

Validité d'une évaluation in situ via la méthode du Magicien d'Oz

Merlin Bruno

IRIT – DIAMANT
118 Route de Narbonne
31062 Toulouse
05.61.55.63.05
merlin@irit.fr

RÉSUMÉ

Dans le cadre d'évaluations de dispositifs à commandes multimodales sur Pocket PC et d'études d'usage in situ de la multimodalité, nous avons jugé opportun d'étudier la possibilité d'utiliser la méthode du « Magicien d'Oz » en milieu non contrôlé. Des pré-expérimentations, destinées à évaluer l'usage de claviers logiciels multimodaux sur Pocket PC, nous ont amenés à tester la validité de cette approche Magicien d'Oz en contexte d'observation *in situ*.

Mots clefs : Magicien d'Oz, mobilité, multimodalité, méthodologie d'évaluation.

ABSTRACT

In the context of studies dealing with the use of multimodality in a non-controlled environment, we decided to try to apply the Wizard of Oz methodology in a context *in situ*. Pre-experimentation, supposed to evaluate multimodal software keyboard for Pocket PC, brought us to test the validity of this methodology in this observation context.

CATEGORIES AND SUBJECT DESCRIPTORS : H.5 [CHI], J.4 [SOCIAL AND BEHAVIORAL SCIENCES].

GENERAL TERMS : Experimentation, Human Factors, Verification

Keywords : Wizard of Oz, mobility, multimodality, evaluation methodology.

INTRODUCTION

Des techniques d'interaction comme la reconnaissance vocale ou la reconnaissance gestuelle et la multimodalité

présentaient, initialement, le charme de leur nouveauté. Elles ont été confrontées aux contraintes et restrictions liées à leur usage (cf. les 10 Mythes d'Oviatt [12]) et ne se sont finalement que peu développées dans les applications grand public. Le développement des technologies d'informatique mobile, à travers la démocratisation des PDAs, l'omniprésence et la poly-potence des téléphones, ou encore l'apparition des *wearable computers*, repose la question de l'usage des modalités d'interaction. Des modalités qui ne se présentaient que dans de rares contextes comme une alternative efficace aux périphériques d'interaction classiques (claviers/souris en entrée, écran en sortie), pourraient obtenir une importance nouvelle en contexte de mobilité. La multimodalité pourrait ainsi permettre de palier le handicap, en terme d'interaction, procuré par le contexte de mobilité et la taille réduite des IHMs.

Consécutivement à cet essor des technologies d'informatique mobile, un volet d'études, concernant l'usage de la multimodalité en situation de mobilité, a vu le jour. Cependant, l'implémentation des technologies d'interaction, nécessitant un investissement technique et financier non négligeable, reste contraignante, incitant à mener des expérimentations sous la forme de séance de Magicien d'Oz [1].

L'évaluation d'un dispositif destiné à un usage *in situ* nous a amené à sortir la méthodologie du Magicien d'Oz de son contexte normal d'utilisation pour expérimenter la réalisation d'une séance de Magicien d'Oz en milieu non contrôlé. Nous avons donc mis en place une pré-expérimentation destinée à étudier les facteurs permettant de favoriser le bon déroulement et le respect des contraintes méthodologiques, ainsi que l'obtention des données relatives à l'expérimentation.

MULTIMODALITE & EVALUATION

En matière d'étude des usages de la multimodalité, la technique du Magicien d'Oz présente l'avantage principale de permettre une simulation des modalités d'interactions

qui, pour des raisons techniques ou pratiques, peuvent ainsi rester non implémentées [13]. Un « Compère » situé dans une pièce voisine communique à distance, soit directement avec l'utilisateur (dans le cadre de la simulation d'un système vocal interactif [9], [7], par exemple) ou avec l'application de l'utilisateur, en répercutant à distance les commandes effectuées par l'utilisateur (applications multimodales par exemple [11]). Dans les deux cas, cette méthode implique que le compère puisse soit voir, par l'intermédiaire de caméras, soit seulement entendre l'utilisateur de manière à percevoir l'information nécessaire à la simulation.

Plusieurs contraintes sont cependant induites par la séance de Magicien d'Oz, l'optique étant de permettre au « Compère » d'opérer la simulation de l'usage des modalités sans que l'utilisateur sujet d'expérimentation ne soit conscient de cet aspect simulé. Le but est que l'utilisateur interagisse le plus naturellement possible avec le dispositif évalué [7]. En effet, des études ont montré que la connaissance de l'aspect simulé de l'interaction, notamment pour l'interaction en langage naturel, influe de façon importante le vocabulaire utilisé par l'utilisateur [3].

Les contraintes méthodologiques se posent ainsi :

- En terme d'observation : sans être intrusif, comment accéder aux données pour opérer la simulation, comment récolter les données de l'observation ;
- En terme de communication : comment simuler l'interaction de l'utilisateur.

Dans un grand nombre de projets existants traitant de l'étude de la multimodalité en situation de mobilité, comme l'évaluation d'un système vocal d'aide au guidage [10], ou l'étude de l'interaction naturelle dans une voiture [2] (VICO), ou encore dans le cadre de NIM [5], les contraintes liées à la séance de magicien d'Oz ont été résolues par la limitation de l'observation à l'étude en milieu contrôlé. Les utilisateurs étaient placés dans un milieu techniquement adapté au Magicien d'Oz.

Nous avons évalué la validité de l'application d'une séance *in situ* de Magicien d'Oz à travers la réalisation d'une pré-expérimentation. Seuls les résultats concernant la méthodologie seront abordés dans le cadre de cet article.

METHODE

Tâche & Support d'expérimentation

Le support de l'expérimentation est un éditeur de texte muni d'un clavier logiciel développé sur Pocket PC. Le clavier logiciel est de type *fullscreen keyboard* [12], dont les touches sont accessibles par l'intermédiaire des deux pouces, auquel a été ajouté une liste de complétion.

Deux modes de saisie sont envisagés par l'éditeur de texte : un mode de saisie susceptible de favoriser à proprement parler l'activité de saisie ; et un mode classique (cf. Figure 1) où l'utilisateur est plutôt supposé réaliser des tâches d'édition.

Langage d'interaction multimodale

Plusieurs modalités d'interaction, destinées à améliorer l'accessibilité de certaines fonctions du clavier, ont été ajoutées au dispositif. Celles-ci sont soit implémentées, soit simulées par le compère.

Une modalité *embodied* [13] (tangible dont l'artefact est l'ordinateur lui-même) permet, tout d'abord, de passer d'un mode de saisie à l'autre. Cette modalité se justifie par la nécessité pour l'utilisateur d'effectuer une rotation lorsqu'il souhaite passer du mode de « saisie » au mode d'« édition » (cf. Figure 1). Elle permet à l'utilisateur : de joindre une commande à la manipulation de l'artefact et de ne pas nécessiter d'interaction supplémentaire pour activer le clavier. Comme nous ne disposons pas d'un module de reconnaissance de mouvement, celle-ci sera en conséquence simulée. A noter qu'il existe sur le marché un périphérique de reconnaissance tangible susceptible d'être adapté à cet usage [15] (cf. Figure 2).

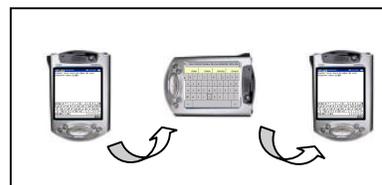


Figure 1 : Utilisation de la reconnaissance tangible



Figure 2 : ECER TiltControl

La reconnaissance vocale, également simulée, permet de réaliser des opérations d'édition et de sélection dans la liste de complétion. Ces commandes permettent de supplanter une et surtout plusieurs opérations tactiles (comme les opérations d'édition. Ex : ctrl + c, ou ouverture d'un menu suivi de la sélection d'un item du menu). Une modalité de reconnaissance de geste 2D (cf Figure 3), implémentée, offre des possibilités équivalentes aux commandes offertes par la reconnaissance vocale.

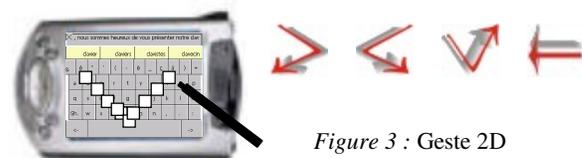


Figure 3 : Geste 2D

Lors de l'expérimentation, la tâche de l'utilisateur est de relever, dans le campus universitaire, des informations affichées sur les panneaux d'information.

Architecture : observation et compère

Compère. Plusieurs impératifs viennent contraindre le développement du poste du compère. Le compère doit pouvoir observer visuellement l'utilisateur, a priori de façon médiée ou non, dans le but de simuler l'interaction *embodied*. La prohibition de moyens invasifs, comme les caméras, sous-entend l'impossibilité d'une interaction visuelle médiée et la nécessité d'une interaction visuelle directe entre le compère et l'utilisateur. L'application du compère se doit donc d'être embarquée. Développée sur Pocket PC, l'application du compère communique avec l'application de l'utilisateur par WiFi. Le compère peut ainsi rester près de l'utilisateur le long de son parcours.

Le nombre réduit de commandes à simuler (commandes sans sémantique complexe) permet d'exploiter simplement la totalité de l'interface du Pocket PC. Celle-ci est divisée, en fonction du mode actif (saisie ou édition), en un nombre limité de boutons représentant les commandes vocales. Un bouton supplémentaire change le mode actif, simulant parallèlement l'interaction *embodied* chez l'utilisateur.

Le contact presque direct de l'utilisateur et du compère soulève la question de l'anonymat du compère et pourrait remettre en cause la validité de la méthodologie. Nous expliquerons plus loin l'approche que nous avons mise en œuvre pour contourner ce problème.

Observation des données. De façon à pouvoir être collectées de manière non intrusive, les données de l'interaction entre l'utilisateur et l'éditeur de texte (éditeur + clavier) ainsi que les interactions médiées entre le compère et l'utilisateur sont acquises automatiquement par les supports logiciels. Les traces de l'interaction sont ensuite transmises sous forme de messages textuels, par réseau, à une plate-forme d'observation distante. Cette plate-forme d'observation capture et présente les données de l'interaction en vue d'analyses qualitatives et statistiques.

Évaluation de la méthodologie

La pré-expérimentation a engagé 10 personnes. Ces personnes étaient toutes issues du domaine de l'IHM et familières des techniques de Magicien d'Oz. L'origine des personnes constituait un risque supplémentaire d'échec et d'invalidation de la méthode. En effet, une personne au courant de l'état de l'art ainsi que des techniques courantes d'évaluation est plus spontanément portée à soupçonner l'activité du compère et le contexte Magicien d'Oz. Nous devons par conséquent trouver un rôle et une justification à la présence du compère. Deux possibilités se sont offertes : présenter le compère comme un autre utilisateur passant simultanément l'évaluation sur un autre prototype ; ou bien désigner le compère comme une personne accompagnatrice

destinée à prendre de simples notes et veiller au bon déroulement de l'expérimentation.

La première solution présentait plusieurs intérêts, et particulièrement celui de détacher complètement l'utilisateur du sentiment d'être observé. Cependant cette solution risquait aussi de motiver une communication de l'utilisateur vers le compère, cette communication pouvant être dommageable tant pour le rôle du compère que pour l'activité de l'utilisateur. Cette solution interdit par ailleurs l'omniscience du compère et son pouvoir d'apporter une solution en cas de problème technique ou circonstanciel.

Le compère a été finalement introduit comme un observateur (cf. Figure 7) destiné à guider l'utilisateur dans le campus et recueillir d'éventuelles remarques spontanées de la part de celui-ci. Un bouton physique du Pocket PC permet en outre au compère de pouvoir masquer son interface par une fausse fenêtre de prise de notes pré-remplie.

Utilisateur

Compère



Figure 7 : Déroulement de la séance de magicien d'Oz

Afin d'écartier le plus possible les soupçons qui auraient pu peser sur le rôle exact du compère, les consignes données aux utilisateurs insistaient sur la nécessité de s'appliquer à la réalisation des interactions sur des périphériques peu fiables. Après le déroulement de la pré-expérimentation, un débriefing a été effectué avec les utilisateurs. Nous occulterons ici les résultats concernant le reste de la pré-expérimentation pour centrer nos remarques sur l'évaluation de la méthodologie Magicien d'Oz.

Ce débriefing était tout d'abord destiné à obtenir des informations sur leurs connaissances des techniques de Magicien d'Oz. Le débriefing devait ensuite permettre de déterminer la conscience que les utilisateurs avaient eue de l'aspect simulé de l'expérimentation. Il a été réalisé sous la forme d'un questionnaire destiné à amener progressivement l'utilisateur sur la piste de la séance de Magicien d'Oz. L'aspect progressif des questions permettait de mesurer le degré de « conscience » qu'avait l'utilisateur de l'aspect simulé des interactions.

RESULTATS

Malgré le rapport des utilisateurs à l'IHM et leurs connaissances approfondies des techniques de Magicien

d'Oz. La majorité des utilisateurs (7 sur 10) n'a pas soupçonné l'usage de celle-ci.

Le facteur ayant permis à deux des trois personnes d'identifier l'usage de la méthode de Magicien d'Oz était une très bonne connaissance des PDAs et des dispositifs permettant l'interaction tangible. Cette connaissance leur a permis d'identifier que le Pocket PC n'était pas équipé pour permettre une reconnaissance *embodied* et de déduire le rôle du compère. Dans une expérimentation future utilisant cette méthodologie d'évaluation, il sera donc impératif de tenir compte de cet aspect et d'éventuellement installer un leurre sur le PDA.

Le facteur ayant permis à la troisième personne de soupçonner simplement l'usage du Magicien d'Oz était lié aux trop bonnes performances de la reconnaissance vocal relativement au cadre bruyant de l'environnement.

La proximité du compère avec les utilisateurs, dans un contexte naturel et public (induisant la présence d'autres personnes extérieures à l'expérimentation), ne s'est pas montrée intrusive. Elle a ainsi permis à celui-ci de remplir convenablement son rôle. Par ailleurs, les facteurs qui semblent avoir favorisé le fait que 7 personnes sur les 10 n'aient pas soupçonné la séance de Magicien d'Oz ont été les consignes données lors de la prise en main du dispositif et l'aspect naturel de l'interaction, renforcé par le cadre naturel du déroulement de l'expérimentation. Le contexte non contrôlé et hors laboratoire, ainsi que l'absence de dispositifs ostensibles d'observation ont permis à l'utilisateur de pouvoir faire abstraction du cadre expérimental de son activité.

CONCLUSION & PERSPECTIVES

On doit préalablement noter que nos contraintes expérimentales étaient assez restreintes. Restreintes en matière d'observation, dans la mesure où les données statistiques acquises au moyen de la plate-forme d'observation étaient suffisantes et où nous n'avions donc pas besoin d'enregistrements vidéos, et restreintes en matière de langage d'interaction à simuler. Celui-ci étant fini et restreint, l'application du compère et son usage en situation de mobilité étaient simplifiés.

Cependant, bien que nous ne puissions généraliser l'usage de cette méthode, nous avons mis en évidence non seulement la possibilité de réaliser des séances de Magicien d'Oz en contexte *in situ*, mais également le bénéfice apporté par le contexte *in situ* sur la méthodologie elle-même. Elle nous engage également à privilégier un dispositif simple d'observation, intégré au milieu naturel, par rapport à un dispositif intrusif contraignant physiquement l'utilisateur et susceptible de le transformer en objet de curiosité pour son environnement.

REMERCIEMENTS

Je remercie J. Phalip, S. Chamberlhac et L. Benech, (ex-) étudiants en DESS IHM pour l'aide qu'ils m'ont apportée dans l'élaboration du support de cette expérimentation.

REFERENCES

1. Amalberti, R., Valot, C. Le Magicien d'Oz, CERMA, 1992.
2. Bensen, N. O., Dybkjær, L. Exploring natural interaction in the car. Workshop on Information Presentation and Natural Multimodal Dialogue, Verone, décembre 2001, pp. 75-79.
3. Brennan, S.E. "Conversation as Direct Manipulation: An Iconoclastic View". In "The Art of Human-Computer Interface Design", B. Laurel ed., 1990.
4. Buisson, M., Bustico, A., Chatty, S., Colin, F.R., Jestin, Y., Maury S., Mertz, C., Truillet, P. Ivy : un bus logiciel au service du développement de prototypes de systèmes interactifs. IHM'02, Poitiers.
5. Calvet, G., Kahn, J., Zouinar M., Salembier P., Briois J-C., Nigay, L., Rey., L. Pasqualetti Étude empirique de l'usage de la multimodalité avec un ordinateur de poche. IHM-HCI2001, CEPADUES Publ., pp. 5-8.
6. Dahlbäck, N., Jönsson, A. "Empirical studies of discourse representations for natural language interfaces". Chapter of the ACL, 291-8, 1989.
7. Dahlbäck, N., Jönsson, A., Ahrenberg, L. Wizard of Oz Studies-Why and How. IUI, 1993.
8. Fraser, N. M., Gilbert G. N. Simulating speech systems. Computer Speech and Language, 1991.
9. Fraser, N., Gilbert N., McDermid, C. "The Value of Simulation Data". ANLP, Trento, Italie, avril 1992.
10. Manstetten, D., Krautter, W., Grothkopp, B., Steffens, F., Geutner, P. Using A Driving Simulator To Perform A WOZ Experiment On Speech-Controlled Driver Information Systems. HCTSC 2001.
11. Nigay, L., Coutaz, J. A Design Space For Multimodal Systems: Concurrent Processing and Data Fusion Laurence. INTERCHI '93, 1993, pp. 172-178.
12. Oviatt, S.L. Ten myths of multimodal interactions. ACM Vol. 42, N°11, Novembre, 1999, pp. 74-81.
13. Salber, D., Coutaz, J., Applying the Wizard of Oz Technique to the study of Multimodal Systems. HCI'93, août 93.
14. Spb Software House, SPB Full Screen Keyboard, <http://www.spbsoftwarehouse.com/products/fsk/?en>
15. TiltControl, ECERTech. <http://www.ecertech.com/tiltcontroldevice.aspx>