INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Project-Team MErLIn

Methods for Interactive Software Ergonomics

Lorraine - Rocquencourt

THEME 3A

Activity Report

2003

Table of contents

1. Team

Scientific Leaders

Dominique Scapin [DR, INRIA Rocquencourt, Scientific Leader]
Noëlle Carbonell [Pr., Henri Poincaré University, LORIA, Scientific co-leader]
Jean-Claude Sperandio [Pr., René Descartes University, Scientific co-leader]

Administrative assistants

Christiane Demars [AI, INRIA Rocquencourt, part-time until 31/03/2003]
Stephanie Aubin [TR,INRIA Rocquencourt, part-time from 1/04/2003]
Danielle Marchand [TR, Henri Poincaré University, LORIA, part-time]

University personnel

Christian Bastien [Assistant Pr., René Descartes University]
Antonio Capobianco [ATER, IUT Montbéliard, from 09/01/03]

Junior Technical Staff

Daniel Gepner [from 12/01/01 until 11/30/03]
Antonio Capobianco [from 11/01/02 until 06/30/03]

Ph.D. Students

Cédric Bach [INRIA grant, from 01/01/01, Metz University]
Noelly Grondin [INRIA grant, from 01/01/02, René Descartes University]
Suzanne Kieffer [INRIA-Région grant, from 10/01/01, Henri Poincaré University]
Vincent Lucquiaud [INRIA grant, from 01/01/02, René Descartes University]
Jérôme Simonin [DGA grant, from 10/01/03, Henri Poincaré University]
Charles Tison [INRIA grant, from 01/12/00, Henri Poincaré University]

Students interns

Marie-Laure Belin [Masters, René Descartes University, 8 months]
Cyrille Mennessier [Masters, René Descartes University, 8 months]
Émeric Wiederkehr [Masters, René Descartes University, 8 months]
Jérôme Simonin [DEA, Inst. Nat. Polytechnique de Lorraine, 04/01/03-08/31/03]

2. Overall Objectives

Key words: *ease of learning*, *electronic commerce*, *ergonomic criteria*, *ergonomic quality of interactive software*, *formal task description*, *HCI*, *hypermedia*, *information systems*, *interaction languages*, *interface design*, *interface evaluation*, *knowledge modeling*, *learning*, *methods*, *mixed-reality multimedia*, *multimodal interaction*, *online help*, *software ergonomics*, *standardization*, *task models*, *usability laboratory*, *user performance*, *user preference*, *user satisfaction*, *user testing*, *virtual environments*, *www*, *3D interaction*.

The goal of the MErLIn project is to contribute to the improvement of the Ergonomic Quality of Interactive Software. Two sub-goals contribute to that general goal :

- Study, through empirical studies [1], users' interactions with software-based systems in order to improve such systems. It is about increasing available knowledge about users' activities and cognitive characteristics as well as about the usability of software systems.
- Study and improve ergonomic design and evaluation methods, thereby contributing to the overall improvement of technical systems by providing software designers with sound methodological elements helping the incorporation of user-centered concerns within the design process life cycle. It is about increasing available knowledge on such processes, together with defining new methods or complementing existing ones.

[1] i.e., resulting from experience, through various methods, including controlled experiments.

Considering interactive computing systems for human use, i.e., ergonomic optimization of interactive software, requires to make progress both on fundamental knowledge and on methods in HCI (Human-Computer Interaction), and Ergonomics. The scientific contributions of the MErLIn project include scientific literature on users and task modeling, on empirical studies, on design and evaluation methods, on ergonomics recommendations, as well as software (e.g., mock-ups, test-prototypes, tools supporting design and evaluation methods). These various contributions are aimed at disseminating current ergonomic results, knowledge, and know-how to the national and international scientific community, but also to standards and to technology transfer through industrial contracts, collaborations and consulting activities.

Currently, the MErLIn project investigates two main research directions :

- The study, design, assessment and set-up of ergonomic methods for designing and evaluating interactive software. This corresponds to the need for integrating available ergonomic results into the computer systems life cycle. The main current topics relate to task-based and criteria-based methods.
- The study of usability issues raised by « new » computer applications: new user populations, new application domains, new forms of interaction (often new technology raises new usability problems). This corresponds to the need for acquiring novel ergonomic results on innovative computer systems, and to further increase current knowledge on usability. The main current topics relate to multimodal interactions, and virtual reality.

3. Scientific Foundations

The scientific domains characterizing the activities of the MErLIn project are essentially : Ergonomics, particulary Software Ergonomics and HCI. Four definitions apply to the research activities of the MErLIn project:

Ergonomics [2] (or Human Factors) is the scientific discipline concerned with the understanding of interactions among humans and other elements of a system, and the profession that applies theory, principles, data and methods to design in order to optimize human well being and overall system performance.

Ergonomics contributes to the design and evaluation of tasks, jobs, products, environments and systems in order to make them compatible with the needs, abilities and limitations of people.

Derived from the Greek ergon (work) and nomos (laws) to denote the science of work, ergonomics is a systems-oriented discipline which now extends across all aspects of human activity.

Domains of specialization within the discipline of ergonomics are broadly the following:

- Physical ergonomics is concerned with human anatomical, anthropometric, physiological and biomechanical characteristics as they relate to physical activity (Relevant topics include working postures, materials handling, repetitive movements, work related musculoskeletal disorders, workplace layout, safety and health).
- Cognitive ergonomics is concerned with mental processes, such as perception, memory, reasoning, and motor response,as they affect interactions among humans and other elements of a system (Relevant topics include mental workload, decision-making, skilled performance, human-computer interaction, human reliability, work stress and training as these may relate to human-system design).

[2] Definition from IEA (International Ergonomics Association) (http://www.iea.cc/ergonomics/)

   - Organizational ergonomics is concerned with the optimization of sociotechnical systems, including their organizational structures, policies, and processes (Relevant topics include communication, crew resource management, work design, design of working times, teamwork, participatory design, community ergonomics, cooperative work, new work paradigms, virtual organizations, telework, and quality management).

Software Ergonomics (or HCI Ergonomics) inherits from the main characteristics of ergonomics. It is a science that contributes to the knowledge necessary to software design, and more generally to computer-based environments, with the overall perspective of human security and well-being, but also with the perspective of effectiveness, efficiency, and productivity, for instance by facilitating users' tasks, limiting learning time, reducing errors and the cost of errors. Software Ergonomics focuses on the improvement of human-computer interactions mainly in terms of cognition, as the main human activity involved with software interactions is mental. However, as novel interaction techniques (e.g., multimodality) and novel environments (e.g., Virtual Reality) arise, some aspects of physical ergonomics are starting to be considered as well.

HCI [3] is also at the center of the MErLIn project research activities : "Human-Computer Interaction can be defined as the set of hardware, software, human and environmental elements that influence the efficiency of systems and products, both from a technological and a human point of view".
In addition, the MErLIn project aiming at the optimization of Software Ergonomics the *Ergonomic Quality of Interactive Software*, the following definition applies as well:

Ergonomic Quality of Interactive Software [4] covers all software aspects which have an influence on the users' task completion: it therefore covers usability in the widest sense, or ease of use, i.e., the extent to which the users can easily reach their interaction goals (which usually refers to presentation and dialogue aspects of the interaction (modes, interface features, dialogue, etc.)) but also what is sometimes called utility, i.e., the extent to which the users can reach their task goals (which usually refers to functional aspects of the interaction such as functions, objects, data, etc.). From a software engineering perspective (e.g., architecture models), it could be said that *Ergonomic Quality* covers not only the classical presentation, dialogue control and application interface aspects, but also some application kernel aspects: those that have an influence on the users reaching their goals.

3.1.1. National and International Context.

- At the international level, HCI ergonomics has been very dynamic for a number of years: young researchers, strong implication of industrial partners and academia, many job offers, major conferences, large audiences, many scientific journals. Most renowned universities and large software companies have HCI groups.
- In France research centers are still very few and are often mono-disciplinary ; the development of multidisciplinary (ergonomics and software engineering) research in HCI is quite recent.

3.1.2. Approach.

- The Human-Computer Interaction, particular type of Human-Machine Interaction, can be viewed along three complementary aspects : the human point of view, the computer point of view, and the interaction point of view. The MErLIn project considers all three points of view : the human and interaction aspects belong to ergonomics ; the computer and interaction aspects belong to the HCI part of computer science. Our research views computer systems (software, interfaces, environments) as a set of tools provided for human use.

[3] Definition from AFIHM (Association Francophone d'Interaction Homme-Machine) (http://www.afihm.org)
[4] Definition from: Scapin, D. L., and Bastien, J. M. C., (1997).Ergonomic criteria for evaluating the ergonomic quality of interactive systems.*Behaviour and Information Technology*, *16*, pp 220-231

- The MErLIn project uses methods from Ergonomics and from Computer Science, with a strong background and orientation in experimental approaches and methods (in the sense of experimental sciences, with hypotheses testing and proving, basic orientation of sciences such as medicine, biology, physics, etc.).
- The project contributes to the rationalization of the ergonomic methods, from experimental testing in the laboratory or field simulations, using performance data (e.g., learning time, task duration, usage frequency, error frequency, navigation types, level of recall, etc.), analysis of verbal protocols, analysis of preferences. The modeling activities are also centered on the production of computer models.
- The appropriateness and accuracy of such models compared to reality always goes through ergonomic evaluations.
- Research work starts usually from the observation of real tasks, on selected fields of activity, often in parallel with particular practical problems to be solved. Data gathering is based on activity and interaction analyses, case studies, critical incidents, automatic logs and records.

3.1.3. Focus.

Research work at MErLIn has also three additional characteristics.

- The focus is on methods dedicated to designers that are not necessarily skilled in ergonomics, even though such methods can also improve the activity of the ergonomists themselves. More specifically, the project deals with the integration of ergonomics approaches within the computer systems life cycle through sets of recommendations, methods, software support tools, and involvement in standardization, teaching and consulting.
- The focus is on users who are not computer-specialists. That user population is the major target of current software developments, whether it is the large public (e.g., interactive booths, electronic commerce, mobile systems) or professional experts in various domains (e.g., nuclear power plants, railways systems, textile design). A particular focus is on accessibility which promotes increased effectiveness, efficiency, and satisfaction for people who have a wide variety of capabilities and preferences.
- The focus is not only on "classical" work situations, but also on new computers uses, not yet all well defined, such as : consumer products (e.g., electronic commerce), information retrieval (e.g., tourism), mobility, etc.

4. Application Domains

This year, the main application domains have been : web sites (comparison of methods), design activities (3D design, initially textile industry), regulation activities (railway and subway systems). See section 6 for the specific results and the industrial partners involved.

6. New Results

6.1. Introduction

The research work conducted this year is presented along two main topics:

- Ergonomics methods for the evaluation and design of software interactions.
- Ergonomics of multimedia and multimodal interactions.

6.2. Ergonomic methods for the evaluation and design of software interactions

6.2.1. Adapting Ergonomic Criteria to Virtual Environments

Participants: C. Bach, D. L. Scapin.

The long term goal of this research is to elaborate an ergonomic inspection method for the design and the evaluation of 3D applications and Virtual Reality. To reach that goal, we have defined and tested a set of 20 Ergonomic Criteria (E. C.) derived from a compilation of ergonomic recommandations for virtual environments published last year. These criteria are described in a draft document (used in the experiments) through definitions, justifications, examples and counter-examples of recommandations.

Two experiments were conducted to validate the criteria. The requirement for the first experiment was to measure the intrinsic validity of the criteria within an assignment task. The requirement for the second experiment was to measure the contribution of the Ergonomic Criteria to the evaluation diagnosis during an inspection, compared to two other evaluation methods (a user test and a free inspection).

The first experiment was based on an assignment task consisting in the matching of concepts (the criteria) with their potential instances (usability problems). Ten experts in software usability (5.65 years of experience; SD = 5.4) but not experts in Human Virtual Environment Interaction (HVEI) had to assign the criteria to forty Virtual Environment (VE) usability problems. The results of this experiment show that the global performance of the subjects is: 68% of correct assignment to the main criteria (corresponding to theoretical assignments) and 59.5 % correct assignment to the elementary criteria. A deeper analysis of the results identified the confusions between the different criteria. These results are being used to improve the Ergonomic Criteria as modifications are made to the definitions, justifications and more illustrative examples and counter-examples are added. Two papers have been published on that experiment [2], [3].

The second experiment is based on the comparison of three ergonomic evaluation methods used to evaluate two VEs. The two VEs evaluated run on a classical computer, in order to limit familiarization biases. One of the VEs is a tourism application (an interactive visit of the Chamonix valley); the other one is an educational 3D application. The first method is a user test; 5 males and 5 females (age = 21.8; S. D. = 1.5) took part to the experiment. The subjects were familiarized with classical computers but not with 3D application. The subjects performed a set of 10 tasks with the tourism application, and then followed the instructions into the educational application. The subjects had 30 minutes per application to perform the different tasks. The results of this study are currently being analyzed. The second method is a free inspection; 10 students in software usability (DESS) performed an ergonomic inspection on the 2 VEs, just on the basis of their knowledge and expertise. They had to find as many usability problems as possible in each application. The subjects had 30 minutes per application to perform the task. The results of this study are currently being analyzed. The third situation is similar to the previous one, except that, in this case, the E. C. are used to guide the evaluation. Ten different students in software usability read the E. C. before performing the inspection, and they were able to refer to E. C. during the experiment. The results of this study are currently being analyzed.

The results from the three different experimental sessions are also being compared quantitatively (mainly types and number of usability problems diagnosed) and qualitatively (mainly comments on the problems, on the criteria, etc.) in order to compare their relative scope and efficiency in terms of diagnosis.

6.2.2. Identifying and specifying mental workload factors

Participants: N. Grondin, D. L. Scapin.

In the context of a PREDIT project with SNCF (French Railways) and RATP (Paris Subway System) which aims at identifying and predicting factors involved in mental workload from the identification of task characteristics, the study followed a two-step process: first analysis the literature, secondly collecting field data.

A review of the literature has been published [29]. The review is divided into two parts. First, a literature survey looks at the concept of mental workload. A number of definitions are provided about mental workload; also, we identified a number of general factors potentially leading to workload situations (related to individual, social and working conditions factors) and their underlying theories (e.g., “ Information Theory” , “Mulptiple

Resources Theory"). Also different workload measures are described (subjective measures, performance measures, physiological measures and analytic methods). The second part focuses on the identification of the workload factors. This review reveals that research seldom deals with task configurations that are problematic from the point of view of mental workload. However, we attempted to extract and classify workload factors from the literature along four types of variables: physical issues (e.g., visual information, environment), human issues (e.g., motivation, age, experience), time-related issues (e.g., biological rythms, work organization, time-related regulations); and tasks issues. The classification presents workload as being multidimensional and distinguishes between overload and underload characteristics.

The data gathering of field data consisted first in a pilot study at Rennes railway regulation center in order to test an initial interview methodology, then in conducting interviews at three other railway regulation centers (Château- Landon, Paris Gare du Nord and Versailles- Chantier). The semi-directed interviews were conducted as follows:

- Presentation of the study and of the interviewer;
- Presentation of the aim of the interview;
- Explanation of the concepts of "Mental Workload", Underload and Overload.

After a leading question asking: "what does mental workload mean for you in your activity?", we used a "Why and How" technique to elicit precise descriptions of workload cases experienced by the subjects, as well as their evaluation of these situations.

The data gathering is in progress (70 interviews so far) as well as, the analysis of the data. It appears quite difficult to find clear-cut cases of underload and overload, for two main reasons. One is that workload varies a lot from one person to a other, especially considering levels of experience. The other is that there are differences within the different sites observed: although the technology is the same (PRCI : " Poste tout Relais à Commande Informatique"), the work organisation and the role of the human operators differ from one site to the other. The initial analyses, however, seem to show that overload is not related mainly to the complexity of the tasks but to the simultaneity of severals tasks with the added factor of time constraints.

The selected instances of workload that have been identified and detailed further are being modeled with MAD (Méthode Analytique de Description des tâches, i.e., *Analytical Task Description Method)* in order to formalize task configurations in a systematic way.

6.2.3. MAD task modeling and tool

Participants: V. Lucquiaud, D. L. Scapin.

This year, a software tool was implemented using the JAVA language (16592 lines, 128 classes). This tool helps to describe an activity under user tasks, which are broken down into a hierarchy in order to build a task tree. The tasks are specified according to the MAD grammar. This grammar and its semantics were specified and enriched for the task attributes, the conditions (triggering pre-condition, execution pre-condition, stop pre-condition and post-condition), the definitions of abstract objects, objects instances and users.

This tool uses an ARCH architecture which has several modules. The updates and development are thus simplified.

- The Dialogue Component is an implementation of the H4 model. This controller is driven through four components: the tokens, the monitor, the questionnaires and the diagets. This is an abstracted module that can be used by another application. An XML file permits to describe the specific dialog for this tool. It contains the declaration of the functionalities of the software, their parameters and the associated name of the input command. This file read by the H4 parser generates, through automata creation, the dialog controller, the menus and the contextual menus.

- The Domain-Specific Component is an implementation of the EXPRESS language, where only our requirements were implemented. It uses the preliminary and differed connections (a part is then reusable for all EXPRESS specification) and a parser was built. A meta-model was designed for the representation of the "state of the world". Schemata are carried out for the whole description (task, state of the world, condition). This module controls the coherence of MAD descriptions.
- The Interaction Toolkit Component uses the Swing toolkit.
- The adapters (Presentation Component and Domain-Adaptor Component) allow the connections between the Interaction Component and the Domain-Specific Component. These are specific to this tool and are static. The calls to the adapters are driven by the Dialogue Component.

Today the tool allows a complete graphical description of a user activity: task description, state of the world, conditions, description of the users. Several types of format (spf, XML, HTML) are available in order to translate data sets for others tools of the same field (EUTERPE, AWG, ISOLDE, CTTE). Simulation remains to be implemented. Analysis of the user needs for user queries and their implementation also remains to be investigated. Usability must be improved and few service functionalities such as for example printing are still to be added in order to distibute and fully test the tool.

6.2.4. *Performance and preferences: assessing the role of learning in the light of eye tracking data*

Participants: J. M. C. Bastien, M. L. Belin, C. Mennessier.

Usability of an interactive software is related to several usability parameters that can be measured. Usability parameters fall into two broad categories: subjective user preference measures, assessing how much the users like the system or prefer one system over another, and objective performance measures, which measure how users use the interactive systems. Intuitively, a positive correlation between subjective preference and objective performance would be expected since people can be expected to prefer using interactive systems that help them accomplish their tasks rather than hinder them. A meta-analytic review published in 1994 by Nielsen and Levy indicated that preference and performance were positively associated in most cases. In some other cases, users did not prefer the system that would seem to be better based on the objective performance measures. For example, some users were found to consistently use a certain selection method in situations in which they had earlier been measured as being faster when using another method during practice trials. Some other results indicated that users preferred the slower of two data entry methods as long as it was not 20% slower than the faster method. More recent studies shed new light on the correlations between effectiveness, efficiency, and satisfaction. When asking users to accomplish complex tasks, the correlation between efficiency, as indicatedby task completion time, and effectiveness, as indicated by quality of solution, was negligible. Generally, the correlations among the usability aspects depend in a complex way on the application domain, the user's experience, and the use context. More generally, studies involving complex tasks account for only one or two aspects of usability. When these studies make claims concerning overall usability, they rely on risky assumptions about correlations between usability aspects.

To further investigate the relationship between performance and preference, two studies were conducted this year. These studies were conducted to pursue last year's study which was designed to assess the effect of the repetition of tasks on the relationship between performance and preference. This year, the study aimed at investigating in more details the experimental design and we used eye tracking technology to get data on visual exploration of the interface during task completion. Participants in both studies were asked to buy a ticket for a play on two different Web sites. The two sites allow the purchase of the same ticket but were different in terms of design, structure, and complexity. In both studies [30], each participant completed the task three times. In one study15 participants (n = 15) completed the task three times on one site and then three times one the other. The order of the sites were counterbalanced. In both studies, performance measures were: task completion time, number of Web pages explored, number of user actions (mouse clicks, page scrolls, etc.), number of eye fixations, mean fixation time, percent of fixation time, etc.

In the other study [28], each participant (n = 13) completed the task on one site and then on the other, and this was repeated three times. Again, the order of presentation of the sites was counterbalanced. Globally, the results of both studies indicate that users preferred the Web site for which the initial performance, that is the performance on the first trial, was the best and for which the increase in performance was the lowest. However, the analysis of individual performances indicated that users preferred the Web site for which the increase in performance was the lowest across the three trials but only in cases where the differences in performances on the two Web sites were significant.

The eye movement data indicate that the increase in performance was characterized by a decrease in the number of Web elements explored. However, the mean duration of the fixations did not change over trials. Visual exploration patterns were not analysed yet.

6.2.5. *Using conversation analysis, contingency analysis and pragmatics to analyse and evaluate natural language interfaces: an exploratory study*

Participants: J. M. C. Bastien, E. Wiederkehr.

Since people already have extensive communication skills through their own native or natural language many believe that Natural Language Interfaces (NLIs) can provide the most useful and efficient way for people to interact with computers.

The goal for most natural language systems is to provide an interface that minimizes training required for users. To most, this means a system that uses the words and syntax of a natural language. However, it has been shown that people converse differently with computers than they do when their counterpart is another person.

Laboratory studies have shown that users do relatively well if: they are knowledgeable about the domain or are given good feedback about the domain; are given language-specific training; and are given tasks that have been generated by the experimenters. Users perform poorly when training is absent, the domain knowledge is limited and the system is functionally impoverished.

To assess the ergonomic quality of a NLI, different approaches have been put forth. One of the most often used technique is user testing conducted with or without the Wizard-of-Oz simulations. Although this method allows one to identify exactly where in a dialogue sequence problems arise, it does not tell the evaluators how “ natural” is the interaction. To investigate this aspect, some work has been conducted. One approach has been to apply the Grice’s maxims, which allow to characterize naturally occurring dialogues, to evaluate whether NLI behaved accordingly.

We attempted to apply conversation analysis concepts and the three-term contingency of behaviour analysis which has been applied to the pragmatics and semantics of naturally occurring verbal interactions to a NLI. The conversation analysis concepts are divided into the *Pre-sequence* or preliminaries formalities, the *Main-sequence* and the *Post-sequence* or concluding formalities. These are themselves sub-divided into *turns* in which one person speaks and the other provide *response tokens* which are either *continuers* which maintain the speaker in turn, or *terminaters* which acknowledge the completion of the speaker’s turn and allow the listener to take over. An adjacency pair is defined as an utterance followed by an appropriate response token, or two successive turns in which the second is one the response demanded by the first one. Examples of such pairs are Request/Compliance and Question/Answer. Concepts borrowed from behaviour analysis are the *Antecedent (discriminative stimuli, establishing conditions)* conditions which call for some *Behaviour (mand, tact)* to be emitted and the actual or anticipated *Consequences* of so behaving.

In this study [34], 47 participants had to interact with a NLI every day. A total of 207 dialogues were recorded. Forty seven dialogue were chosen at random. The dialogues were then transcribed and analysed. The results indicate that the number of dialogue turns varies as a function of the request’s complexity. Results also indicate that, on the total number of dialogue turns produced by the users, more than 95% were single bivalent turns. The NLI produced essentially double bivalent turns. The results also indicate that one out of five adjacent pairs shows a violation of one of the Grice’s maxims. However no correlation were observed between a violation of a maxim and a failure of the dialogue. This should be further investigated.

This experiment has provided some interesting results. Although it was an exploratory study, the dialogue understanding we got when applying these analysis grids was interesting. These approaches should allow us to make interesting guidelines for the design and evaluation of NLI.

6.2.6. Contextual online help

Participants: A. Capobianco, N. Carbonell, J. Simonin.

This study, which began in 1999, addresses the issue of how to design online help that will really prove effective and, most of all, that will actually be used. Our approach is based on the assumption that online help systems implementing human experts' strategies will prove most effective.

We first elicited the contextual strategies used by human experts for helping novices in the general public to master the use of standard application software, from the analysis of a corpus of expert-novice help dialogues. Written transcripts of 15 dialogues between novice and expert users of MS Word were annotated and analysed using an ad hoc taxonomy for characterizing information exchanges. To validate the conclusions of this analysis, we then designed and performed an experimental ergonomic study. We analysed and compared the performances and subjective satisfaction of 18 novice users of Word who performed text processing tasks using two help systems successively (in counterbalanced order), a contextual one that emulated human expert strategies, and a non contextual standard one. Both help systems were simulated by a human operator (Wizard of Oz technique) who was assisted in his task by dedicated software.

This year, results of this experimental study have been presented at an international conference [25] and published in an international journal [18]. We have also designed and performed an additional experimental study that aims at comparing the respective efficiency of speech+graphics help messages and text+graphics standard ones. This study is described in subsection 6.3.2 which is focused on our work on multimodal human-computer interaction.

6.3. Ergonomics of multimedia and multimodal interactions

6.3.1. Multimodal input for user interaction with 3D design environments

Participants: C. Tison, D. L. Scapin, N. Carbonell.

Considering the recent developments of new graphical 3D techniques, Virtual Environments, etc., considering hopes and efforts that those developments stimulate, many questions arise, in particular from the ergonomic point of view, and more especially concerning the HCI techniques that would permit to interact with such environments. Our rationale is that in order to achieve 3D objects manipulation within such environments, classical interaction techniques and devices (mouse and keyboard) are not necessarily the most suitable.

Our ongoing research deals with user interaction with 3D Virtual Environments (VEs), for design activities. In this context, one must consider the type of activity to be supported, the range of available devices, techniques and metaphors, and the problems and constraints still inherent to these technologies from both human factors and application architectural points of view.

An initial survey of design activities in the field of clothing revealed a particular organisation of the design process and the possibility to extend the speech act theory to this kind of design activity where the user, in command of the system, may obtain an expected result by ordering it through a reduced set of input modalities. A survey of diverse available input techniques and devices was performed afterwards in order to identify the current problems encountered in such environments as well as the applicability of such an input multimodal combination. This initial litterature survey was published early this year [33].

Taking into account the absence of any tactilo-kinesthetic interaction, the most suitable input modalities in this context appear to be vocal commands synergistically combined to gaze or head direction and hand gestural deictic or mimetic movements. It is then possible, similarly to speech acts to define an input multimodal "expressive act" the user displays in order to obtain, from the system, an expected action on any selected component of the displayed 3D environment. An additional identification of user tasks within an immersive environment through such an interaction metaphor helped afterwards to design a User Tasks-Input modalities combination table. In order to experiment such a multimodal input, a 3D environment has been designed and implemented in JAVA so as to more precisely evaluate the user input multimodal expressions when requesting the system to perform some desired action within the environment.

6.3.2. Multimodal human-computer interaction: novel modalities

Participants: S. Kieffer, S. Simonin, N. Carbonell.

Spoken natural language may appeal to users from the general public, since it is the main modality used, together with pointing gestures or gaze, in face-to-face human communication. Speech+gestures based multimodality has been extensively studied, both from a software and an ergonomic point of view. However, speech+graphics as an output form of multimodality has raised fewer research studies, especially regarding the utility and usability of speech as a substitute modality for text in text+graphics presentations, or as supplementary modality to graphics. On the other hand, pointing hand gestures have the same expressive power as gaze as regards the selection of objects in very large displays (e.g., electronic blackboards, reality centres or caves, etc.), namely: both modalities can only specify directions in this context, if used spontaneously as in real life. Our current work on multimodality addresses the three following issues:

- How to design oral messages that help visual search in cluttered displays?
- How to design multimodal command languages that use information on gaze movements to disambiguate oral commands, especially those including deictic phrases?
- Are voice+graphics help messages more effective than standard text+graphics ones? Does this form of output multimodality actually improve the effectiveness and efficiency of online help?

Concerning oral support to visual search, we have completed the quantitative and qualitative analysis of a preliminary experimental study. This study aimed at determining whether oral information on the location of a visual target in the display of a complex scene could improve the efficiency (accuracy and selection time) of its identification. Targets were either familiar (visual presentation of the isolated target prior to scene display) or unfamiliar (oral characterisation of the target only, prior to scene display). Results have been presented at IHM'03 [27], and will be detailed in a chapter of a collective book edited by Kluwer (to be published next spring).

We are currently preparing a second experiment focused on the influence of the scene spatial organisation on the effectiveness of oral messages for familiar target detection, as a first contribution to the Micromegas Project (ACI 'Masses de données', 1st call). The development of the software necessary for automating scene construction, experimental setup and subject data (measures) recording, is completed. This experiment which will involve 20 voluntary computer science students will take place before the end of the year.

Concerning speech+gaze multimodal interaction, a software tool for recording and analysing speech+gaze interactions with virtual reality (VR) applications generated by the ORIS VR development software has been developed during this year, and demonstrated (at local scientific events). We have also completed the design of the protocol of an experimental study meant to collect data on gaze movements during unconstrained oral interaction with VR interactive applications. The user interface is partly simulated (Wizard of Oz technique), and the 3D interactive animated scenes have been created using ORIS. The multimodal data collected will be used both for eliciting users' gaze strategies, and for designing, implementing and testing robust algorithms intended for extracting information on the user's current goal from their gaze movements during oral human-computer interaction, and using this information for solving ambiguities and deictic phrases in their speech commands (multimodal fusion).

Regarding online help, we have performed an experimental study aimed at comparing the behaviours of 8 novice users who interacted with an animation software (Flash) using successively speech+graphics and text+graphics online help messages. First results are presented in Jérôme Simonin's DEA report [32]. We are currently implementing software tools for annotating and analysing the recorded interactions semi-automatically, including gaze fixations. Our aim is to elicit the possible influence of the modalities used on the usage of this information by novice users.

6.3.3. Universal Access to the Information Society

Participants: A. Capobianco, S. Kieffer, N. Carbonell.

The main objective of the thematic European network IS4ALL (5th PCRD, IST Programme, from 10/01/00 to 12/31/03) is to collect and disseminate best practice methods and techniques for implementing the concepts of "Universal Design" and "Universal Access" in Telemedicine and Health Care. Partners are currently completing the chapters of a handbook based on the experience gained in this area. We are in charge of the chapter on multimodal interface design.

8. Other Grants and Activities

8.1. National projects

- Participation to the PREDIT program (Ministry of Transportation) together with SNCF (French Railways) and RATP (Paris Subway System) : study of mental workload based on task characteristics (N. Grondin, V. Lucquiaud, D. L. Scapin).
- Participation to the 'Pôle Intelligence logicielle' of the 'Contrat de plan Etat-Région Lorraine': projects 'Assistance à l'apprentissage des langues' and 'Interactions multimodales' under the theme 'Téléopérations et assistants intelligents' (N. Carbonell).
- Participation to the RTP-CNRS 16 'Méthodes et Outils pour l'Interaction Homme-Machine'; AS Méthodes et outils pour les systèmes mixtes (C. Bach, D. L. Scapin).
- Participation to the RTP-CNRS 32 'Acceptabilité, ergonomie et usage des TIC' (N. Carbonell member of the 'Comité de pilotage').
- Participation to the Micromégas project, ACI 'Masses de Données', since July 2003 (N. Carbonell, D. L. Scapin).
- Participation to the project PrescInfo (computerization of the therapeutic prescription in the hospital) program RNTS in collaboration with EVALAB, LAMIH-Percotec, SIB de Rennes, C2A informatique and the Regional University Hospital of Lille (J. M. C. Bastien).
- Participation to the project CIINEMA ^{tm} project, a RIAM program (Research and innovation in audio-visual and multimedia), in collaboration with NovoCiné S.A., the ENST and Dorémi. (J. M. C. Bastien)
- Participation to the project « the discovery of the computer in infancy », in collaboration with the laboratory of Cognition and Development and Génération 5. (J. M. C. Bastien).

8.2. Projects and actions supported by the European Commission

- Thematic European IS4ALL Network (5th PCRD, IST Programme) on Universal Design and Universal Access in Health Telematics (N. Carbonell).

8.3. Networks and international working groups

- ERCIM Working Group 'UI4ALL' (N. Carbonell member of the Steering Committee).
- 6th PCRD, Strategic Objective 'Multimodal Interfaces', 1st Call, Bruxelles, May 12-20, 2003 and Luxembourg, June 10-11, 2003 (N. Carbonell expert).
- Member of the EFMI WG (European Federation for Medical Informatics) and IMIA WG on IT Evaluation in Health Care (J. M. C. Bastien).
- Member of the WWCS (Work With Computer Systems conference) Group (D. L. Scapin).

- AFNOR X3SE (Ergonomie des Logiciels Interactifs) ; (Chair : D.L. Scapin ; Author of the Introduction to the AFNOR's book and CD on Computer Systems Ergonomics Standards).
- ISO/TC 159/SC4/WG5 (Software ergonomics and human-computer dialogues) (D. L. Scapin expert).
- ISO/TC 159/SC4/WG6 (Human-centred design processes for interactive systems) (D. L. Scapin expert).
- CEN/TC 122/WG 5 (Software ergonomics and human-computer dialogues) (D. L. Scapin expert).

9. Dissemination

9.1. Animation of the scientific community

9.1.1. Organization of scientific events

- EPIQUE'2003, Journées d'Etude en Psychologie Ergonomique, Boulogne-Billancourt, 2-3 octobre 2003 (J.M.C. Bastien, chair and local organizer).

9.1.2. Journals Editorial Boards

- Behaviour and Information Technology. (Member of the Editing Committee: D. L. Scapin).
- Interacting with Computers, (Reviews: D.L. Scapin).
- International Journal of Cognitive Ergonomics. (Member of the Editing Committee: J.-C. Sperandio).
- International Journal of Human-Computer Interaction. (Member of the Editing Committee: J.-C. Sperandio).
- International Journal of Human-Computer Studies. (Reviews: D.L. Scapin).
- International Journal of Universal Access in the Information Society (UAIS). (Members of the Editorial Board: N. Carbonell, D. L. Scapin).
- Le Travail Humain (Membre du Comité de Rédaction: J.-C. Sperandio ; Membres du Comité de Consultants: N. Carbonell, D. L. Scapin.
- Revue d'Interaction Homme-Machine. (Membres du Comité de Rédaction : J. M. C. Bastien ; D. L. Scapin).
- Revue Information, interaction, intelligence (13.) (Membre du Comité de Rédaction: N. Carbonell).

9.1.3. Conference Programme Committees

- ACM Conference on Human Factors in Computing Systems (CHI 2003), Fort Lauderdale, USA (Reviews, D. L. Scapin).
- ACM Conference on Universal Usability (CUU'03), Vancouver (N. Carbonell. Programme Committee member).
- Cide.6 : 6ème Colloque International sur le Document Electronique, Caen, 24-26 novembre 2003 (J.M.C. Bastien, Program committe member).
- EPIQUE'2003, Journees d'Etude en Psychologie Ergonomique, Boulogne-Billancourt, 2-3 octobre 2003 (J.M.C. Bastien, D. L. Scapin, membres du Comité Scientifique).
- INTERACT'03, 1-5 Sept. 2003, Zurich, Switzerland. (Member program Committee & Member Jury Doctoral Consortium (D. L. Scapin).

- International Conference on User Modelling (UM'03), Pittsburg (N. Carbonell, Programme Committee member).
- International Workshop "Exploring the design and engineering of Mixed Reality Systems", joined with ACM IUI,2004 & CADUI,2004, Funchal, January 13-16, 2004 : (Member Program Committee D. L. Scapin).
- 2nd International Conference on Universal Access in Human-Computer Interaction (UAHCI'03), Crete:(N. Carbonell, D. L. Scapin Programme Committee members).
- 5th ACM International Conference on Multimodal Interfaces, Vancouver (N. Carbonell, Programme Committee member).
- 15ème Conférence Francophone sur l'Interaction Homme-Machine (IHM'03), Caen: (N. Carbonell Co-chair full papers, J. M. C. Bastien (Co-chair informal papers), D. L. Scapin (reviews)).
- OZCHI 2003, Brisbane, Australia, November 26-28 2003. (Member Program Committee, D. L. Scapin).

9.1.4. Scientific Societies Membership

- ACM, Association of Computing Machinery ; Special Interest Group on Computer-Human Interaction (SIGCHI). Members : J. M. C. Bastien, N. Carbonell, D. L. Scapin.
- AFIA (Association Française d'Intelligence Artificielle). Membre : N. Carbonell.
- AFIHM (Association Francophone d'Interaction Homme-Machine). Membres : D. L. Scapin (membre de la Commission de Pilotage des Manifestations Scientifiques) ; J. M. C. Bastien (membre du C.A.) ; C. Bach, N. Carbonell , V. Lucquiaud.
- APS (American Psychological Society) : Member : J.M.C. Bastien.
- ARCo (Association pour la Recherche Cognitive). Membre : N. Carbonell.
- HFES (Human Factors and Ergonomics Society). Members : J. M. C. Bastien et D. L. Scapin.; HFES-CSTG (Computers Systems Interest Group). Members : J. M. C. Bastien et D. L. Scapin.
- IEEE (Institute of Electrical and Electronics Engineers). Member : N. Carbonell.
- ISCA (International Speech Communication Association). Member : N. Carbonell.
- SABA (Society for the Advancement of Behavior Analysis). Member : J. M. C. Bastien.
- SELF (Société d'Ergonomie de Langue Française). Members : J. M. C. Bastien; D. L. Scapin; J.-C. Sperandio.
- SFP (Société Française de Psychologie). Member : J.M.C. Bastien.

9.1.5. Ph.Ds and HdRs examining boards

- Marjolaine VILLEY-MIGRAINE : *Multimedia et carto-geographie: Ergonomie des interfaces de navigation hypermédia dans les systèmes documentaires*. Doctorat en Sciences de l'information Spécialité “images, médias et électronique de la communication”, 3 Juin 2003, Université Paris 2 (Panthéon-Assas). D. L. Scapin examinateur.

9.2. Teaching

- DEA d'Informatique, Ecole Doctorale IAEM, N. Carbonell, responsable de la filière 'Perception, raisonnement, traitement automatique des langues', et membre permanent du jury de soutenance.
- DESS Information Scientifique et Technique, Universités de Nancy: N. Carbonell (70h)
- DEA 'Chimie Informatique et Théorique', DEA national: N. Carbonell (15h).
- DEA MIASH, Université René Descartes, J.M.C. Bastien (12h).
- DESS Ergonomie, Université René Descartes, J.M.C. Bastien (56h cours, 12h TD).
- DESS Ingénierie de la santé, Université René Descartes, J.M.C. Bastien (12h).
- DU Bases Facteurs Humains pour la Conception de Systèmes Homme-Machine en Aéronautique, Université René Descartes, J.M.C. Bastien (3h).
- Maîtrise Psychologie, Université René Descartes, J.M.C. Bastien, (12h).
- DEUG Psychologie, Université René Descartes, J.M.C. Bastien (4 h); N. Grondin (20h Cours ergonomie ; 14 h TD ergonomie ; 12 heures TD statistiques).

9.3. Participation to conferences, workshop, invited talks

- ACM Conference on Human Factors in Computing Systems (CHI 2003), April 5-10, Fort Lauderdale, USA (D. L. Scapin).
- HCI International 2003, Crete, June 22-27, 2003 (A. Capobianco).
- IFIP TC13 International Conference on Human-Computer Interaction (INTERACT'03), Zurich, September 1-5, 2003: (C. Bach, A. Capobianco, D. L. Scapin).
- 5th ACM International Conference on Multimodal Interfaces, Vancouver, November 5-7 (N. Carbonell).
- ACM Conference on Universal Usability (CUU'03), Vancouver, Novembre 10-11 (N. Carbonell).
- Second Working Meeting of the SIG-SEM ACL Working Group on the Representation of Multimodal Semantic Information, Nancy, September 22-23, 2003 (N. Carbonell, invited talk).
- 15ème Conférence Francophone sur l'Interaction Homme-Machine (IHM'03), Caen (C. Bach, N. Carbonell, S. Kieffer., D. L. Scapin).
- École d'été du CNRS, Théories et méthodes pour la conception, l'évaluation et l'usage d'environnements informatiques pour l'apprentissage humain, Autrans, 4-9 juillet 2003 (J.M.C. Bastien, invited speaker).
- CETSISEEA, 4ème Colloque sur l'Enseignement des Technologies et des Sciences de l'Information et des Systèmes, Toulouse (J.M.C. Bastien, invited speaker).
- Seminaire MErLIn, Rocquencourt; 22 Sept. 2003. Exposé invité de Frank Vetere (University of Melbourne, Australia).
- Seminaire MErLIn, Rocquencourt ; 29 Sept. 2003. Exposés invités de D. Trevisan et J. Vanderdonckt (Université de Louvain, Belgique); Exposés de C. Bach, C. Tison.

10. Bibliography

Major publications by the team in recent years

[1] R. AMALBERTI, N. CARBONELL, P. FALZON. *User representations of computer systems in human-computer speech interaction.* in « International Journal of Man-Machine Studies », volume 38, 1993, pages 547-566.

[2] J. M. C. BASTIEN, D. L. SCAPIN. *A validation of ergonomic criteria for the evaluation of human-computer interfaces.* in « International Journal of Human-Computer Interaction », volume 4, 1992, pages 183-196.

[3] J. M. C. BASTIEN, D. L. SCAPIN. *Evaluating a user interface with ergonomic criteria.* in « International Journal of Human-Computer Interaction », volume 7, 1995, pages 105-121.

[4] N. CARBONELL. *Acquisition et formalisation de stratégies d'aide à la mise en œuvre de logiciels grand public.* Technical report, Inria, Rocquencourt, France, février, 1997, Contribution au Rapport final du Projet de Recherche "Dialogue et coopération" soutenu par le PRC 'Sciences Cognitives'.

[5] A. POLLIER. *Évaluation d'une interface par des ergonomes : diagnostics et stratégies.* in « Le Travail Humain », volume 55, 1992, pages 71-96.

[6] D. L. SCAPIN, J. M. C. BASTIEN. *Ergonomic criteria for evaluating the ergonomic quality of interactive systems.* in « Behaviour and Information Technology », volume 16, 1997, pages 220-231.

[7] D. L. SCAPIN, C. PIERRET-GOLBREICH. *Towards a method for task description.* in « Proceedings of Working With Display Units », Elsevier, L. BERLINGUET, D. BERTHELETTE, editors, pages 371-380, Amsterdam, The Netherlands, 1990.

[8] D. L. SCAPIN. *Situation et perspectives en ergonomie du logiciel.* J.-C. SPERANDIO, editor, in « L'Ergonomie dans la conception des projets informatiques », Octares, Toulouse, France, 1993, pages 7-68.

[9] D. L. SCAPIN. *The need for a Psycho-engineering approach to HCI.* in « II Congresso Latino Americano e VI Seminario Brasileiro de Ergonomia, Florianopolis, Brasil », pages 3-22, 10-13 octobre, 1993.

[10] S. SEBILLOTTE, D. L. SCAPIN. *From users' task knowledge to high level interface specification.* in « International Journal of Human-Computer Interaction », volume 6, 1994, pages 1-15.

[11] S. SEBILLOTTE. *Décrire des tâches selon les objectifs des opérateurs. De l'interview à la formalisation.* in « Le Travail Humain », volume 54, 1991, pages 193-223.

[12] J.-C. SPERANDIO. *L'ergonomie cognitive.* in « Psychologie Française », volume 40, 1995, Numéro spécial.

[13] J.-C. SPERANDIO. *Apports de l'ergonomie à la conception et aux choix des aides techniques pour personnes handicapées.* in « Performances Humaines et Techniques », avril, 1996, pages 39-43, Hors série : Situations de handicaps : Nouvelles approches ergonomiques.

[14] J.-C. SPERANDIO. *L'apport de la psychologie du travail.* P. CAZAMIAN, F. HUBAULT, editors, in « Traité d'ergonomie », Octares, Toulouse, France, 1996.

[15] J.-C. SPERANDIO. *L'ergonomie face aux changements technologiques et organisationnels du travail humain.* Octares, Toulouse, France, 1996.

[16] S. ROBBE. *Étude ergonomique de contraintes d'expression orales et gestuelles dans un environnement multimodal d'interaction homme-machine.* Ph. D. Thesis, Université Henri Poincaré, décembre, 1998.

Articles in referred journals and book chapters

[17] J. M. C. BASTIEN, D. L. SCAPIN. *La conception de logiciels interactifs centrée sur l'utilisateur : étapes et méthodes.* in « Traité d'ergonomie », 2003.

[18] A. CAPOBIANCO, N. CARBONELL. *Online help for the general public: specific design issues and recommendations.* in « International Journal Universal Access in the Information Society, Special issue on countering design exclusion », October, 2003, pages 265-279.

Publications in Conferences and Workshops

[19] C. BACH, D. L. SCAPIN. *Adaptation des Critères Ergonomiques aux Interactions Homme Environment Virtuel.* in « 15ème Conférence Francophone sur l'Interaction Homme-Machine 2003 - IHM'03 », New York: ACM Press, pages 24-31, Caen, France, Novembre, 2003.

[20] C. BACH, D. L. SCAPIN. *Adaptation of Ergonomic Criteria to Human Virtual Environment.* in « Ninth IFIP TC13 International Conference on Human-Computer Interaction - INTERACT'03 », Amsterdam: IOS Press, pages 880-883, Zurich, Switzerland, September, 2003.

[21] C. BACH, D. L. SCAPIN. *Obstacles and Perspectives for Evaluating Mixed Reality Systems Usability.* in « Submitted to the Workshop MIXER "Exploring the Design and Engineering of MR system", IUI-CADUI 2004 », Funchal, Portugal, January, 2004.

[22] J. M. C. BASTIEN. *Approches et méthodes ergonomiques pour la conception de systèmes interactifs.* in « École d'été du CNRS : Théories et méthodes pour la conception, l'évaluation et l'usage d'environnements informatiques pour l'apprentissage humain », Autrans, France, 4-9 juillet, 2003.

[23] J. M. C. BASTIEN. *Les nouvelles technologies de l'information et de la communication (NTIC) : nos méthodes d'analyse du travail sont-elles toujours adaptées ?.* in « Actes du XXXVIIIème Congrès de la SELF », G. VALLERY, R. A. (EDS.), editors, pages 85-86, Paris, France, 24-26 septembre, 2003, Paris, Société d'Ergonomie de Langue Française.

[24] J. M. C. BASTIEN. *Étapes et méthodes clés pour une conception ergonomique des logiciels interactifs.* in « CETSIS-EEA, 4ème Colloque sur l'Enseignement des Technologies et des Sciences de l'Information et des Systèmes », Toulouse, France, 12-14 novembre, 2003.

[25] A. CAPOBIANCO. *Questioning the effectiveness of contextual online help: some alternative propositions.* in « Ninth IFIP TC13 International Conference on Human-Computer Interaction - INTERACT'03 », Amsterdam:

IOS Press, pages 65-72, September, 2003.

[26] N. CARBONELL. *Multimodal HCI Corpora: a tentative listing of the specific information content of annotations.* in « Second Working Meeting of the SIG-SEM ACL Working Group on the Representation of Multimodal Semantic Information », pages 6, Nancy, France, September, 2003, invited talk.

[27] S. KIEFFER, N. CARBONELL. *Assistance orale à la recherche visuelle dans des affichages complexes.* in « 15ème Conférence Francophone sur l'Interaction Homme Machine 2003, IHM'03 », New York: ACM Press, pages 126-133, Caen, France, Novembre, 2003.

Internal Reports

[28] M.-L. BELIN. *Utilisabilité des sites Web: liens entre performance, apprentissage et préférence, et apport de l'investigation oculométrique.* Mémoire de Maîtrise, Université René Descartes, Paris, 2003.

[29] N. GRONDIN, D. L. SCAPIN. *Revue de Questions sur la Charge Mentale de Travail : Notions, Mesures et Prédiction.* Rapport de contrat INRIA, Convention de Recherche INRIA- SNCF- RATP, Projet PREDIT, INRIA, Rocquencourt, Octobre, 2003.

[30] C. MENNESSIER. *Performance, apprentissage et exploration visuelle : une approche des déterminants de la préférence dans l'évaluation comparative des sites Web.* Mémoire de Maîtrise, Université René Descartes, Paris, 2003.

[31] D. L. SCAPIN. *Rapport Final Convention de Recherche INRIA-LECTRA.* Rapport de contrat COMEDIA, INRIA, Rocquencourt, janvier, 2003.

[32] J. SIMONIN. *Présentation de documents procéduraux : Apport de la multimodalité à la mise en oeuvre de consignes.* Rapport de DEA, stage LORIA, INRIA, Nancy, 2003.

[33] C. TISON, D. L. SCAPIN, N. CARBONELL. *Techniques et modalités d'interaction homme-machine avec des environnements et des représentations d'objets 3D : revues de questions.* Rapport de contrat COMEDIA, INRIA, Rocquencourt, janvier, 2003.

[34] É. WIEDERKHIER. *L'analyse d'interaction vocale homme-machine : étude comparative exploratoire de l'analyse des contingences et de l'analyse conversationnelle.* Mémoire de Maîtrise, Université René Descartes, Paris, 2003.