

Efficiency of Multimodal Metaphors in the Presentation of Learning Information

Marwan Alseid

Department of Computing, School of Informatics
University of Bradford
Bradford, BD7 1DP, England, UK
+44 7515363039

M.N.K.Alseid@Bradford.ac.uk

Dimitris Rigas

Department of Computing, School of Informatics
University of Bradford
Bradford, BD7 1DP, England, UK
+44(0) 1274 235131

D.Rigas@Bradford.ac.uk

ABSTRACT

The comparative study described in this paper has been conducted to investigate the effect of including multimodal metaphors on the usability of e-learning interfaces. Two independent groups of users were involved to evaluate two different interfaces of an experimental e-learning platform. The control group used the visual only interface that presents information about class diagram notation in textual approach, and the experimental group used the multimodal interface in which a combination of recorded speech sounds, non-speech sounds (earcons), and avatar with simple facial expressions were employed to communicate the same information. Three usability parameters which are efficiency, effectiveness, and users' satisfaction were considered in the study. The scope of this paper is to discuss the results that related to efficiency only, which has been measured by task completion time. It was found that the multimodal interface group taken significantly less time to complete the experimental tasks compared to the visual only interface group. These results encouraged for further exploration to examine the contributing role of each of the applied multimodal metaphors.

Categories and Subject Descriptors

H.5.2 [Information Interfaces and presentation]: User Interfaces – *auditory (non-speech) feedback, graphical user interface, interaction styles, Screen design (e.g., text, graphics, color)*. K.3.1 [Computer and Education]: Computer Uses in Education – *Computer-assisted instruction (CAI)*.

General Terms

Performance, Design, Experimentation, Human Factors.

Keywords

Efficiency, Multimodal interaction, E-learning, Usability, Avatar.

1. INTRODUCTION

User interfaces of the existing computer applications typically concentrate on the users' visual channel in communicating the presented information causing usability problems in these interfaces. Other human senses could be involved in HCI by incorporating multimodal interaction metaphors. Examples of these metaphors are: speech sounds (recorded or synthesized), short musical sounds (earcons), environmental sounds (auditory icons), and avatars with facial expressions and body gestures. Previous studies found that using speech sounds and earcons improved the usability of computer application in many different domains including e-learning. Other studies showed that employing avatar with facial expressions could be useful in e-learning. However, involving auditory and other multimodal metaphors in human computer interaction is still limited. The two-group experimental study reported in this paper aimed at investigating the usability aspects of e-learning interfaces that combines the inclusion of speech sounds (recorded), non-speech sounds (earcons) and the use of avatar with simple facial expressions, to deliver learning information. The following section presents an overview of the related work in e-learning and multimodality. Section 3 describes the experimental e-learning platform and section 4 details the design of the experiment. The obtained results are presented and discussed in section 5. Finally, section 6 concludes this paper and highlights the directions of future work.

2. E-LEARNING AND MULTIMODALITY

2.1 E-learning

The continuous developments in information and communication technology (ICT) caused accelerated developments in educational technology and facilitate the access to the learning materials. Many terms have been used to describe the use of technology in learning. In this paper, the term e-learning is used. A comprehensive definition of e-learning was introduced by the e-learning group of NCSA [1]. Other definitions of e-learning were introduced and most of them agreed that it is a method of learning in which learning experience can be enhanced by the help of ICT [2, 3].

Many benefits could be gained from e-learning. It offers better adaptation to individual needs, gives the ability to learn anywhere and anytime, and facilitate monitoring the student's knowledge and skills [4]. In addition, e-learning can offer common environment to formulate online learning communities where users can learn collaboratively [5]. Furthermore, educational material in e-learning can be updated and then distributed again easily and quickly and everyone can receive

© The Author 2008.

Published by the British Computer Society

the same content in the same way [6]. Also, it could be applied with different pedagogical methodologies [7]. However, e-learning also has defects. Some examples are: the technology needed in e-learning is not always available for access and use [8] and students, sometimes, are not satisfied with the use of ICT in the learning process and feel lack of personal interaction [9]. Therefore, students accessibility and satisfaction in e-learning interfaces should be enhanced [9] and pedagogical principles must be applied [10] to insure the success of ICT employment for learning purposes.

2.2 Multimodal Interaction

Focusing only on users' visual channel and neglecting other human senses (i.e. auditory channel) during the human-computer interaction process resulted in usability problems in the interface of many computer applications including e-learning. For example, users' visual channel could be overloaded [11] and important communicated information could be missed [12]. Therefore, involving more human senses in HCI will result in a more natural interaction which suit different individual abilities [13]. As a result, the amount of information delivered by one communication channel will be reduced [14], more than one sense could be used to receive different types of information [15], and user interfaces will be more usable.

Using interaction metaphors that utilize human auditory channel in computer applications has been evaluated by many studies. Earcons are short musical sounds [16]. It was found to be efficient in enhancing the interaction with user interface components [17-19]. Also, it was successfully used to provide auditory feedback regarding interaction events [20, 21]. Other studies demonstrated that earcons can improve students' attitude towards educational curriculum and enhance their understanding [22]. Earcons was successfully combined with recorded speech in the interface of multimedia on-line learning and assisted users to complete complex learning tasks more successfully [23].

Avatar is another interaction component through which auditory and visual human senses are used. It is a computer-based character that could be utilized to play the role of human being [24] with the ability to express feelings, emotions, and other linguistic information through facial expressions and body gestures [25]. It could be employed in e-learning environments to enhance users' attitude towards online courses [26] and to improve students' achievement by making the learning experience more interesting and motivating [27].

Short term memory is very important in the learning process for the remembrance and retention of the latterly perceived information [28]. Its capacity is one of the main cognitive traits [29]. Therefore, incorporating as much as possible of human senses in e-learning interfaces might help in perceiving and remembering the presented information and in turn enhancing learners' performance.

3. EXPERIMENTAL PLATFORM

An experimental e-learning platform was developed with visual-only and multimodal interface versions both of which were designed to communicate the same information about class diagram representation of specific problem. This information included explanation about classes, associations among classes and the multiplicity of a given class in the diagram. Three examples of class diagrams were also used in both versions, each of which was in a separate interface display.

These examples were gradually increasing in complexity and chosen to contain as much as possible of the notation frequently used in class diagram. In visual-only interface, this information was presented only by text and the interaction with the user occurs only via the visual channel; when the mouse cursor placed over a notation in the class diagram, a textual description of that notation appears in the provided notes text box.

The same design was used in the multimodal interface but the textual descriptions were removed and replaced with a combination of visual and audio interaction metaphors. An avatar with simple facial expressions (neutral, happy, sad, and smile) was used to communicate the information about classes, recorded speech was used to deliver the information about associations, and non-speech sounds (earcons) to represent the multiplicity. All of the auditory messages were communicated to users using headphones. The inclusion of these audio visual metaphors aimed to enable users to employ both of their visual and auditory senses in obtaining the communicated information.

The design of earcons used in the study was based on the suggested guidelines [30, 31]. Simple tones were used to create six earcons in order to represent six types of multiplicity needed in the three examples of class diagram representations. Four earcons were designed with two parts separated by a very short time of pause in between, each of which communicated one of the multiplicities: zero or 1 (0..1), one or more (1..*), two or more (2..*), and one or two (1..2). The other two earcons were designed of only one part to communicate the multiplicities one (1) and many (*) which represents zero or more. Therefore, there was a need to musically illustrate each of the values 0, 1, 2, and * in order to compose the appropriate earcons. One note of seashore sound was used to represent the value zero and different numbers of piano tones with rising pitch were used to represent the other values as follows: one tone to communicate 1, two tones to communicate 2, and four tones to communicate many (*).

4. EXPERIMENT

In order to accomplish the aim of the study, usability level in both interface versions of the experimental platform were compared in terms of efficiency. Efficiency was measured with the task completion time which has been considered as the dependant variable. The independent variable was the method used in the presentation of the learning information. Therefore, this variable has two levels: visual-only and multimodal.

4.1 Participants

Thirty users participated in the experiment in an individual basis. All of them were of scientific background and used the experimental platform for the first time. The majority of users were regarded as experts because they use computer ten or more hours a week. Also, most of them had limited or no background in class diagram notation.

4.2 Tasks

Each participant performed six common tasks. These tasks were gradually increasing in terms of its complexity and equally divided into easy, moderate, and difficult. It also covered all types of the presented information. Each task comprised a set of requirements each of which asked the user to place the mouse cursor over a given notation in the displayed class diagram and to receive the presented information related to that notation.

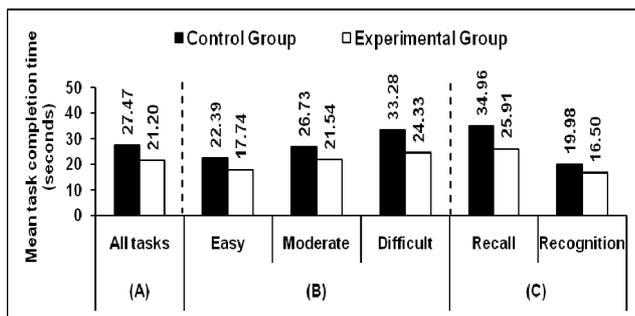


Figure 1. Mean values of completion time for all tasks (A), according to task complexity (B), and according to task type (C) in both groups.

The number of requirements in each task was proportional with the level of task complexity. After completion of the requirements, each task was evaluated with two questions; the first was a recall question in which the user needed to retrieve the answer from his/her memory. The second one was a recognition question in which the correct answer(s) should be recognized among a given set of 2 to 4 multiple choices. As a result, a total of twelve questions were required from each user; half of which were recall and the other half were recognition.

4.3 Procedure

Participants in the experiment have been divided into two independent groups; control group tested the visual only interface and experimental group tested the multimodal interface. Each group had 15 users. Users were randomly assigned to one of the two groups and each of them participated in the experiment individually. The experiment was explained to each user and started by filling the user profile. Then, 5-minute tutorial about class diagram notation was presented followed by 2-minute demonstration of the applied interface. Thereafter, the experimenter informed the user to start executing the required tasks according to the provided requirements. Once the participant completed a task, he/she was instructed to answer the two questions related to that task. At the end of the last task, each user answered the satisfaction questionnaire. This procedure was repeated for each participant.

4.4 Hypothesis

The hypothesis stated that the multimodal interface users will spend lower time to complete the required tasks in comparison with their counterparts who used the visual only interface.

5. RESULTS AND DISCUSSION

In the data collection and analysis, each question was considered as a task and the time spent by each user to answer each of the twelve questions was observed and analyzed using t-test. Although the visual only interface offered a simpler interaction, the results of the experiment shown in Figure 1A demonstrated that the overall mean completion time for all tasks was significantly lower in the multimodal interface group in comparison with the visual only interface group ($t=1.74$, $cv=1.72$, $p<0.05$). The reason behind this result was attributed to the inclusion of recorded speech, earcons, and avatars in the multimodal interface. During the experiment, it was noticed that users of the visual only interface have been compelled to keep moving their visual focus between the notes text box and the class diagram to understand the delivered information, which may result in scattering their attention and overloading

their visual channel. On contrary, users of the multimodal interface were able to focus their visual attention on the class diagram while hearing the communicated auditory information, giving them better concentration and accordingly allowing them to respond faster to the required tasks.

Figure 1B shows a breakdown of the results according to the complexity of the task. This Figure demonstrate the direct relationship between the time required to complete the task and the complexity of that task, which is logical, but the more important aspect is that the difference in completion time between the two groups increased as the task complexity increased. This difference was found to be significant in easy ($t=2.27$, $cv=1.66$, $p<0.05$), moderate ($t=2.12$, $cv=1.66$, $p<0.05$), and difficult ($t=2.99$, $cv=1.66$, $p<0.05$) tasks. These results explain the gradual effect of recorded speech, earcons, and avatar in users' achievement of higher complexity tasks.

Another detailed analysis of task completion time is presented in Figure 1C. Users in the experimental group spent significantly lower time than users in the control group to perform both of recall ($t=3.94$, $cv=1.65$, $p<0.05$) and recognition tasks ($t=2.42$, $cv=1.65$, $p<0.05$) but the difference was lower in the latter one. This indicated that adding the multimodal metaphors to the experimental interface helped users much more in the recall activities which are more difficult than the recognition ones. In recall tasks, users may have taken more time trying to retrieve the required communicated information which is not the case in recognition tasks where users needed only to choose the answer among the given options.

Overall, it can be summarized that using multimodal metaphors such as recorded speech, earcons, and avatar with simple facial expressions was more efficient than using text only in the presentation of learning information. As a result, accepting what has been hypothesized. This was evidenced particularly in the recall activities and when the required tasks become more complex.

6. CONCLUSION AND FUTURE WORK

This paper has shown that adding a combination of recorded speech, earcons, and avatar with simple facial expressions can enhance the efficiency of e-learning interfaces. The time taken to complete the required learning tasks was significantly reduced when these multimodal interaction metaphors were used to communicate learning information about class diagram notation. These results suggest that the addition of communication metaphors which employ both of visual and auditory channels would improve the usability of e-learning interfaces. Therefore, it is recommended to take these metaphors into consideration when designing user interfaces of e-learning applications. The results of the experiment did not clearly bring out the contribution made by each of the applied multimodal metaphors in the usability enhancement. This will be investigated in the next experimental work.

7. REFERENCES

- [1] Wentling, T.L., et al., *E-learning-A review of literature*. Knowledge and Learning Systems Group, University of Illinois at Urbana-Champaign learning. nesa. uiuc. edu/papers/elearnlit. pdf, 2000.
- [2] Alexander, S., *E-learning developments and experiences*. Education and Training, 2001. **43**(4-5): p. 240-248.

- [3] Yu, D., W. Zhang, and X. Chen, *New Generation of E-Learning Technologies*. Proceedings of the First International Multi-Symposiums on Computer and Computational Sciences (IMSCCS'06), 2006.
- [4] Mikic, F. and L. Anido, *Towards a Standard for Mobile E-Learning*. International Conference on Systems and International Conference on Mobile Communications and Learning Technologies, 2006. ICN/ICONS/MCL 2006. International Conference on Networking, 2006: p. 217-217.
- [5] Correia, A.P. and P. Dias, *Criteria for evaluating learning web sites: how does this impact the design of e-learning*. Actas da II Conferência Internacional de Tecnologias da Informação e Comunicação na Educação: Desafios Challenges, 2001: p. 521-528.
- [6] Rosenberg, M.J., *E-learning-strategies for delivering knowledge in the digital age*. 2001: McGraw-Hill.
- [7] Spalter, A.M., et al., *Considering a full range of teaching techniques for use in interactive educational software: a practical guide and brainstorming session*. Frontiers in Education Conference, 2000. FIE 2000. 30th Annual, 2000. 2.
- [8] Brady, L., *Fault lines in the terrain of distance education*. Computers and Composition, 2001. 18(4): p. 347-358.
- [9] Shaw, G. and N. Marlow, *The role of student learning styles, gender, attitudes and perceptions on information and communication technology assisted learning*. Computers & Education, 1999. 33(4): p. 223-234.
- [10] Govindasamy, T., *Successful implementation of e-Learning Pedagogical considerations*. The Internet and Higher Education, 2001. 4(3-4): p. 287-299.
- [11] Brewster, S.A., *Using non-speech sound to overcome information overload*. Displays, 1997. 17(3-4): p. 179-189.
- [12] Oakley, I., et al., *Putting the feel in 'look and feel'*. Proceedings of the SIGCHI conference on Human factors in computing systems, 2000: p. 415-422.
- [13] Dix, A., et al., *Human-Computer Interaction (3rd Edition)*. 2004, Prentice Hall.
- [14] Brown, M.L., S.L. Newsome, and E.P. Glinert, *An experiment into the use of auditory cues to reduce visual workload*. Proceedings of the SIGCHI conference on Human factors in computing systems: Wings for the mind, 1989: p. 339-346.
- [15] Sarter, N.B., *Multimodal information presentation: Design guidance and research challenges*. International Journal of Industrial Ergonomics, 2006. 36(5): p. 439-445.
- [16] Blattner, M.M., D.A. Sumikawa, and R.M. Greenberg, *Earcons and Icons: Their Structure and Common Design Principles*. Human-Computer Interaction, 1989. 4(1): p. 11-44.
- [17] Brewster, S.A., P.C. Wright, and A.D.N. Edwards, *The design and evaluation of an auditory-enhanced scrollbar*. Proceedings of the SIGCHI conference on Human factors in computing systems: celebrating interdependence, 1994: p. 173-179.
- [18] Brewster, S.A., et al., *The sonic enhancement of graphical buttons*. Proceedings of Interact, 1995. 95: p. 43-48.
- [19] Crease, M. and S. Brewster, *Making Progress With Sounds-The Design and Evaluation Of An Audio Progress Bar*. Proceedings of ICAD, 1998. 98.
- [20] DiGiano, C.J., R.M. Baecker, and R.N. Owen, *LogoMedia: a sound-enhanced programming environment for monitoring program behavior*. Proceedings of the SIGCHI conference on Human factors in computing systems, 1993: p. 301-302.
- [21] Rigas, D.I. and J.L. Alty, *Using sound to communicate program execution*. Proceedings of the 24th EUROMICRO Conference, 1998. 2: p. 625-632.
- [22] Upson, R., *Educational sonification exercises: Pathways for mathematics and musical achievement*. International Conference on Auditory Display, Kyoto, Japan, 2002.
- [23] Rigas, D. and D. Hopwood, *The Role of Multimedia in Interfaces for On-Line Learning*. 9th Panhellenic Conference on Informatics (PCI'2003). , Thessaloniki, Greece, 2003.
- [24] Sheth, R., *Avatar Technology: Giving a Face to the e-Learning Interface*. The eLearning Developers' Journal, 2003.
- [25] Beskow, J., *Animation of Talking Agents*. Proceedings of AVSP, 1997. 97: p. 149-152.
- [26] Annetta, L.A. and S. Holmes, *Creating Presence and Community in a Synchronous Virtual Learning Environment Using Avatars*. International journal of instructional technology and distance learning, 2006. 3: p. 27-43.
- [27] Theonas, G., D. Hobbs, and D. Rigas, *Employing Virtual Lecturers' Facial Expressions in Virtual Educational Environments*. International Journal of Virtual Reality, 2008. 7(1): p. 31-44.
- [28] Graf, S., T. Lin, and Kinshuk, *Analyzing the relationship between learning styles and cognitive traits*. Seventh IEEE International Conference on Advanced Learning Technologies (ICALT 2007), 2007: p. 235-239.
- [29] Lin, T. and P.A. Kinshuk, *Cognitive Trait Model-A Supplement to Performance Based Student Models*. Proceedings of International Conference on Computers in Education, Hong Kong, 2003: p. 629-632.
- [30] Brewster, S.A., P.C. Wright, and A.D.N. Edwards, *Experimentally derived guidelines for the creation of earcons*. Adjunct Proceedings of HCI, 1995. 95: p. 155-159.
- [31] Rigas, D.I., *Guidelines for Auditory Interface Design: An Empirical Investigation*. 1996: PhD thesis, Loughborough University of Technology.