



**HAL**  
open science

# Étude empirique de l'interaction multimodale en mobilité: approche méthodologique et premiers résultats

Guillaume Calvet, Pascal Salembier, Julien Kahn, Moustafa Zouinar

## ► To cite this version:

Guillaume Calvet, Pascal Salembier, Julien Kahn, Moustafa Zouinar. Étude empirique de l'interaction multimodale en mobilité: approche méthodologique et premiers résultats. IHM'05, Sep 2005, Toulouse, France. hal-04319584

**HAL Id: hal-04319584**

**<https://hal.science/hal-04319584v1>**

Submitted on 3 Dec 2023

**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.

# Etude empirique de l'interaction multimodale en mobilité : approche méthodologique et premiers résultats

*Guillaume Calvet, Pascal Salembier*

GRIC-IRIT  
118, rte de Narbonne  
31062, Toulouse Cedex, France  
calvet, salembier@irit.fr

*Julien Kahn, Moustapha Zouinar*

France Telecom R&D  
38-40, rue du Général Leclerc  
92794, Issy-les-Moulineaux, France  
kahn, zouinar@rd.francetelecom.com

## RESUME

Cet article présente les premiers résultats d'une étude de l'usage d'un système multimodal en mobilité dans le contexte d'une situation de jeu collaboratif. Après avoir rappelé les caractéristiques du dispositif de recueil et d'analyse des données, nous donnons quelques éléments de résultats centrés sur la gestion de la multimodalité.

**MOTS CLES :** Mobilité, Multimodalité, Plateforme d'évaluation, Jeu collaboratif

## ABSTRACT

This paper sums up the first results of an empirical study on the use of a multimodal device in mobility in the context of a collaborative game. The main features of the data gathering and analysis platform is described, then some results on the use of multimodality are presented.

**CATEGORIES AND SUBJECT DESCRIPTORS:** H.5.2 [Information Interfaces and Presentation]:User Interfaces--Evaluation/methodology, Voice I/O, Ergonomics, Interaction Styles

**GENERAL TERMS:** Human Factors ; Experimentation

**KEYWORDS:** Mobility, Multimodality, Evaluation platform, Collaborative game.

## INTRODUCTION

La multiplication des systèmes interactifs mobiles contribue à renouveler le champ des problématiques investiguées dans le domaine de l'interaction Homme-machine. La prise en compte de la mobilité peut être traitée de plusieurs points de vue : gestion coordonnée du mouvement [1], de la navigation dans un espace physique [8] et de l'interaction avec le système ; effet des

changements de contexte induits par l'extension de l'espace d'activité potentiel [11] ; conséquences de la mobilité sur les processus sociaux de constitution de réseaux d'acteurs [5]. Les artefacts mobiles posent également de nouveaux problèmes relatifs à la conception des dispositifs et des méthodes d'évaluation à mettre en œuvre pour l'étude empirique de leur usage en situation naturelle. On distingue ainsi par convention deux approches paradigmatiques : les études menées en laboratoire et les études de terrain. Les études en laboratoire se focalisent sur le contrôle de variables susceptibles d'influer sur l'interaction ; elles requièrent de recréer artificiellement une situation de mobilité [1] qui se traduit généralement par un appauvrissement du contexte. Les études naturalistes insistent quant à elles sur le critère de validité écologique et s'efforcent de préserver la richesse contextuelle des activités mobiles. Récemment certains auteurs ont proposé sous les termes de « field experiments » [8] et de « quasi-experiments » [10] une voie médiane qui tentent de concilier contrôle expérimental des variables et réalisme de la situation. Cette approche demeure essentiellement quantitative et s'inscrit dans la tradition orthodoxe de la psychologie expérimentale appliquée à l'interaction Homme-machine. Le travail présenté ici se rattache à la catégorie des études naturalistes, mais il se donne des contraintes de rigueur quant aux phases de recueil et d'analyse des observables. Il s'efforce en outre d'articuler les données quantitatives (qui constituent des éléments de synthèse commodes et permettent de résumer les tendances générales des résultats) avec des données contribuant à donner du sens aux comportements observés (verbalisations des sujets).

## OBJECTIFS

Cette étude s'inscrit dans la continuité d'un premier travail mené sur l'utilisation d'un système portable multimodal en laboratoire [3]. Notre objectif est cette fois de tester empiriquement les effets potentiels de la mobilité sur l'usage d'un système de ce type en situation réelle, et de documenter empiriquement en quoi la versatilité des systèmes multimodaux, souvent présentée comme un de leurs avantages principaux [7], peut aider l'utilisateur à gérer les contraintes induites par les changements de contexte inhérents à la mobilité (modification de l'environnement physique et social notamment). Dans

cet article nous décrivons le dispositif de recueil et d'analyse des données mise en oeuvre au cours d'une étude menée dans l'espace visiteurs de la Cité des Sciences de La Villette dans le cadre de la plateforme LUTIN [6]. Quelques uns des premiers résultats obtenus seront également présentés.

## DESCRIPTION DU SYSTEME

Pour cette expérimentation, une application de messagerie multimodale et multimédia a été développée sur des PDAs de type Ipaq© équipés d'une carte réseau sans fil 802.11. Les messages échangés peuvent contenir conjointement un certain nombre de médias tels que du son, des images, du texte dactylographié ou encore du texte manuscrit ou des dessins. Trois modalités d'interaction ont été implémentées : la **modalité tactile** « **stilet** » classique qui permet de pointer sur les boutons correspondants à des commandes et sur les zones de texte. La **modalité vocale** : l'application est capable de reconnaître une trentaine de mots clés. Il est possible de combiner certaines commandes vocales, par exemple « Ouverture du message 5 ». La **modalité geste 2D** : l'application est également capable de reconnaître une demi-douzaine de gestes en deux dimensions, effectués à l'aide du stilet sur l'écran du PDA. Ces gestes sont prédéfinis et limités essentiellement à quelques commandes d'édition (copier, coller, supprimer, nouveau message,...). En sortie, la messagerie dispose d'une synthèse vocale qui peut être activée et désactivée. Lorsque les commandes sont vocales, il y a systématiquement un feedback vocal de réussite ou d'erreur. La synthèse peut également, à la demande, lire les informations affichées dans l'écran courant, l'en-tête des messages ou leur contenu.

## METHODE

### Tâche et consigne.

La tâche consistait pour des couples de sujets à répondre collaborativement aux questions d'un jeu de type quiz. Les réponses devaient être trouvées dans des espaces distincts couvrant chacun approximativement 500 m<sup>2</sup> (les sujets ne se rencontraient donc pas physiquement). Chacun des sujets avait cinq questions sur son PDA, et disposait de cinq indices correspondant aux questions de l'autre sujet. D'un point de vue expérimental l'objectif était ici double : favoriser les interactions entre les sujets via le PDA et provoquer un engagement réel dans l'exploration de l'environnement.

### Sujets et protocole expérimental

Après une session de familiarisation avec le système où l'ensemble des fonctions et des modalités leur étaient montrées, 6 sujets ont effectué trois sessions d'interaction de 40 à 50 minutes, toutes espacées de plusieurs jours. Les sujets ont été recrutés et rémunérés pour les besoins de l'expérimentation. Ils étaient tous familiarisés avec l'usage du téléphone portable et des PDAs. Chaque session était suivie d'une autoconfrontation individuelle.

## Dispositif de recueil des données

Le dispositif est composé de trois modules de recueil<sup>1</sup> : un système de prise de vue large qui permet de situer l'utilisateur dans son contexte global d'usage (vue contextuelle observateur) ; un système de prise de vue du champ visuel de l'utilisateur (vue « subjective ») réalisé au moyen de lunettes-caméra (figure 1) ; un système de capture automatique des actions de l'utilisateur sur le système et les sorties de celui-ci (capture log).



Figure 1 : Vues « contextuelle » et « subjective »

## Analyse et codage des données

Nous avons utilisé pour coder les données audio-vidéo une version optimisée du système d'analyse de données Actogram<sup>TM</sup> [4] qui ouvre plusieurs fenêtres vidéo actives simultanément, et permet de piloter et de visualiser de façon synchrone les deux prises de vue (contextuelle et subjective). Cet outil permet en outre de réaliser directement le codage des observables définis au préalable par l'analyste ; il permet également de construire des chroniques d'activité. Les données qui ont fait l'objet de ce codage systématique comprennent : l'usage des modalités, les conditions de mobilité (statique, micromobilité, déplacement), les directions de regard lors de l'utilisation du système, le type d'activité en cours, les informations contextuelles (selon une catégorisation inspirée de [9] qui distingue entre environnement, contexte et situation).

## PREMIERS RESULTATS

Les premiers résultats présentés ici s'appuient uniquement sur des données quantitatives. La grille de lecture utilisée dans ce premier temps de l'analyse s'efforce de donner des éléments de réponse aux questions suivantes : quel est le poids respectif de chaque modalité dans l'utilisation du système par les sujets ? Y'a-t-il des différences interindividuelles marquées ? Constate-t-on une évolution entre la première et la dernière session expérimentale ? Observe-t-on un début de spécialisation (association récurrente d'une modalité et d'une commande) ? La mobilité joue-t-elle un rôle sur l'usage des modalités ?

### Usage global des modalités (stilet, vocal, g2d)

La figure 2 indique un usage prédominant de la modalité tactile stilet (75,6%, n=4077) ; la modalité vocale ne représentant que 22,7% des commandes envoyées par les sujets. On retrouve ici des résultats classiques [7] qui

<sup>1</sup> Pour plus de détails sur le dispositif d'observation et d'analyse voir [12].

viennent contredire des positions « naïves » relatives à la primauté supposée du vocal dans les interfaces multimodales. La part du geste 2D est quant à elle quasi négligeable (1,7%).

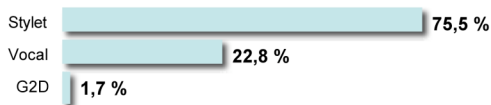


Figure 2 : Usage global des modalités toutes commandes, toutes sessions et tous sujets confondus

### Usage des modalités par sujets

La figure 3 indique une forte variabilité inter individuelle dans l'usage des modalités. On peut néanmoins distinguer trois groupes de sujets. Certains sujets (1 et 2) utilisent très peu la modalité vocale (en dessous des 4%). A l'inverse les sujets 4 et 6 l'ont beaucoup exploitée (46,3% pour le sujet 6 et 42,5% pour le sujet 4).

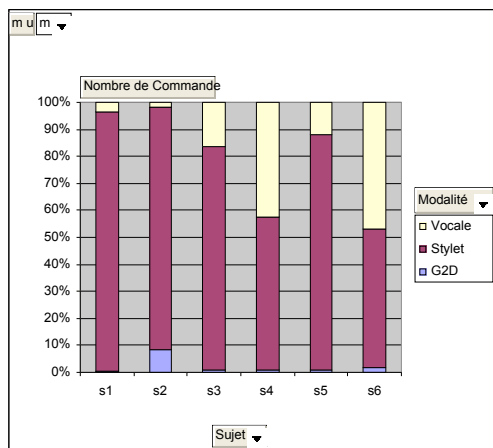


Figure 3 : Usage global des modalités par sujet, toutes commandes et toutes sessions confondus

Les sujets 3 et 5 occupent une position intermédiaire (avec respectivement 16,2% et 11,8% d'usage de la modalité vocale. En ce qui concerne le geste 2D, seul le sujet 2 atteint 8,7% alors que son usage chez les autres sujets reste anecdotique.

### Evolution des usages par session

Les résultats obtenus sur les trois sessions permettent de distinguer 2 catégories de sujets ; les sujets pour lesquels les choix de modalité sont quasiment acquis dès la première séance, et les sujets pour lesquels on observe une évolution importante entre la première et la dernière session. Il s'agit essentiellement d'un accroissement significatif de l'usage de la modalité vocale au détriment de la modalité tactile « stilet ». Si l'on considère par exemple le sujet 4, on note une inversion totale de la modalité dominante : l'usage de la modalité tactile « stilet » passant de 83,7% à la première interaction à 36,9% à la der-

nière session. Corrélativement, la modalité vocale passe de 13,6 à 63,1%.

### Usage spécialisé global de modalités par commandes

Les résultats mettent en évidence qu'un nombre important de commandes sont effectuées uniquement avec la modalité stilet (figure 4). Pour l'ensemble des sujets on n'observe pas de spécialisation intégrale de la modalité vocale pour une commande donnée. Il existe, cependant, une certaine tendance à la préférence vocale pour quelques commandes telles que « envoyer », « insérer », « ouvrir », etc.

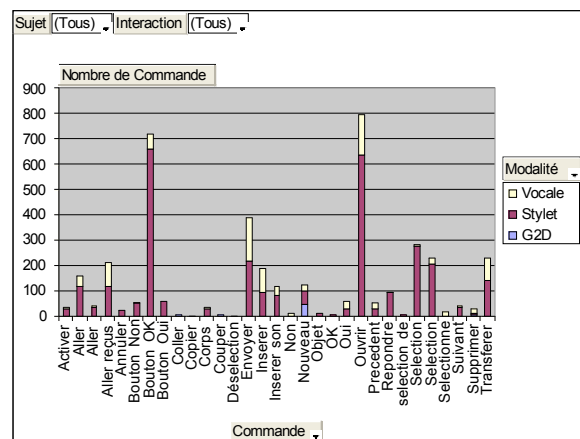


Figure 4 : Usage global des modalités par commandes tous sujets et sessions confondus

Si l'on considère les résultats sujet par sujet on constate que le phénomène de spécialisation est davantage marqué. Ainsi le sujet 4 montre par exemple une préférence marquée pour la modalité vocale (commandes <Envoyer>,...) bien que certaines commandes ne sont jamais effectuées avec cette modalité (<Précédent>, <Répondre>, <Annuler>, <Sélection>,...). Cette tendance se confirme au cours des sessions : le G2D est ainsi totalement abandonné et certaines commandes réalisées principalement en vocal dans la première session, le sont exclusivement à la dernière session.

### Effets de la mobilité sur l'utilisation des modalités

Les résultats sont encore incomplets, le dépouillement des données étant en cours (plus du tiers des résultats déjà codés). Cependant, des tendances intéressantes se dégagent déjà. Lors du codage de la mobilité, nous avons distingué trois catégories : déplacement (déplacement rapide d'un point à un autre avec une destination établie), micromobilité (piétinement, déambulation sans but clair), statique (aucun déplacement).

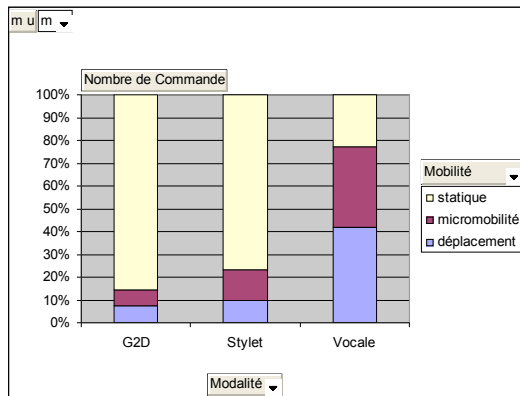


Figure 5 : Usage des modalités en mobilité

Les données disponibles (figure 5) indiquent comme on pouvait s'y attendre que les modalités qui requièrent l'utilisation du stylet (G2D et Tactile Stylet) et conséquemment un niveau de contrôle visuel précis de l'action sont effectuées à 80% en statique alors que la modalité vocale est utilisée quasiment à 80% en mobilité (« déplacement » plus « micromobilité »).

## CONCLUSION

Les premiers résultats présentés dans cet article devront être affinés et complétés. La catégorisation des informations contextuelles permettra notamment de tester certaines hypothèses sur l'effet possible des facteurs physiques d'ambiance et de la présence de tiers sur la gestion de la multimodalité. Plus globalement l'analyse des verbalisations des sujets permettra la mise en perspective systématique des données quantitatives de synthèse et de leur signification du point de vue des sujets. L'objectif étant ici de réduire la fracture entre approche externaliste de l'évaluation des systèmes (sur des critères tels que temps de réalisation, nombre d'erreurs) et point de vue en première personne (jugement du sujet sur son vécu expérientiel) qui donnent parfois des résultats contradictoires [2].

## REMERCIEMENTS

Cette étude a bénéficié d'un financement France Telecom R&D (convention MOTIVE) et d'un financement RNRT (plateforme LUTIN).

## BIBLIOGRAPHIE

1. Beck, E., Christiansen, M., Kjeldskov J., Kolbe, N. & Stage, J. Experimental Evaluation of Techniques for Usability Testing of Mobile Systems in a Laboratory Setting. In *Proceedings of OzCHI 2003*, Brisbane, Australia, 2003.

2. Cahour, B., Salembier, P., Brassac, Ch., Bouraoui, J.L., Pachoud, B., Vermersch, P. & Zouinar, M. Methodologies for evaluating the affective experience of a mediated interaction. *Workshop on Innovative Approaches to Evaluating Affective Interfaces, CHI 2005*, April 2-7, Portland, Oregon, 2005.
3. Calvet, G., Kahn, J., Salembier, P., Zouinar, M., Pasqualetti, L., Nigay, L., et al. Etude empirique de l'usage de la multimodalité avec un ordinateur de poche. In A. Blandford, J. Vanderdonck & P. Gray (Eds.), *People and Computers XV - Interaction without frontiers, Joint proceedings of HCI 2001 and IHM 2001* (Vol. 2), Springer, London, 2001.
4. Kerguelen, A. Quels outils concevoir pour aider au relevé d'observation sur le terrain ? *XXXIIIème Congrès de la SELF, "Temps et Travail"*, Paris, 1998.
5. Licoppe, Ch. & Inada, Y. Les usages émergents d'un jeu multijoueurs sur terminaux mobiles géolocalisés. In J. Coutaz et S. Lecomte (eds) : *UBIMOB'2005, Actes des Deuxièmes Journées Francophones Mobilité et Ubiquité*, 31 mai-3 Juin, Grenoble, 2005.
6. LUTIN. <http://www.lutin.utc.fr/>
7. Oviatt, S. L. Multimodal System Processing in Mobile Environments. In *Proceedings of the UIST '00 conference*. San Diego, CA: ACM, 2000.
8. Pirhonen, A., Brewster, S. & Holguin, C. Gestural and audio metaphors as a means of control for mobile devices. In *Proceedings of ACM Conf. on Human Factors in Computing Systems CHI'2002*, ACM Press Addison-Wesley, Minneapolis, 2002, pp. 291-298.
9. Quéré, L. La situation toujours négligée? *Réseaux* (85), 1997.
10. Roto, V., Oulasvirta, A., Haikarainen, T., Kuorelahi, J., Lehmuskallio, H., & Nyysönen, T. *Examining mobile phone use in the wild with quasi-experimentation*, (HIIT Technical Report No. 2004-1). HIIT, Helsinki, 2004.
11. Salembier, P., Kahn, J., Calvet, G., Zouinar, M. & Relieu, M.(2005). "Just follow me". Examining the use of a multimodal mobile device in natural settings. In *Proceedings of the HCI International Conference*, July 22-27, Las Vegas. Lawrence Erlbaum Associates.
12. Zouinar, M., Relieu, M., Salembier, P., & Calvet, G. Observation et capture de données sur l'interaction multimodale en mobilité. In *Actes des premières journées francophones Mobilité et Ubiquité 2004, 1-3 juin, Nice, Sophia-Antipolis*: ACM, 2004.