysfonctionnelles afin de corriger des erreurs provenant soit du sujet, du système ou bien du compère.

### **Conclusions et perspectives**

Dans cet article nous avons présenté une étude exploratoire de l'usage d'un système multimodal portable. Les résultats obtenus recourent ceux obtenus dans des travaux précédents : usage effectif de la multimodalité (l'ensemble des modalités a été utilisé), différences interindividuelles, apparition de tendances préférentielles), changement de modalité dans les situations dysfonctionnelles. Au-delà de cette convergence, un résultat original de cette étude est que les sujets s'approprient assez facilement des modalités non usuelles (telle que la modalité "embodied" par exemple).

De par leur nature les objets communicants sont susceptibles d'être utilisés dans des contextes très variables. Se pose donc le problème de leur usage dans des situations où les modalités d'interaction ne vont pas être également adaptées. La multimodalité prend alors une importance cruciale dans la mesure où offre aux utilisateurs un moyen de régulation des variations contextuelles (contraintes environnementales, contraintes sociales,...) par l'adoption de la modalité la plus appropriée. C'est ce que illustre cette étude en ce qui concerne par exemple la récupération des dysfonctionnements.

De plus pour chacun des objets usuels considérés, il faudra identifier la ou les modalités d'interaction à intégrer en entrée comme en sortie. Ainsi, au sein d'un environnement comprenant de multiples objets communicants, il faudra définir des mécanismes permettant de spécifier l'objet destinataire de la commande. De ce point de vue la multimodalité pourrait apporter des solutions ; il reste néanmoins à étudier la viabilité conceptuelle et pratique de stratégies d'assignation par construction d'une modalité particulière à un objet ou un type d'objet dans un ensemble d'objets communicants.

## **Bibliographie**

Calvet, G., Kahn, J., Nigay, L., Rey, G., Pasqualetti, L., Salembier, P. & Zouinar, M. (2001) *HOURIA-Nouvelles interactions multimodales*, Rapport final contrat de recherche FT-R&D, CLIPS-IMAG, GRIC-IRIT. FT-R&D DIH-UCE, Issy-les-Moulineaux.

Calvet, G., Kahn, J., Salembier, P., Zouinar, M., Pasqualetti, L., Nigay, L., Rey, G. & Briois, J.C. (2001). Etude empirique de l'usage de la multimodalité avec un ordinateur de poche. *Actes de la conférence "IHM-HCI 2001"*, 10-14 septembre, Lille, France.

DeAngeli, A., Wolff, F., Lopez, P. & Romary, R., Relevance and perceptual constraints in multimodal referring actions (1999) In *Proceedings of the Workshop on Deixis, Demonstration and Deictic Belief, Eleventh European Summer School in Logic, Language and Information" (ESSLLI-99)*, August 9-20, Utrecht, The Netherlands.

Guyomard, M., Le Meur, D., Poignonnet, S. & Siroux, J. (1995) Experimental work for the dual usage of voice and touch-screen for a cartographic application. *Proceedings of the ESCA tutorial and research workshop on Spoken dialog systems*, Vigso, Denmark, 30 may-2 june, pp. 153-156.

Martin, J.C. & Béroule, D. (1993) Types et buts de coopérations entre modalités dans les interfaces multimodales. *Actes des 5èmes journées sur l'ingénierie de l'interaction Homme-Machine*, 19-20 octobre, Lyon, France.

Mignot, C & Carbonell, N. (1996) Commande orale et gestuelle : étude empirique. *Technique et Science Informatiques* 15, pp.1399-1428.

Nigay, L. & Coutaz J. (1996) Espaces conceptuels pour l'interaction multimédia et multimodale, *Technique et Science Informatiques*, spécial Multimédia et Collecticiel, AFCET & Hermes Publ., Vol 15(9), pp. 1195-1225.

Oviatt, S.L. (1999) Ten myths of multimodal interactions, *Communications of the ACM*, Vol. 42, N°11, November, pp. 74-81.

Oviatt, S.L., DeAngeli, A. & Kuhn, K. (1997) Integration and synchronization of input modes during multimodal human-computer interaction. In *Proceedings of Conference on HumanFactors in Computing Systems CHI '97* (March 22-27, Atlanta, GA). ACM Press, New York, , pp. 415-422.