27

MULTIMODAL USER INTERFACE SYSTEM FOR BLIND AND "VISUALLY OCCUPIED" USERS: ERGONOMIC EVALUATION OF THE HAPTIC AND AUDITIVE DIMENSIONS.

Aude Dufresne, Odile Martial** and Christophe Ramstein***

* Université de Montréal, C.P.6128, Succ. Centre-Ville, Montréal, Qc. Canada, H3C 3J7
E-Mail: Dufresne@iro.umontreal.ca
** Centre d'Innovation des Technologies de l'Information

KEYWORDS: Haptic and auditory interfaces, multimodal interfaces, force-feedback, physical model, user interface development system, blind people.

ABSTRACT: An ergonomical evaluation of a multimodal windows-oriented interface that gives blind and "visually occupied" persons access to graphical user interfaces. Physical models are used to represent interface features providing haptic informations, thanks to a force-feedback device, and non-speech audio. In the absence of vision, both auditory and haptic modalities are used, to perform manipulation tasks. Three conditions were compared: sound feedback alone, force feedback alone and bimodal feedback. Measures of usability were timing, error rates with subjective satisfaction. Results from the experiment show that multimodality was associated with better performance for blind and sighted subjects and that it is ranked as the best interface.

INTRODUCTION

As Graphical User Interfaces (GUI) are becoming a standard in Human-Computer interaction, it becomes important to develop alternate modes of interaction for blinds or "sight occupied" users. Since sight is the most rapid and precise way to convey graphical information (Welch, & Warren, 1987), it appears essential to replace it not by the use of one method but many. To test this hypothesis, this research investigates the use of physical sounds and force feedback devices as a substitute for GUI for blind and "visually occupied" users". A multimodal user interface system (MUIS) is used to develop and manage the multimodal feedback interface.

RESEARCH CONTEXT

Much research is currently being done in the field of multimodal interfaces to enrich visual information with audio and tactile feedback (Blattner, Sumikawa, & Greenberg, 1989). While some ergonomic evaluations proved that adding auditory feedback to visual information (ie scroll bar) increases appreciation, performance and control (Brewster, Wright, & Edwards, 1994), and others noticed the same for a tactile complement to visual information (Heller, 1982), there are no systems or ergonomic evaluations that provide concrete support for combining auditory and haptic feedback,[1] or blind and for "visually occupied" persons.

For blind people the question needs to be solved urgently (Lazzaro, 1991) and some alternatives are currently being explored to enrich the sensorial feedback given to them: mouse on a graphic tablet with speech and sound output (Martial, & Dufresne, 1993), tactile or Braille cells on a mouse (Göbel, 1995), the GUIB project (Weber, 1993) using vertical and horizontal braille displays, speech and sound output, a touch sensitive tablet and keys for exploration. Commercial products based on command language and voice synthesis feedback on the screen are already providing blind users some access to GUI. Unfortunately they are incomplete transpositions of the direct manipulation environment of GUI: many do not have pointing devices, and most graphical characters are not translated or are described verbally, so the real-world metaphor is lost (Martial, & Dufresne, 1994). Blind users find using a computer a very difficult task and

[1] The word haptic relates to two senses: the tactile sense that gives an awareness to stimuli on the surface of the body, the kinesthetic sense that provides information on body position and movement (Gibson, 1966; Appelle, 1991).

K. Nordby et al. (eds.), *Human-Computer Interaction*

demanding with regard to concentration and memory. Thus it is important to give them more feedback on their activity, to present them with an environment with more ways of interacting

OBJECTIVES

This research is intended to verify the usability of a force-feedback device and audio feedback as a substitute for vision to facilitate direct manipulation of GUIs. The general hypothesis is that in the absence of vision, both feedbacks can be used, but that bimodality would lead to better performance and satisfaction than either modality alone.

SOFTWARE ENVIRONMENT

The test application was developed in a multimodal user interface system (MUIS) enabling the development and management of an interface that provides every object (windows, icons and menus) with a visual, auditory and/or haptic quality. The original MUIS elements are the force feedback and multimodal representation based on physical modelling (Ramstein, to be published).

The Pantograph: a force-feedback device

A force-feedback device is both a sensor and a motor. It filters hand movements as well as allowing users to feel the force feedback and pressure. Examples of force feedback are given in (Cadoz, Lizowski, & Florens, 1990; Minski, Ouh-Young, Steele, & Brooks, 1990). The MUIS consists of a force-feedback device with three degrees of freedom (DF) called the Pantograph. The device, both sensor and motor, was designed for human-computer interaction based on iconic representation (Ramstein, & Hayward, 1994). It is ergonomic and inexpensive. The Pantograph's basic mechanical structure allows the user to move a point in a 10 cm by 16 cm space. Conversely, two powerful and accurate motors synthesize 20 Newtons of vertical force in this space (see Hayward, Choksi, Lanvin, & Ramstein, 1994).

The Pantograph is a haptic, tactile and kinesthetic device because the forces returned stimulate the skin sensors of the hand, as well as those of the muscles, tendons and joints. However, the Pantograph cannot stimulate the skin like a standard braille display. To perceive a form, a user is obliged to cover its whole surface. It makes it possible to perceive form, texture and distance in three dimensions.

Multimodal representation

The representation of objects (window, icons and menus) is based on physical modelling inspired by research conducted in the context of artistic creation (Cadoz, et al., 1990). We feel that the physical model provides a type of conception that fits well with multimodal feedback, providing an intrinsic guarantee of the coherence of these stimuli. A physical object is in fact an object that is usually perceived, when carrying out an action, through its various sensory qualities. Theses are linked by a causal relation to become coherent. Redundancy and complementarity are intrinsically defined.

The physical models chosen represent the morphology of the objects. The interface must be given material aspects. For example, the icon buttons (closing box, resizing box) are modelled like real mechanical buttons with perceptible edges and a pressing or releasing action depending on the shapes and behaviors of the different icon buttons that are used in standard interfaces. The same is true for the menu icons, items and headers. The size and position parameters are similar to their graphic counterparts, the only variation being the viscosity (conversely to their surface) and height (or depth). The window frame is defined as a rectangular gutter whose shape is demarcated by an increased force feedback. The mouse pointer, controlled by the Pantograph, is comparable to the weight of a submissive mass (orthogonal to the screen surface) that can be moved about the surface (visible) of the objects. The force is constantly calculated and fed back to the user through the Pantograph, resulting in the following forces: weight, friction and object reaction applied to the pointer. The user senses the morphology of the objects in a form of 3D: visual and haptic.

From the auditory aspect, the interface is considered to be a complex musical instrument that has vibrant structures and a stimulative (the pointer). In the physical modelling suggested by Cadoz et al., a vibrant structure can be analyzed as the composition of a collection of masses interconnected by visco-elastic elements. Consequently, when the pointer touches an object, it stimulates its vibrant structure, which oscillates according to proper modes and behavior, thereby producing audible sounds.

PRE-EXPERIMENT OF THE PANTOGRAPH

The Pantograph was pre-experimented using blinds and sighted subjects without vision (Ramstein, Martial, & Dufresne, 1994). The model and the values of the parameters corresponding to the various element of the GUI were adjusted, taking into account satisfaction and performance. It was tested with a tablet of three different sizes.

The results showed that the larger tablet lead to better performance ($p<.05$). They also showed that fatigue was an important component using an haptic device, since performance was lower at post-test ($p<.001$). The fatigue may be explained by the fact that the experimentation was long (mean = 1h30) and included adjustment of the physical parameters, requiring that the users made a major effort to adapt.

Finally in the pre-experimentation, the blinds performed significantly better than the sighted ($p<.05$), and there was an interaction effect, their performance seeming not to decrease as much with time.

HYPOTHESIS

In this research, in order to find a usable interface for the blinds, three conditions were tested offering the user different modalities of interaction of feedback with the GUI: audio (sound feedback), haptic (force feedback) or bimodal (sound and force feedback).

Since the preliminary experiment with the haptic only condition showed that blind subjects performed better in general, the first hypothesis was that they would perform generally better than sighted subjects for all three conditions.

The second hypothesis was to test the usability of the three interaction conditions. It was postulated that all three conditions could be used to interact with the GUI, but that the bimodal interaction would be easier to use than the audio only or haptic only interface. Performance would be better (less error, less time) and also satisfaction would increase (absolute and relative ratings).

DESCRIPTION OF THE EXPERIMENT

Subjects

In order to test the hypothesis, two equal groups of subjects were used, of 12 sighted and 12 blind subjects (7 who had lost sight recently and 5 blind from birth).

Task

The experiment was designed to evaluate the audio and haptic dimensions of the interface in tasks corresponding to typical GUI manipulation. The environment was composed of icons, menus along the top of the screen, windows with a title bar (move button), a closing box on the left, and a resizing box on the right. Objects could be in two states: normal or selected. The subjects were asked to perform specific tasks typical of GUI: select objects, move windows and icons, open resize or close windows, pull menus up and down and select items in menus (table 4).

The preliminary experiment had shown that it was necessary for the user to have some feedback on his performance in the training phase, in order to learn to link the feedback (bimodal, audio or force) to the semantics of objects and actions. Consequently, voice synthesis feedback was added during the training phase ("icon", "menu item"), but was removed during the testing phases, where it was only used to standardize the description of tasks.

Process

In order to verify all the hypothesis and minimize individual differences, each subject tested the three conditions: haptic, audio, bimodal. The order of presentation of the conditions was varied in order to control the effect of learning (Latin square). The 19 test question were identical for each condition.

First the user explored the system and tried the various manipulation. Sound, force and voice synthesis feedback were given, as the user was guided in trying the system. Each sensorial condition was than tested following the sequence: exercise, test, questionnaires on satisfaction. Finally the subjects were asked to rank the three conditions of interaction to have a relative measure of satisfaction.

RESULTS

The measures of usability were performance (% of tasks completed successfully), mean time of execution, satisfaction and rank of the three conditions for six questions on usability

Table 1
Influence on performance of sight (blind vs "visually occupied") and modality (bimodal, haptic, audio).

Sight:	Sighted 12	Blind 12	Totals: 24
Audio	68%	61%	64%
Haptic	78%	71%	74%
Bimodal	83%	78%	80%
Totals:	76%	70%	73%

Analysis of Variance

Source:	F-test:	P value:
Sight (A)	3.684	0.0593
Modality (B)	7.608	0.0011**
AB	0.02	0.9801

Comparing blind and sighted users

The first hypothesis was to evaluate the usability of an interface for the blinds compared to sighted people. We first tried to verify differences between subjects who had lost their sight recently and those who had been born blind. No significant differences were observed for performance, mean time of execution, or general satisfaction between the two groups of blind users, so the data were grouped.

When blind subjects are grouped and compared to sighted users, results show that contrary to the hypothesis blinds did not performed better than the sighted, they even had slightly lower performance (table 1).

We investigated the possible effects of age and experience with GUIs, since blind subjects were older and few had experience with GUIs. Both factors appeared to give significant advantage to the sighted in our experiment (table 2).

For age, the analysis of variance suggest a significant relation to performance: younger subjects made less errors and took less time to complete the tasks . When the age factor is considered, the difference in performance between sighted and blinds is even smaller (p=.41). As for satisfaction, there is an interaction effect between age and sight, showing that middle aged blind subjects are less incline to be critical of the interface. Performance measures also show that blind subjects did not perform as well as sighted subjects in that age category (satisfaction often decreased with performance).

Experience with the GUIs appears to be related to faster performance (p<.01) and all sighted subjects had experience using them (12), while few blind subjects did (3).

Table 2
Influence on Performance, Time per exercise and Satisfaction of age, experience with GUI, sight and modality.

Analysis of variance (P values)			
Influence of age	Perform.	Time	Satisfaction
Modality (A)	0.0006**	0.1191	0.1772
Sight (B)	0.4108	0.5596	0.2248
AB	0.9702	0.4493	0.535
Age (C)	0.0178*	0.0003**	0.7728
AC	0.8365	0.7453	0.4722
BC	0.0178*	0.3737	0.0379*
ABC	0.6771	0.5536	0.345
Influence of experience			
Modality (A)	0.0004	0.3541	0.2983
GUI Experience (B)	0.1474	0.0076	0.0941
AB	0.2064	0.3521	0.8703
GUI only	p = .1873	p = .0081	p = .0907

Effect of the modalities used

The next objective was to investigate the efficiency of the interface depending on the modalities used to give feedback to the user Results in table 1 show that the influence of modality on performance is significant (p>.001) and that bimodality is the best condition followed by haptic only and then audio only condition.

There are no significant differences in mean time of execution or satisfaction between the conditions. Interestingly though, the ranking of the interfaces for the six questions of satisfaction are all significant (p<.001) (see table 3).

The results of the distribution of the modalities in the three ranks (number of times a modality appeared ranked as the best). It shows that the bimodal feedback is clearly preferred to the monomodal interfaces; although for a few users it causes some physical fatigue (4/24) and takes more concentration (2/24). If haptic alone was sometimes preferred, it might be due to the quality of the audio interface which was not pre-tested and adjusted as well as that of the haptic interface.

Table 3
Distribution measures (χ^2) of the ranks for the three interfaces and number of times each interface is ranked as the best..

Rank from the best to the worst ?		χ^2	NB ranked as best interface: Bimodal	Haptic	Audio	Totals
Q1	The interface is easy to use.	p=.0001	23	1		24
Q2	I would use it for daily tasks.	p=.0001	23	1		24
Q3	It is pleasant to use.	p=.0001	24	0		24
Q4	It takes less concentration.	p=.0001	21	1	2	24
Q5	It causes less mental fatigue.	p=.0001	22	1	1	24
Q6	It causes less physical fatigue.	p=.0001	20	2	2	24

There was no effect of interaction between the modality used and blindness on all measures of usability.

Modalities and difficulty of specific tasks

A detailed analysis of the tasks was done in order to highlight the effectiveness of the different interfaces for the various operations on the GUI. Table 4 shows the tasks which are the more difficult in the absence of visual feedback for the three conditions. It appears that closing and resizing windows are the most difficult tasks, due to the difficulty of positioning the pointer into a small box using only sound.

Table 4
Performance (percentage of success) for each tasks in general and for each modality.

		Mean	Audio	Haptic	Bimodal
Q10	Resize the window	40%	29%	42%	50%
Q7	Close the window	56%	33%	58%	75%
Q13	Resize the window	56%	33%	50%	83%
Q14	Close the window	57%	54%	50%	67%
Q11	Select the other window	61%	75%	54%	54%
Q6	Select the last menu item	64%	58%	54%	79%
Q12	Move window to the left	67%	63%	63%	75%
Q19	Select the second item	68%	54%	63%	88%
Q4	Move window to the left	74%	71%	71%	79%
Q5	Select the first menu	78%	67%	83%	83%
Q9	Open the icon	79%	71%	88%	79%
Q17	Open the icon	81%	63%	96%	83%
Q1	Select the lower icon	83%	83%	79%	88%
Q18	Choose the third menu from the left	83%	83%	88%	79%
Q8	Select the higher icon	85%	71%	92%	92%
Q2	Move icon down	86%	67%	96%	96%
Q15	Select the higher icon	89%	83%	92%	92%
Q3	Select the window	90%	83%	100%	88%
Q16	Move it to the left	92%	79%	96%	100%
Totals		73%	64%	74%	81%

As in general performance, using a bimodal interface gives the best scores for those task. We may postulate that adding more complex haptic and audio feedback might add precision in those cases.

DISCUSSION AND CONCLUSION

In this situation where a visual modality was replaced by one or two modality(ies), our research showed that multimodality was associated with better performance for blind and sighted subjects and that it is ranked as the best interface. The sound only feedback appears to be associated with the worst performance. These results may be related to the results of (Brewster, et al., 1994) for an audio-visual feedback (sounds reinforcing visual information), where the authors found more confort with multimodal feedback. Multi modality was reported as the best condition by almost all subjects. A very few people reported that it required more concentration and was more physically demanding. We suppose that as with other motor ability, this difficulty is likely to diminish with longer training.

There is no difference between blind and sighted people in any of the measures, especially when the effect of age and GUI experience is taken away. We can hypothesize that blind people would be less affected by fatigue for longer use (as in the preliminary experiment). They might also profit more from training since their GUI experience would then increase.

Improvements of our system will take these results into account. Force and sound representations may be reviewed to consider textures, a third force-feedback dimension and richer auditory aspects. Consequently, the Pantograph must become more haptic, incorporating the best possible tactile feedback, and possibly other kind of sound synthesis technics to the physical one. Finally, we aim to provide this multimodal feedback on common graphical operating systems such as MS-Windows.

REFERENCES

Appelle S. (1991). "Haptic perception of form: activity and stimulus attributes". In The psychology of Touch edited by Morton A. Heller and William Schiff, (p.169-188).

Blattner, M., Sumikawa, D., & Greenberg, R. (1989). Earcons and icons: their structure and common design principle. In *Human-Computer Interaction* Hillsdale, N.J.: Lawrence Erlbaum Ass. (p. 11-44).

Brewster, S. A., Wright, P. C., & Edwards, A. D. N. (1994). *The design and evaluation of an auditory-emhenced scrollbar. CHI'94,* Boston:ACM (p. 173-179).

Cadoz, C., Lizowski, L., & Florens, J. (1990). Modular feedback keyboard. *Computer Music Journal, 14*(2), (p. 47-50).

Gibson J.J (1966). "The senses considered as a perceptual system". Boston: Houghton Mifflin.

Göbel, M.(*to be published.*). Tactile Feedback applied to Computer-Mice. *Inyternational Journal of Human-Computer Interaction.*

Hayward, V., Choksi, J., Lanvin, G., & Ramstein, C. (1994). *Design and multi-objective optimization of a linkage for a haptic device. Proc. of the 4th workshop on Advances in Robot Kinematics, ARK'94,* Ljubljana, Slovenia:Kluver academic.

Heller, M. A. (1982). Visual and tactual texture perception: intersensory cooperation. *Perception &psychophysics, 31,* (p.339-344).

Lazzaro, J. (1991). Windows of Vulnerability. *Byte,* (June), (p.416).

Martial, O., & Dufresne, A. (1993). *Audicon: Easy Access to Graphical User Interfaces for Blind Persons - Designing for and with People. Human Computer Interaction,* Orlando, Florida:Elsevier (p. 808-813).

Martial, O., & Dufresne, A. (1994). *Les interfaces graphiques un espoir pour les aveugles. ErgoIA'94,* Biarritz.

Minski, M., Ouh-Young, M., Steele, O., & Brooks, F. P. J. (1990). Feeling and Seeing: Issue in force Display. *Computer Graphics, 24*(2), (p. 235-243).

Ramstein, C. ((to be published)). MUIS: multimodal user interface system with force feedback and physical models.

Ramstein, C., & Hayward, V. (1994). *The Pantograph: a large workspace haptic device for multimodal interaction. CHI'94, Interactive Experience,* Boston:ACM, (p. 57-58).

Ramstein, C., Martial, O., & Dufresne, A. (1994). *Pré-evaluation d'une interface à rétroaction multimodale conçue avec le MUIS.* CITI, 39p.

Weber, G. (1993). *Adapting direct manipulation for blind users. Adjunct to Interchi'93,* Addison Wesley (p. 21-22).

Welch, R. B., & Warren, D. H. (1987). Intersensory Interactions. *Handbook of perception and human performance, 1*(4 Space and motion perception), (p.25-1 - 25-36).