

Usability Study of Multi-Modal Interfaces using Eye-Tracking

Regina Bernhaupt¹, Philippe Palanque², Marco Winckler² and David Navarre²

¹ ICT&S Center, Universität Salzburg, Sigmund-Haffner-Gasse 18,
5020 Salzburg, Austria
Regina.Bernhaupt@sbg.ac.at

² IRT, Université Toulouse III,
{palanque, [winkler. navarre](mailto:winkler.navarre@irit.fr)}@irit.fr

Abstract. The promises of multimodal interaction to make interaction more natural, less error-prone and more enjoyable have been controversially discussed in the HCI community. On the one hand multimodal interaction is being adopted in fields ranging from entertainment to safety-critical applications, on the other hand new forms of interaction techniques (including two-handed interaction and speech) are still not in widespread use. In this paper we present results from a usability evaluation study including eye-tracking on how two mice and speech interaction is adopted by the users. Our results show evidence that two mice and speech can be adopted naturally by the users. In addition, we discuss how eye-tracking data helps to understand advantages of two-handed interaction and speech.

Keywords: Multimodal interfaces, usability evaluation method, two mice, speech.

Introduction

Making interaction more natural is one of the promises of multimodal interfaces. Research on evaluating multimodal interfaces dates back to the mid 1980s [5] and since then it is seen as a promising approach to improve desktop interaction techniques. Multimodal interaction techniques are considered as a potential way to increase communication bandwidth between users and systems and to enhance users satisfaction and comfort by providing a more natural way of interacting with computer systems [5], [6].

Several studies have shown that using two pointing devices in a normal graphical user interface is a more efficient and understandable interaction than using basic mouse and keyboard [8], [19], [28]. On the one hand there exists a strong foundation on two-handed input (see [15] for an overview), on the other hand still two-handed interaction and speech is not in common use. This might have various reasons. In addition to subjective factors like comfort and satisfaction, increasing communication

bandwidth between users and systems can have a significant impact on efficiency. For instance the number of commands triggered by the users within a given amount of time and the error rate (typically both the number of slips and mistakes made by the users [26]), are influenced by the user interface and its underlying interaction techniques. While complementarity of modalities can be used to reinforce and clarify the communication between the users and the system [23], studies of Dillon et al. [9] and by Kjeldskov and Stage [21] “unsurprisingly” revealed that when multimodal interfaces are poorly designed they are neither better understood nor more efficient than any other user interface offering more standard interaction techniques. In order to determine the impact of modalities on interaction, many empirical studies have been carried out:

- *Showing how usability and user acceptance is influenced by new devices and novel interaction techniques:* [7], [16];
- *Showing that the perceived usability is impacted according to the kind of tasks performed* [10], [18] *and according to the context of use* (e.g. indoor x outdoor conditions, mobile applications) [2];
- *Trying to assess the accuracy of multimodal interaction for given tasks:* [3], [20].

Based on these findings, this paper reports research work aimed at exploring if two mice and a speech recognition system can be easily adopted by novice users. To investigate the usage of two mice and speech, in terms of usability we conducted a usability study including eye-tracking. We used eye-tracking to explore eye-gaze and hot spots when using various combinations of two mice and speech. We additionally explored possible interferences of multimodal interaction with cognitive load. This evaluation has been carried out on a multimodal interface of a safety critical interactive command and control system for satellite operations.

Evaluating Two Mice and Speech

We want to explore if two mice and speech (which a rarely used combination of input devices) can be naturally adopted by novice users. Our hypothesis is that two mice and speech are adopted quickly and new forms of interaction techniques are discovered by non-experienced users easily. We assume that speech becomes natural to interact with the system, even though combined with two mice as graphical input devices.

To prove our hypothesis that two-handed interaction and speech is adopted quite easily and quickly by users, we performed a usability study. Usability evaluation studies seem to be a preferred strategy for the evaluation of multimodal interfaces because it allows the investigation of how users adopt and interact with multimodal technology providing valuable information about the general usability and the user experience. Several types of user testing have been conducted for multimodal interfaces in both usability laboratories and in the field revealing user preferences for interaction modalities based on factors such as acceptance in different social contexts,

noisy and mobile environments [18]. An overview on new forms of usability evaluation methods for multimodal interfaces is presented in [4].

We used eye-tracking to gain some insight in usage patterns for two mice, and how speech influences the usage of the two mice. Additionally we explored eye-tracking as a mean to measure users' cognitive load. Using eye-tracking in usability studies combined with some forms of think-aloud is still discussed controversially in the literature. One of the reasons is that think-aloud might increase people self-consciousness and thus prevents them from behaving naturally [11]. Additionally, we will evaluate to which extend eye-tracking data adequately represents cognitive load [25].

Usability Evaluation Study

The goal of this usability evaluation study was to show that two handed interaction and speech can be naturally adopted by users, and that two-handed interaction or speech helps to reduce the cognitive load and necessary attention compared to normal mouse input. Next section presents informally the system on which the usability evaluation has been carried out. Then, the experiment (participants, setup and procedure) will be detailed. The section is concluded with the presentation of the results of the experiment and their interpretation according to the hypothesis presented above.

Informal description of the system used

This system has been built within a research project funded by CNES (French National Study Center on Space). The project was meant at defining a formal description technique capable of describing multimodal interfaces and multimodal interaction techniques in a complete and unambiguous way.

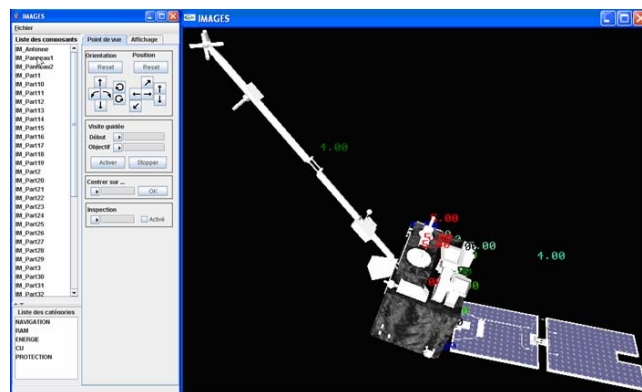


Fig 1: Screen shot of the system showing a 3D representation of Demeter satellite for the study of ionospheric disturbances.

This work defining extensions to an existing formal description technique called Interactive Cooperative Object (ICO) [1], which is based on object Petri nets and Objects. These extensions allowed us to describe unambiguous temporal behaviour of the speech-based interaction technique and how speech commands impact the temporal evolution of the graphical representation. The formal specification of the behavior of such multimodal interaction techniques is given in [22]. Beyond these software engineering aspects, the usability of the system was one of the main concern of the part of the project reported in this contribution. Indeed, such safety critical systems require certification by independent authorities prior to deployment and human factor certification is a critical aspect of that certification phase.

The system lies in the domain of Space Ground Systems that can be found in satellite control rooms [24]. The main goal of the system is to monitor a satellite making it possible for the controller to access any component of the satellite. This goal is separated into three main tasks: monitoring temperature, monitoring electricity consumption and locating hardware components in the satellite by moving the 3D image representation (see Figure1). In this application multimodal interaction takes place both while using special buttons (Figure 2) for navigating and changing the point of view of the 3D model, and while interacting with a range slider widget for selecting the temperature and the consumption (see Figure 4). The speech recognition system recognizes two words: “fast” and “slow”, which are used to increase or decrease the speed of visual display (movement of the 3D graphical representation of the satellite). The user can use speech interaction at any time she wants. Two-handed synergistic interaction is available on some interactors (see Figures 2 and 4) allowing users to manipulate both left and right input device at the same time.

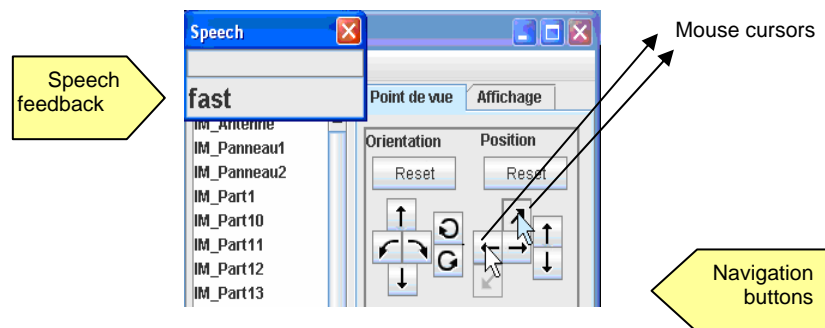


Fig 2. Possible multimodal interactions using 2 mice for tow-handed interaction (rotating the satellite) and speech input.

Participants

We asked 11 participants to take part in our usability evaluation study. Participants were recruited within the university including students and staff members. We selected four female and seven male participants, four of them with former experience on new forms of interaction techniques (pen input, special controllers and joy-sticks)

and seven with low or no experience of new forms of interaction techniques. We excluded people wearing eye-glasses in the selection process. Thus we could unfortunately not reach an equivalent distribution of male and females. 8 participants use the right hand for mouse input, 3 participants the left hand for mouse input. Age ranged from 21 to 26 years, with a mean of 22.9 years. All participants used the PC for more than 4 years (mean 10 years of usage), around 4.5 hours per day for work and 2.95 hours per day for work with the Internet at every day of the week (except one participant using the PC only 5 days/week). 9 out of 11 participants stated that they play games, about one hour per day (mean).

Set-up

For the usability evaluation study the participant was seated in front of a table, the eye-tracking device in front, and a laptop on a box in front of the participant. Two mice (left and right from a normal keyboard) were positioned on the table (see Figure 3). The participant used a headset around the neck for speech input. Additional to the eye-tracking, we used two cameras and a microphone for recording the scene and the user interaction. Additionally we captured the screen of the application.



Fig 3. Set-up for the usability study (including a third mouse for the leader of the experiment to start and stop applications)

Procedure

The participants got a short introduction about the usability study. We asked them to describe in the beginning of the task, what they intended to do, and if possible to think aloud during task completion. If participants naturally stopped the think-aloud because of the speech interaction, we did not force them to continue. We decided to use think-aloud as additional direct mean of providing useful insights in the usage of multimodal interaction, especially to offer insights in possible difficulties encountered while using the system. The data from eye-tracking thus has to be interpreted carefully, as think-aloud might interfere with cognitive load. However, the interaction technique was not influenced by the think-aloud, as speech was only used in two tasks of the study and simply consisted of two words.

Starting the usability study, participants answered some preliminary questions regarding the usage of other interaction techniques and games. Then the eye-tracking system was calibrated using 16 points on the screen to make calibration as good as possible. We used the Tobii x50 eye-tracker. None of the participants was wearing eye-glasses. One participant had lenses. The calibration was acceptable for all users.

After the calibration an introduction to the system and its basic functions was given (see Figure 1 for a screen-shot of the prototype in use). During the following training phase the user was introduced to the possible usage of the system using two mice and speech. The usability study included 6 tasks. After the completion of the tasks we made a short interview and participants filled out a questionnaire. Task 1 was to make only one component of the system visible. This task was a slight variation of one of the tasks during training. All participants solved this task, showing that they were able to perform basic interaction on the system.

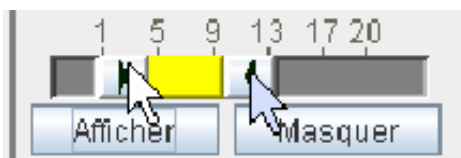


Fig 4. Two-Handed synergistic interaction on a temperature-slider.

In task 2 all components of the satellite with a temperature higher than 4 degrees °C should be made visible. This task is only solvable if the participant goes through the long list of all the components and requesting display of the temperature of each component. Task 2 should involve two-handed interaction. Task 3 was to locate a special component and to show all the components surrounding it (using two-handed interaction and speech) Task 4 was a variation of task 2. All components between 4 and 9 degrees should be shown and the number of components should be given. The main goal was to show that after short training (in task 2) two-handed input becomes more frequent in use. In task 5 participants were requested to show the antenna of the satellite on the full screen as quickly as possible, and task 6 to show the anchor of the antenna as quickly as possible too. Task 5 and 6 should involve speech input. To enforce speech input usage, we made the system considerably slower displaying graphical output.

After the performance of the tasks, an interview was carried out. During that interview the participants were asked to fill in two questionnaires. One questionnaire addressed system's usability (SUS). The other questionnaire was used to indicate how active or passive the system was during usage. This active/passive questionnaire was adopted from the AttrakDiff [14].

Results

Table 1 gives an overview of number of tasks successful completed, the tasks completed with help (help was defined for each task, for example during task 3, after 3 minutes all participants needed a hint to look more closely on the available functions in the interface) and tasks not completed at all or not completed because of reaching the time limit. In addition, the mean duration for performing the tasks is given. During the task we tried to give hints for function and usage of the program, to avoid, that functions not found interfere with the usage of the interaction technique. This explains the high percentage of tasks completed with help. Participants started to perceive tasks as easier during the test. We interpret the better ratings as a natural adoption of the interaction technique. The fact of easier usage was also reported by nearly all participants during the interview.

Task	Completed	Completed with help	Not Completed	Rating (Mean/SD)	Task duration (mn:sec)
T1	11	0	0	1.45 (0.52)	0:54
T2	0	5	6	3.91 (0.70)	5:55
T3	0	11	0	3.09 (0.83)	4:36
T4	5	3	3	2.73 (0.90)	3:37
T5	9	2	0	1.70 (1.06)(*)	1:39
T6	7	4	0	2.36 (1.03)	0:55

Table 1. Overview of completed tasks and their ratings.

(*) unfortunately one rating could not be reconstructed during transcription, the mean rating is based on N=10.

Two-Handed Interaction and Speech

The participants adopted the interaction with two mice quite naturally. Most of them started to use two mice interchangeably. They selected a specific component in the list using left mouse, and then performing an action using the right mouse. This behavior (influenced by the relative position on the GUI) was even shown by left-handed participants. Table 2 gives an overview of the interaction techniques used by the participants during the execution of the six tasks.

Task	One mouse	Two mice	Speech used additionally
T1	8	3	0
T2	4	7	0
T3	6	5	3
T4	1	10	0
T5	11	0	7
T6	10	1	9

Table 2. Interaction Techniques used during the performance of the six tasks: First column presents dominant vs non-dominant mouse usage (adding up to 11 participants); last column reports the number of time speech has been used additionally.

During task 1 the participants mainly used the mouse of their dominant hand to solve the task. Task 2 already shows an increase in two-handed interaction. The increased usage of the two-handed interaction-technique was supported by selecting a task including many repetitive elements (each component had to be selected on its own to solve the task). Depending on the strategy of the participant to solve task 3 they started to use two-handed interaction or speech to solve the task. Task 4 was closely related to task 2. The increase of using the two-handed interaction is quite impressive. This might be explained by the fact, that the attention needed for problem solving during the task is shifted, increasing the number of people using the two-handed interaction technique. The problem description of task 5 included the formulation “as fast as possible”. 7 out of 11 participants started to use speech to solve the task faster. The other participants used a specific function in the system to solve the task (just taking seconds to reach the solution). Task 6 again included the request to solve this task in minimal time. 9 out of 11 participants started to use voice to speed up the visualization process.

Noticeable during the whole usability study is that one participant started to move the hand away from the non-dominant mouse during several tasks (for example during task 4; this was the only user not using other interaction forms). On the other hand, two participants started to use the two-handed input not only interchangeably but really synergistically in trying to rotate the point-of-view of the satellite not only for one axis but two (task 6), or to use the slider for the temperature left and right at the same time (task 4). In one case the user started to use two mice in parallel (Figure 4), but completely forgot to use speech (while ALL other users used speech and only one mouse to complete task 6).

The standard usability scale (SUS) indicated that participants thought that the usability of the system could be considerably improved. SUS scored 59,6 on a scale from 0 to 100 indicating quite some potential to improve usability. Ratings on the adjective scales showed, that people thought that the system is best described using attributes like new, original and innovative. The highest ratings for negative attributes referred to boring, displeasing and clumsy.

Eye-Tracking

User intent and cognitive load is very difficult to assess during a normal usability study using thinking aloud [17]. To quantify cognitive load and user intent eye-tracking data is used as a possible objective measure. The three measures most commonly used are the number of fixations and the mean fixation duration, gazing time and saccade rates [25]. The number of fixations in each area of interest (AOI) is said to be negatively correlated with search efficiency and task efficiency [25]. Psychologists investigated eye-tracking as to provide insights on problem-solving, reasoning or search strategies. [17] gives an overview of measures used in eye-tracking studies and related evidence for corresponding cognitive processes. According to [12] the overall number of fixations indicates less efficient search. More fixations on a particular area indicate that this area is more noticeable, or more important to the viewer. Longer fixation duration can indicate difficulty in extracting information. We computed overall numbers of fixations, target fixations (fixations on targets divided by total number of fixations [13]), average fixation time and transitions between areas.

We defined several areas of interest, especially the navigation area, temperature slider, component list and satellite depiction. The overall number of fixations in the respective areas of interest defined, showed no significant trend of less efficient search in any of the areas. Same holds true for differences on number of fixations and target fixations. The average fixation time showed a tendency that areas related to navigation, selection of objects or temperature selection showed longer average fixation times compared with menu options or the satellite depiction. For example in task two the average fixation time in the area of interest “temperature slider” was 386 ms compared to 259 ms for the satellite.

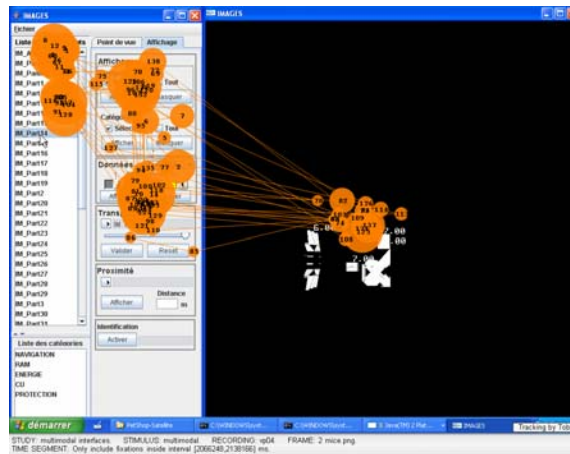


Fig. 5: Gaze plots when interacting with two mice, the left mouse was positioned on the selection of the item, the right mouse positioned on the menu entries and temperature slide

The extensive eye-tracking data analysis showed that average task duration was ways too high to allow meaningful and reliable interpretation of the data captured.

Long task durations and variations within the interaction technique usage might deeply influence the eye-tracking data. We thus decided to look more closely on the specific usages of interaction techniques during task execution.

To interpret the eye-tracking data in more detail, we classified all scenes of two-handed mouse interaction and usage of mouse and speech. The interaction technique was used mainly alternating (selection of a component in a list with the mouse in the non-dominant hand, performing the action with the mouse in the dominant hand). We additionally looked closer at participants trying to interact synchronously (using two mice at the same time, using mouse and speech at the same time). For synchronous usage in technical terms we could only classify a small number of cases. To investigate the fact of synchronous usage in detail we will implement a logging tool.

From the tasks described above we want to explain the selection of the temperature as one typical example. We classified the selection of the temperature based on the interaction technique used. As the selection of temperature could be used in two tasks we could classify 17 cases for usage with only one mouse and 3 cases of multimodal usage. The mean number of fixations while solving the task with one mouse was 41.45 fixations, when using two mice only 9.73 fixations occurred. More overall fixations normally indicate a less efficient search [25]. We interpret this gap as support for multimodal interaction, making solution of the task more efficient when using two mice.

Comparing the multimodal selection of components in a list with the single-modality selection, the same trend can be shown, that multimodal interaction shows less fixations compared to normal interaction with one mouse.

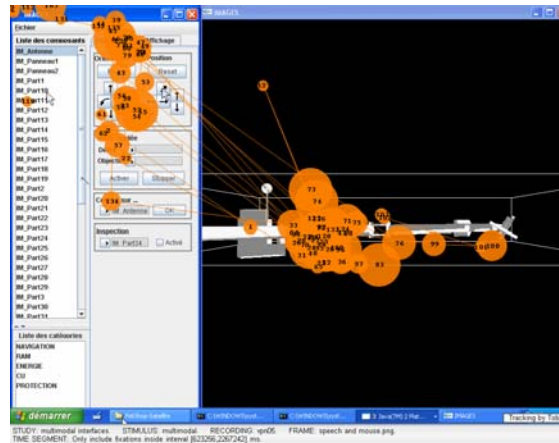


Fig. 6: Hot spots during speech interaction

We identified 14 instances of using *speech while rotating* the satellite. Unfortunately we could only identify one instance of synergistic usage of two mice to solve the same rotation task. Using speech the mean number of fixations in the areas of interest of the satellite (54.09 fixation in the AOI), the orientation control area (4.36 fixations) and the position control area (5.18 fixations) show that number of fixations are lower when using speech compared to multimodal interaction. The interaction with two mice showed more fixations in these areas (62 fixations for the

satellite but 8 for the orientation and 18 for the position). Speech reduces the number of fixations. Taking the overall fixation count, the number of fixations while using two mice is 643 fixations in the menu area compared to only 100 fixations in the same menu area while using speech. People tended to look at the satellite while using speech interaction (758 fixations compared to 332 fixations in the two mouse interaction condition).

Gaze-time shows the same trends as the number of fixation. The average fixation time was about the same for all interaction techniques (ranging from 300 to 500 ms). We currently explore these results in another usability evaluation study to measure time to first fixation more accurately. Observing the gaze during two-handed interaction shows that people are not looking at the mice while interacting. Participants tend to look at the interface elements. Figure 5 shows typical gaze plots when interacting with two mice.

Participants tended to look at the satellite while rotating with speech, ignoring additional input possibilities with the mouse. Figure 6 shows the hot spots during speech interaction.

Summary

The usage of three input devices (two mice and speech) is naturally adopted by users quite rapidly. Users reported in the final interview that they quickly adopted the interaction technique involving these devices, in terms of usage. They reported that continuous use made usage easier. The usability study showed that eye-tracking data can be used to investigate usage of multimodal interaction technique. People tend to ignore visual input possibilities while using speech, on the other hand they forget to use speech, when performing two-handed interactions. Eye-tracking data shows tendencies that speech reduces the number of fixations compared to two-handed interaction. Some cases indicated that eye-gaze is reduced when using two mice, compared to one mouse. Multimodal interaction with two mice shows advantages compared to one mouse, making task completion more efficient. Usage of two mice or usage of mouse and speech shows less cognitive load (in numbers of eye fixations) than when only one mouse is used.

Due to the rather limited number of users, results and interpretations presented above have to be handled carefully. Further experimental studies have to show if the data presented above can be generalized. Overall eye-tracking should be used to further explore multi-modal interaction in experiments to lay the ground for further usage of this method in the area of usability evaluation of safety-critical systems.

Conclusion and Future Work

This paper has presented the evaluation of a multimodal user interface featuring synergistic multimodal interaction techniques exploiting two mice and a speech recognition system. We have shown that multi-modal interaction with two mice and speech can be quickly adopted by novice users. Eye-tracking data indicated that

speech reduces the visual focus. Two-handed interaction does not result in additional eye-gaze compared with single mouse usage.

Some remaining issues are related to the low level interaction technique exploiting the input devices. Indeed, as stated in the introduction, the interaction technique has been fully modeled using the ICO formal description technique. While building the models we had to define aspects like acceleration of the mouse cursors (both for the dominant and non dominant hand) temporal interval for fusing the events produced by the various input devices, etc. Such temporal elements have been modeled using the notation and their actual value might have a significant impact on the usability results presented above. For instance, informal evaluation has shown that acceleration of mouse cursor should be smaller on the non dominant hand with respect to the dominant one.

To explore the eye-tracking data we want to connect the eye-tracking software with our modeling software PetShop (that allows editing, verifying and executing ICO models) to study first fixations and dynamic stimuli in more detail.

This work belongs to a more ambitious work aiming at providing methods, techniques and tools for the engineering of safety critical interactive systems. Beyond the low level interaction technique issues raised above we also consider higher level concerns like meaningful tasks carried out by users. This will be done by exploiting the formal models to support usability evaluation. Such support can be achieved by generating test cases from the models and thus providing pertinent scenarios to be evaluated via usability techniques as the ones presented in this paper. Such scenarios will be designed in such a way that they will cover all the low-level interaction aspect that have to be evaluated and thus enable designer to assess whether the interaction technique requires additional improvement or if the new design has a positive or negative impact on users' performance (with respect to previous designs).

Acknowledgements

This work is partly funded CNES (National Center of Spatial Studies in France) and the European commission via the Network of Excellence ResIST IST-4-026764-NOE (www.resist-noe.org).

REFERENCES

1. Bastide, R., Navarre, D., Palanque, P., Schyn, A., Dragicevic, P. Model-Based Approach for Real-Time Embedded Multimodal Systems in Military Aircrafts. *Proc. ICMI 2004*, ACM Press (2004), 243-250.
2. Baille, L., Schatz, R. (2005) Exploring Multimodality in the Laboratory and the Field. In: Lazzari, G. & Pianesi, P. (eds) ACM international conference on Multimodal Interfaces (ICMI'2005). ACM Press, New York, pp 100-107.
3. Balbo, S., Coutaz, J. Salber, D. (2003) Towards Automatic Evaluation of Multimodal User Interfaces. *Intelligent User Interfaces, Knowledge-Based Systems*, 6(4), pp. 267-274.

4. Bernhaupt, R., Palanque, P., Winkler, M. and Navarre, D. 2006. Supporting Usability Evaluation of Multimodal Safety Critical Interactive Applications using Dialogue and Interaction Models. (to appear) In book: *Maturing Usability: Quality in Software, Interaction and Value*. Law, E., et al. (eds.) Springer, to appear.
5. Bolt, R. A. (1980) Put-That-There: Voice and Gesture at the Graphics Interface. In: *Proceedings of the 7th International Conference on Computer Graphics and Interactive Techniques*. Seattle, USA, pp 262-270.
6. Bolt, R. E., Herranz, E. (1992) Two-handed gesture in multi-modal natural dialog. In: Mackinlay, J. & Green, M. (eds) *Symposium on User Interface Software and Technology (UIST'92)*. ACM Press, New York, United States, pp 7-14.
7. Bowman, D., Gabbard, J., Hix, D. (2002) A Survey of Usability Evaluation in Virtual Environments: Classification and Comparison of Methods. *Presence: Teleoperators and Virtual Environments*, v. 11 (4), pp. 404-424.
8. Buxton, W., Myers, B.A. (1986) A study in two-handed input. In: Mantei M, Orbeton P (eds) *ACM Conference on Human Factors in Computing Systems (CHI'86)*. ACM Press, Boston, Massachusetts, United States, pp 321-326.
9. Dillon, R.F., Edey, J.D., Tombaugh, J.W. (1990) Measuring the true cost of command selection: techniques and results. In: Chew JC, Whiteside J (eds.) *ACM Conference on Human Factors in Computing Systems (CHI'90)*, ACM Press, Seattle, Washington, United States, pp 19-25.
10. Dybkjær, L., Bernsen, N. O., Minker, W. (2004) New Challenges in Usability Evaluation - Beyond Task-Oriented Spoken Dialogue Systems. In: *Proceedings of ICSLP*, (vol. III), pp 2261-2264.
11. Godding, T. (2006) Eye tracking vs. Thinking aloud? Retrieved December 15th, 2006, from <http://staff.interesource.com/priority4/octover2006/eyetracking.htm>
12. Goldberg, H. J. and Kotval, X. P. (1999). Computer interface evaluation using eye movements: Methods and constructs. *International Journal of Industrial Ergonomics*, 24, 631-645.
13. Goldberg, H. J. and Wichansky, A. M. (2003) Eye tracking in usability evaluation: A practitioner's guide. In J. Hyöna, R. Radach and H. Deubel (Eds.) *The mind's eye: Cognitive and applied aspects of eye movements research*, 493 – 516. Amsterdam Elsevier.
14. Hassenzahl, M. (2003). The Thing and I: understanding the Relationship Between User and Product. In: Blythe, M., Overbeeke, C., Monk, A., and Wright, P. Eds.: *Funology: From Usability to Enjoyment*, Kluwer Academic Publishers, Dordrecht, 2003, p. 31-42.
15. Hinkley, K., Czerwinsky, M., and Sinclair, M. 1998. Interaction and Modeling Techniques for Desktop Two-Handed Input. *Proc. IUI 1998*, ACM Press (1998)49-59.
16. Hinckley, K., Pausch, R., Proffitt, D., Kassel, N. F. (1998) Two-handed virtual manipulation. *ACM Transactions on Computer-Human Interaction*, 5 (3), 1998, pp 260-302.

17. Jacob, R. J. K., and Karn, K. S Eye tracking in human-computer interaction and usability research: Ready to deliver the promises (Section commentary). In Hyona, J. The Mind's Eyes: Cognitive and Applied Aspects of Eye Movements. Oxford: Elsevier Science (2003).
18. Jöst, M., Haubler, J., Merdes, M., Malaka, R. (2005) Multimodal Interaction for pedestrians: an evaluation study. In: Riedl, J., Jameson, A., Billsus, D. & Lau, T. (eds) ACM International Conference on Intelligent User Interfaces (IUI'2005). ACM Press, San Diego, United State, pp 59-66.
19. Kabbash, P., Buxton, W., Sellen, A. (1994) Two-handed input in a compound task. In: Plaisant C (ed) ACM Conference on Human Factors in Computing System (CHI'94). ACM Press, Boston, Massachusetts, United States, pp 417-423.
20. Kaster, T., Pfeiffer, M., Bauckhage, C. (2003) Combining Speech and Haptics for Intuitive and Efficient Navigation through Image Database. In Oviatt, S. (eds) ACM International Conference on Multimodal interfaces (ICMI'2003). ACM Press, New York, pp. 180-187.
21. Kjeldskov, J., Stage, J. (2004) New Techniques for Usability Evaluation of mobile systems. *International Journal on Human-Computer Studies*, 60 (5), pp 599-220.
22. Navarre, D., Palanque, P., Dragicevic, P., Bastide, R.. An Approach Integrating two Complementary Model-based Environments for the Construction of Multimodal Interactive Applications. *Interacting with Computers*, 18, 5, (2006) 910-941.
23. Oviatt, S. Ten myths of multimodal interaction (1999). *Communications of the ACM*, Vol. 42 , No. 11, pp. 74 - 81
24. Palanque, P., Bernhaupt, R., Navarre, D., Ould, M., and Rubio, R. 2006. Model-based Measurement of the Usability of Multimodal Man-Machine Interfaces for Space Ground Segment Applications. *Proceedings of SpaceOps 2006*, AAAI Press (2006).
25. Poole, A. and Ball, L. J. Eye Tracking in Human-Computer Interaction and Usability Research: Current Status and Future Prospects. In Ghaoui, C. (Ed.). *Encyclopedia of Human Computer Interaction*. Idea Group (2004).
26. Reason J. (1990) *Human Error*. Cambridge University Press.
27. Salber, D. and Coutaz, J. A Wizard of Oz Platform for the Study of Multimodal Systems. *Interchi 93*, ACM Press (1993), 95-96.
28. Zhai, S., Barton, A.S., Selker, T. (1997) Improving browsing performance: a study of four input devices for scrolling and pointing tasks. In: Howard S, Hammond J & Lindgaard G (eds) *The IFIP Conference on Human-Computer Interaction (INTERACT'97)*. Chapman & Hall, Sydney, Australia, pp 286-292.