A user satisfaction approach to multimodal interfaces for note-taking

Dimitrios Rigas, Mohamed Sallam

Abstract— the comparative study described in this paper to investigate the use of multimodal metaphors in interface of e-learning application. The primary focus is to present and discuss the experimental results related to users' views and evaluation of individual avatar and recording video used in the experimental e-learning tools. This investigation involved two different interface versions of the experimental e-learning tool. The first interface platform (the textual interface) was based on three input modalities, namely text, graphics, and images and was used to deliver information about note-taking. The second experimental platform developed for this investigation was based on visual and auditory metaphors. This platform interface (the multimodal interface) consisted of the three input modalities as well, including; speech, video, and avatar to deliver the same information. This was in a different type of windows for example login, select, optional menu, and assists or word help. The aim of this study is to provide a general guideline for Learning Software developers to help developing multimodal interfaces in order to enhance software usability and provide more information in learning interfaces.
The results obtained from this investigation have shown that the multimodal e-learning interface increased the level of usability as users took significantly less time to complete the given tasks, performed successfully in a higher number of tasks and were more satisfied than when using the textual interface. Also, the results indicate that users much preferred the video as a choice of input modality, while avatar was the second most preferred option for representing information. These input modalities could be used to improve the appeal of note taking which in turn will be reflected in increasing users' motivation and interest in the presented learning material.

***Keywords*—** E-learning, Usability, User interface, Multimodal Interaction.

Manuscript received February 25, 2010;

Mohamed Sallam is with the Department of Media Technology, De Montfort University The Gateway Leicester, LE1 9BH UK, BD7 1DP UK (e-mail: mohamedssallam@DMU.ac.uk).

Professor Dimitrios **Rigas** Head of Department of Media Technology. De Montfort University The Gateway Leicester, LE1 9BH UK +44 (0) 116 257 7485, email: d.rigas@DMU.ac.uk).

I. INTRODUCTION

Nowdays, most online applications are crowded interfaces and convey information to users via visual channel only. Therefore other human senses could be involved in human computer interaction to employ more interaction metaphors within the visual channel, the auditory channel or both "The auditory channel, as a whole, has been neglected in the development of user-interfaces, possibly because there is very little known about how humans understand and process auditory stimuli"[1]. Interfaces that offer interaction using more than one sense are highly demanded. This is called multimodal metaphors. Rigas et al, suggest that the use of multimodal metaphors in application learning interfaces can be more useful to communicate the information that 'needs' to be communicated to the user [2, 3, 4]. They found that the use of speech and non speech in interface application helped the users to make fewer mistakes and reduced the time taken when accomplishing their tasks [5, 6]. Several other studies have been carried out to test the use of multimodal metaphors in visual user interface and to evaluate and examine the effect of these metaphors on the usability of computer applications. Some of these studies suggest that the use of multimodal metaphors such as speech sounds, non speech sound and avatar could improve the usability of computer interfaces in many different ways including in e-learning application. Nevertheless, more investigate sill needed in this field. Using multimodal interaction in a multiple interfaces including e-learning can enhance human-computer interaction [7]. In this experiment, we investigated the effect of including multimodal metaphors such as recorded speech, video and avatar with simple facial expressions to communicate data, and see how the addition of these metaphors affected the usability of an e-learning system [8, 9].
The aim of this study is to provide a general guideline for Learning Software developers to help developing multimodal interfaces in order to enhance software usability and provide more information in learning interface.

II. E-LEARNING

The main objective of the learning process is to develop the individuals maximum potential, and e-learning is a form of learning which uses electronic applications to deliver learning experiences [10, 11]. Also, e-learning can be used as a collective term that describes learning with the use of internet technologies that allows learning to take place without being

constrained by time or location [12]. Many benefits can be gained from e-learning as it accommodates individual needs, access to online learning from anywhere at any time, and reduce the costs of delivering information [12, 13].

Time is valuable for full time workers as they cannot attend classes regularly. Location also is important for student or people who live far away from university or school. In addition, courses fee are not high and this is ideal for students who cannot pay traditional course fees. However, e-learning also has its limitations including reduced social and cultural interaction, technical issues and some courses can be difficult to simulate in e-learning applications. Therefore, the student's use of e-learning interfaces could be advantageous.

E-learning can be classified into four categories: 1) Online learning [14]. 2) Web-based learning [15]. 3) Computer-based learning [16]. 4) Distance learning [17]. Online learning refers to online study with or without instructors and also describes learning via internet, intranet and extranet [14]. Web-based learning refers to the use of both technology and traditional methods in learning [15]. Computer-based learning is the term for self-study using a computer, often in conjunction with a CD-ROM with exercises and self-tests [16]. The term distance learning refers to studying using several methods such as TV, Radio and also describes learning via a broadcast of lectures to distant locations, usually through video presentations and other technology [17, 18]. Recently, the Internet has become the major widely used method of transferring course materials. Overall, the term e-learning can be used for any type of learning that uses a computer [19, 20].

III. NOTE-TAKING

Note-taking is a skill which it can be used in many different ways in our life, at university, school and in the workplace [21]. On the other hand, it is a strategy for storing important information. Note-taking is connecting information to existing knowledge and the more connections we make the better value information we get and understand. For example, when you tell something new, you will understand it and to remember that you will refer it to something you know already. In face to face lecture students might find it difficult to write what is important or what is not. In online or distance learning students want to copy out from textbooks and what they read. Note taking can helps you to remember and reviewed what you have learnt and you only draw or write what you will need later on. Several studies have shown that note-taking can support remembering, thinking, and clearing to encourage learning [22].

A. Anticipated Benefits

According to the literature studied, use of multimodality can influence a student's learning [23]. As e-learning applications are widely used, the expected benefits of using multimedia in e-learning systems, in particular within the area of note-taking, will result in improving student performance by reducing the time required to complete the tasks with fewer errors and to enhance student understanding and satisfaction [24]. Also, it will provide additional usability guidelines for development of multimodal metaphors in e-learning applications.

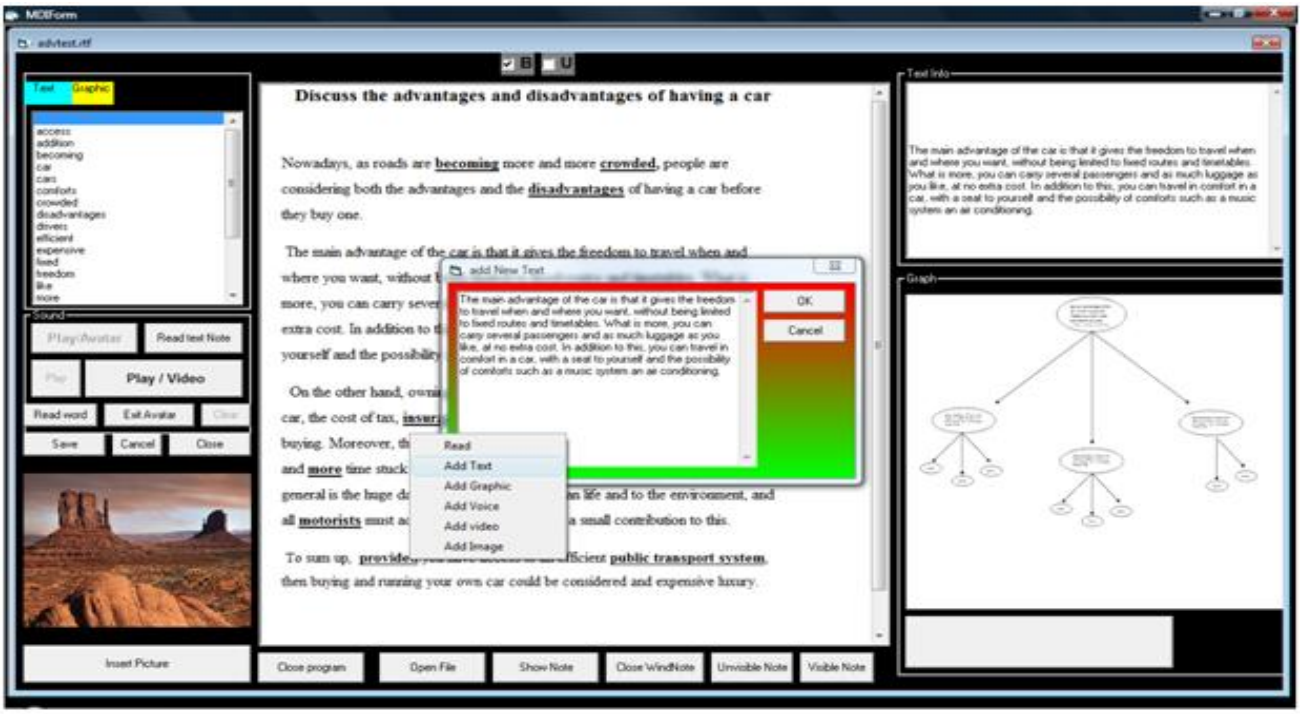

Fig. 1 The textual interface

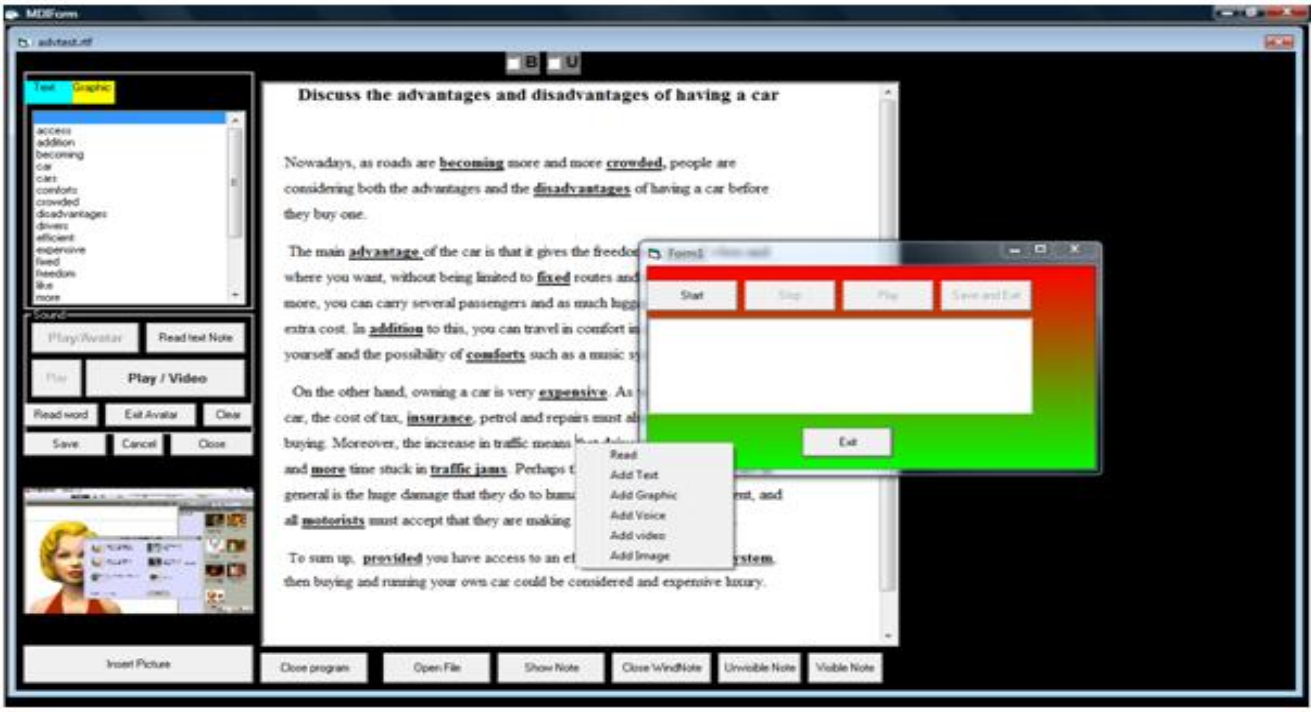

Fig. 2 The multimodal interface

IV. MULTIMODAL METAPHORS

A series of investigations in the use of multimodal metaphors in human computer interaction have taken place. These investigations prove that the usability of user interfaces can be enhanced when multimodal metaphors are included [25, 26]. Users will feel that they interact with e-learning applications more naturally when the visual and auditory senses are utilised in human computer interaction [27].

These auditory and visual senses can be used in e-learning environments to improve students' achievement and enhance users' attitude towards online courses by making the learning experience more stimulating [28]. Furthermore, it has also been proven that providing a multimodal learning environment makes learning more exciting and fun as learners enjoy interaction whilst being taught rather than boring textual delivery of information [25].

In software applications, speech and sounds, after visual output, are the most common methods for communicating a response to the user [29]. It was found that the earcons with short musical sounds are efficient and effective metaphor in the interaction with users' interface [30]. Moreover, using recorded speech and earcons in the interface of multimedia e-learning application assisted users to complete complex learning tasks more successfully [29].

Another visual and auditory interaction was involved in this experiment which is the avatar. It is a computer based character that could be used to play the role of human being and has the ability to express feelings, emotions, and other linguistic information via various facial expressions [31]. It could be utilized in e-learning environments to enhance the usability.

In this experiment video, speech, and avatar are used for enhancing efficiency, effectiveness of the interface and users' satisfaction. This study investigates comparison of the usability experience using the above mentioned modalities with visual-only interaction metaphors including text, graphic, and image.

V. VARIABLES

A. Independent Variable

Independent variables are those which were controlled during the experiment to ensure output results consistency. These variables were:

1. Training session. All participants had the same training session about how to use the interface system. This was ensured by creating two training video to explain how to use each task. Further explanation and help was provided when they needed.
2. Required tasks: All participants had the same number of tasks, four in each interface related to three levels of complexity; easy, moderate and difficult tasks.
3. Required tasks time: All participants had to finish required tasks in required time. Otherwise, the tasks would be regarded as unsuccessful.

B. Dependent Variable

The dependent variables, there are the output results of manipulating the independent variables.

1. Efficiency: Task accomplishment time: this measured

by the actual time taken to complete the task

1.2. Number of mouse clicks: this measured of the number of mouse click in each task.

1.3. Number of errors: this measured of the number of errors which the users made while he/she performance the tasks.

2. Effectiveness: Percentage of successfully completed tasks: the tasks which completed in required time as a percentage of total tasks performed.

2.2. Participants completed tasks: The percentage number of participants who successfully completed all tasks.

3. Satisfaction: Overall satisfaction: After performing all tasks, participants were required to complete a satisfaction questionnaire for each interaction metaphor used in the experiment.

VI. EXPERIMENT

The aim of this experiment was to implement a typical e-learning application to investigate the effect of using multimodal metaphors in e-learning interfaces and how users would interact with the system. The experimental setup was done on two different versions of applications. These versions of the experimental e-learning application were built from scratch to be used in this empirical study.

The first interface platform (textual interface) based on three input modalities, namely, text, graphic, and image was used to deliver information about note-taking. The second experimental platform was developed for this investigation and based on visual and auditory metaphors.

This platform interface (multimodal interface) consisted of the three input modalities as well as including; speech, video, and avatar to deliver the same information. This was in a different type of window such as login, select, optional menu, and assists or help window which included explanations about the use of platforms. In each interface participants were required to make notes about specific words by selecting a word and then right clicking the mouse to display a menu of options (add speech, add video, add avatar in multimodal interface and add text, add graphic, add image in textual interface). For example, in the multimodal interface the participant was required to read and select a word from a passage of text and then make some notes relating to the selected word by speech. The same task was then replaced with a recorded video for adding notes. In the last task, a human-like avatar was included in the multimodal interface to represent the recorded speech.

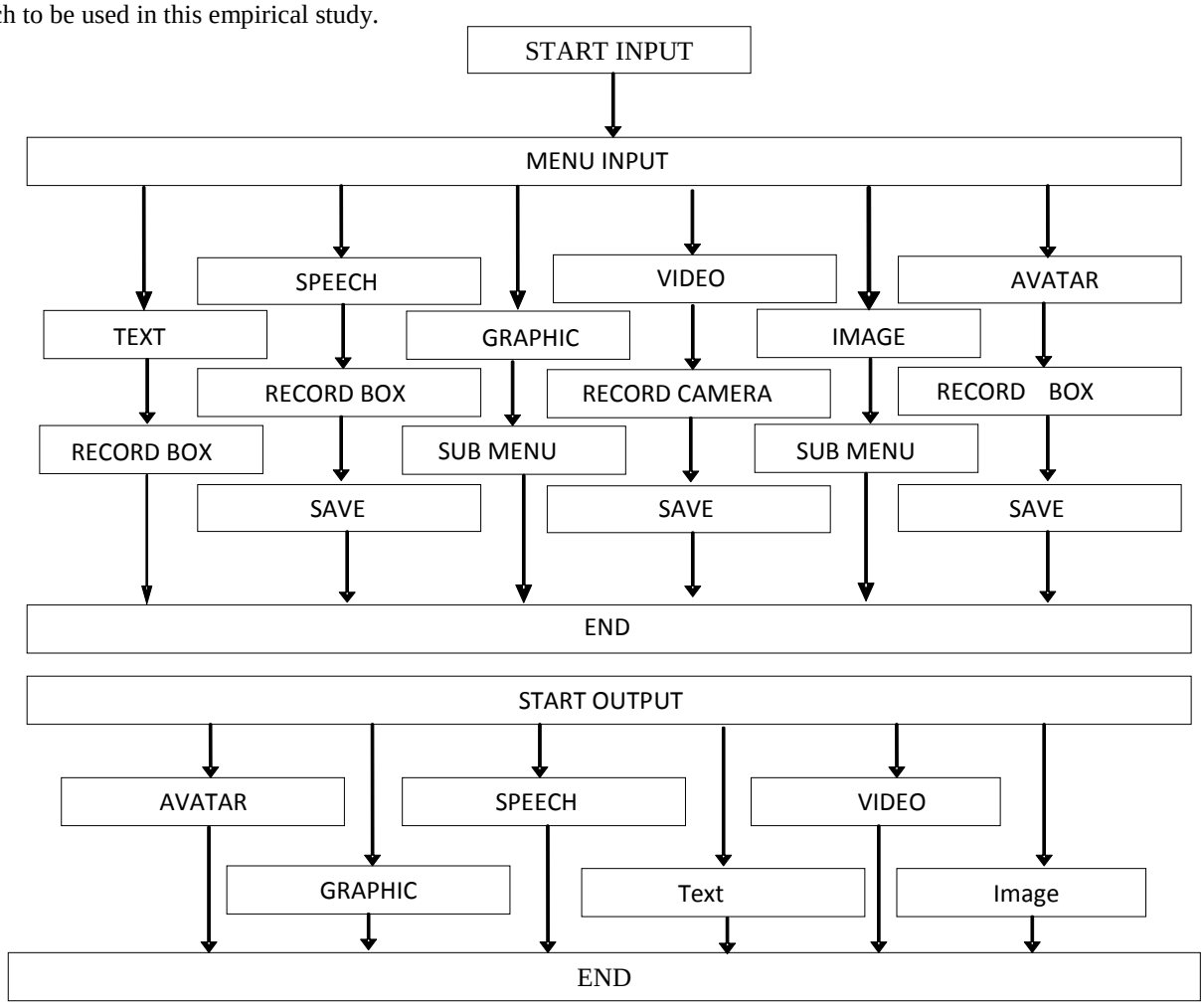

Fig.3 Flowcharts for both interfaces

A. Questionnaire

The participants were asked to read the instructions in the questionnaire and type their name in the space provided. The above action was to enter their gender and area of study. The participants were also asked to go through five tasks. The experiments measured effectiveness, interactivity and user satisfaction. The questionnaire was the same for all two interfaces except that the post –question varied. This was done because there was a difference between the multimodality which was added in each interface. The first platform interface (textual interface) was based on three input modalities; text, graphics, and images. The second platform interface (experimental interface) was based on three input modalities, namely recorded speech, video, and avatar. On completion of the tasks, the participants were required to write down their views about the experiment.

The tasks were designed with the objective of measuring ease of information access and how clearly the descriptions matched the main form. The participant was asked to make some notes on both interfaces. Interface one by either by adding text, graphic, and image. In interface two by adding recorded speech, video, and avatar. The questionnaire attached to the prototype contained detailed instructions and the six tasks the participants were asked to perform.

The participant was asked to select a word after reading the text and then write some notes which related to the word selected. For example, in the first interface platform (multimodal interface) The first task enabled recording of speech, for example, click on the word 'crowded' and then right-click on the mouse to display a menu of options and choose add speech. The user recorded his/her notes verbally which related to the word selected.

In the last task, a human-like avatar was included in the multimodal interface to represent the recorded speech.

After performing each task the participants were asked to rate the user satisfaction, perception of information, and description. Using the scale below to indicate how much he/she agreed or disagreed with the statements by circling the number that most closely describes their view.

B. Participants

Forty participants, consisting of under-graduates and post-graduates were selected to investigate the effects of including multimodal metaphors usability of e-learning interfaces.

A post-experimental questionnaire at the end of the experiment was answered by all participants. Participants consisted of 15% with a bachelor's degree, 30% with a doctorate degree and the remaining 55% had a master's degree.

The participants have been grouped into three categories on the basis of age. The obtained result shows that the majority were aged between 25 and 34 years old (43%) followed by those between 35 and 44 (38%) and the remaining were over 35 years old.

The gender of the participants was 78% male and 23% females. The reason for a low number of female participants was due to scarcity of females meeting the criteria of English as a second language and the requirement of some basic computer competency. The participants also had a scientific background and they were using the experimental platform for the first time. Also the number of participants who had limited knowledge about human computer interaction in the experiment was 35%, the percentage with good knowledge was 35% and about 25% had no knowledge.

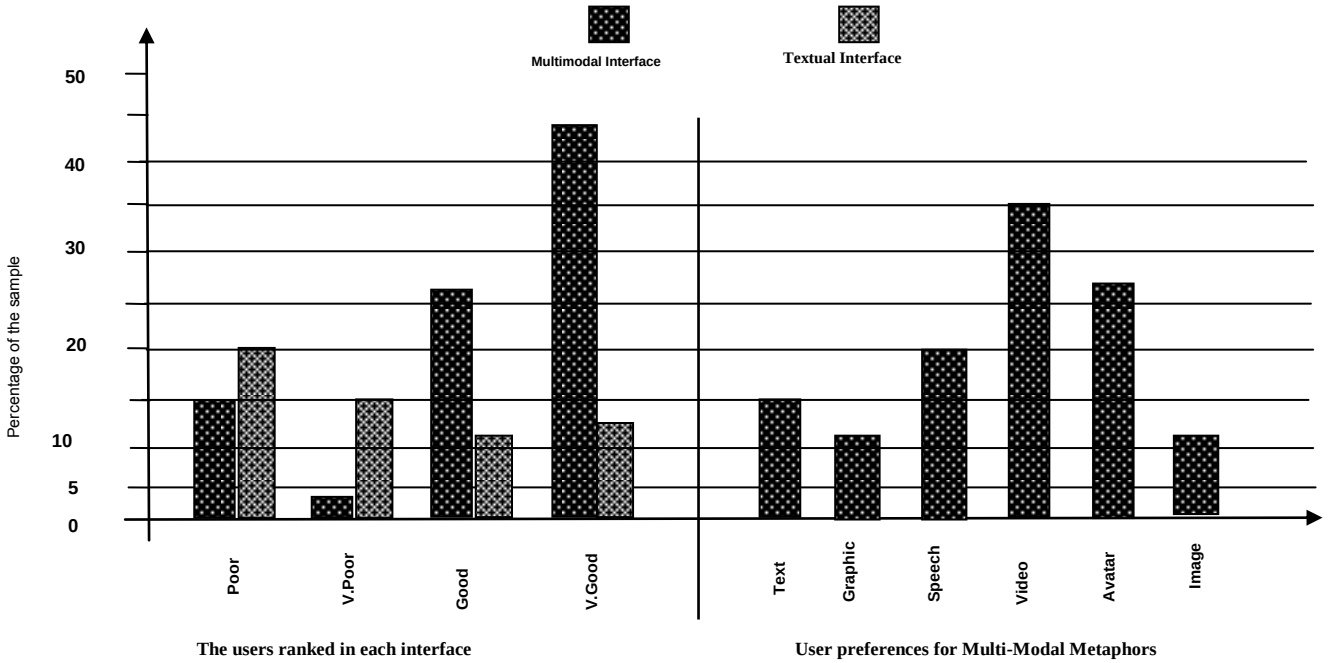

Fig. 4 The users ranked and user preferences of both interfaces

In order for the experiment to be successful, all participants had to fulfill a certain set of criteria. The requirements were a) computer literacy (i.e. used computers for more than 10 hours a week), b) had not used the experimental platform before. c) Spoke English as a second language.

Approximately 43% of users' area of study was computing and 13% was telecommunication. The rest of the samples were based in electronics, engineering, networking and communication. The analysis of the respondents found that 84% of the total used a computer for more than 10 hours.

The numbers who had used a computer for between 1-5 and 6-10 hours were nearly equal however, only 2% of the total selected said that they never used a computer. The average number of participants who used the internet for less than 10 hours a week was 28%. The number increased to 73% for those who used the internet for more than 10 hours a week.

C. Tasks

The participant was asked to go through four tasks in each interface. The tasks were designed with the objective of testing all the three different modalities listed above for the multimodal interface.

For the textual interface the steps were exactly the same for each task. They were given a set of pre-selected words and some notes to add as comments for them. These tasks were gradually increasing in terms of complexity; thus, were equally divided into easy, moderate and difficult.

Each task comprised a set of requirements which asked the user to place the mouse cursor over a selected word.

In the first platform interface (textual interface) the user was required to make notes about a specific word by text, graphic and image.

D. Procedure

In order to fulfil the aim of this experiment, a within-subject approach was employed in carrying out the experiment. After having tutorial on how to use the e-learning application system and receiving a complete explanation for adding notes by different modality, participants had ten minutes to read the questionnaire introduction and answer the pre-experimental questions and choose the right answer such as age, gender and education level. Participants also had to select the number of hours of internet and computer usage per week plus their experience in using e-learning applications, in particular within the area of note-taking. All participants had to go through two screenshot-based training sessions about how to use the interface system. Further explanation and help was provided when they were needed. The purpose of this training was to allow users to become familiar with the multimodal metaphors. The experiment time was recorded for each task and for the whole experiment as well. The number of questions and errors was recorded for all tasks. The participants were then asked to complete the satisfaction questionnaire.

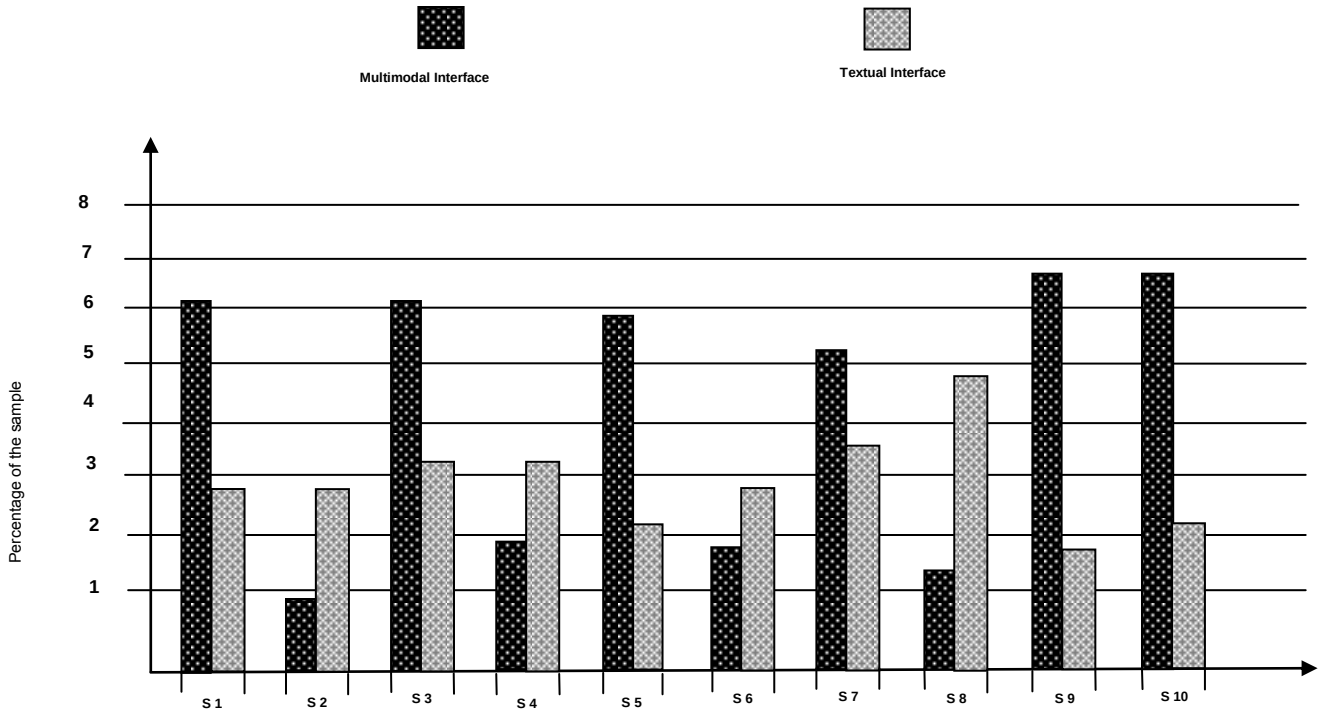

Fig. 5 The mean score of satisfaction for both interfaces

VII. RESULTS AND DISCUSSION

A. Satisfaction

The responses of users in both interfaces to the post-experimental questionnaire were used to measure their satisfaction. This questionnaire was scored 1- 5 on the Likert Scale with ten statements regarding each interface, which fitted all experimental conditions, and the users were required to specify their agreement to these statements. These statements were mainly about the ease of use, complexity, uneasiness and usefulness of each metaphor and overall satisfaction. Users were asked to select their preferred interface and provide an explanation for their choice. The five points Likert scale was used for each statement in the questionnaire, ranging from 1(strongly disagree) to 5-(strongly agree). The score of each statement for each user in the questionnaire was added together to create an overall user's satisfaction. t-Test was used to evaluate the significance of the difference between the two interfaces by the total number of scores to test the difference in the users' satisfaction. The results in figure 4 show that participants were significantly more satisfied in all statements when using the multimodal interface than when using the textual interface ($t = 2.94$, $cv = 1.68$, $p<0.05$).

Consequently, statistical results showed that using multimodal metaphors in e-learning applications, in particular within the area of note-taking, was found to be more satisfactory than the interface with text, graphics, and image.

B. Evaluations

The results of the experiment showed that users preferred to use multimodal such as speech, video, and avatar when they want to make their own comments. Figure 3 shows there were 30%, 25% and 17% of the users in the multimodal interface who preferred to use the application based on video, avatar and recorded speech respectively. The users wanted the multimedia metaphor options such as video and so on to be available in the textual interface. Furthermore there were approximately 78% of participants who participated in the experiment and described the multimodal interface as very good. Only a small percentage expressed a negative view of such use. The experimental interface with multimodal was chosen and ranked positively by almost every user.

Conversely, about 61% described the textual interface as poor and only 11% described it as good or very good. There was a noticeable difference in the successful completion of tasks that involved complexity between the multimodal interface and textual interface. In those tasks, it appeared that for most of the tasks the users performed better in the multimodal interface than in the textual interface.

VIII. CONCLUSION

This experiment investigated the use of multimodal metaphors in the interface of e-learning applications. A combination of multimodal metaphors such as recorded speech, video, and avatar was found to be significantly more satisfactory (in terms of ease of use, confusion, nervousness and overall satisfaction) than using text, graphic, and image only. Also, the results in this paper have shown that incorporating speech, video, and avatar can improve the usability of e-learning applications in particular within the area of note-taking.

References

[1] Rigas, D, D.Memery, et al. (2001). Experiments in using structured musical sound, synthesised speech and environmental stimuli to communicate information: is there a case for integration and synergy? Intelligent Multimedia, Multimedia, Video and Speech processing,

[2] D. I. Rigas and D. Memery, "Utilising audio-visual stimuli in interactive information systems: a two domain investigation on auditory metaphors," Proceedings of the International Conference on Information Technology: Coding and Computing.

[3] Rigas, D.I., Guidelines for Auditory Interface Design: An Empirical Investigation. 1996: PhD thesis, Loughborough University of Technology.

[4] Rigas, D.I. and J.L. Alty, Using sound to communicate program execution. Proceedings of the 24th EUROMICRO Conference, 1998. 2: p. 625–632.

[5] 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.

[6] Rigas, D.I., Guidelines for Auditory Interface Design: An Empirical Investigation. 1996: PhD thesis, Loughborough University of Technology.

[7] S. Oviatt, "Multimodal Interfaces," The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies, and Emerging Applications, 2003.

[8] V. Jones and J. H. Jo, "Ubiquitous learning environment: An adaptive teaching system using ubiquitous technology," Beyond the comfort zone: Proceedings of the 21st ASCILITE Conference, pp. 468-474, 2004.

[9] P. E. Sanderson, "E-Learning: strategies for delivering knowledge in the digital age," The Internet and Higher Education, 2002.

[10] J. McGovern and K. Gray, "Directions for organisation and management of university learning: Implications from a qualitative survey of student e-learning," Proceedings of ASCILITE 2005, 2005.

[11] M. Paiting, E-learning: is it really the best thing

since sliced bread, 2002.
[12] J. M. Rosenberg, "E-Learning: Strategies for Delivering Knowledge in the Digital Age.," 2001.
[13] R. Hamilton, C. Richards, et al., "An examination of e-learning and e-books," 2001.
[14] M. Nichols, "A theory for eLearning," Educational Technology & Society, vol. 6, pp. 1-10, 2003.
[15] T. Bates, Technology, e-learning and distance education: Routledge, 2005.
[16] J. McGovern and K. Gray, "Directions for organisation and management of university learning: Implications from a qualitative survey of student e-learning," Proceedings of ASCILITE 2005, 2005.
[17] J. C. Richardson and K. Swan, "Examining social presence in online courses in relation to students‟ perceived learning and satisfaction," Journal of Asynchronous Learning Networks, vol. 7, pp. 68-88, 2003.
[18] R. D. M. A.Gunasekaran, and D. Shaul, , "E-learning: research and applications, Industrial and Commercial Training," 2002.
[19] G. Singh, J. O'Donoghue, and H. Worton, "A study into the effects of elearning on higher education," Journal of University teaching and learning practice, vol. 2, pp. 13-24, 2005.
[20] A. P. Rovai, "Sense of community, perceived cognitive learning, and persistence in asynchronous learning networks," The Internet and Higher Education, vol. 5, pp. 319-332, 2002.
[21] Francoise Boch, A. p. (2004). "Note Taking Learning: A Summary of Research."
[22] Wirth, M. A. "E-notes: Using Electronic Lecture Notes to support Active Learnign in Computer science." (2003).
[23] N. B. Sarter, "Multimodal information presentation: Design guidance and research challenges,," 2006.
[24] R. Wagner, J. Werner, and R. Schramm, "An evaluation of student satisfaction with distance learning courses," 2002.
[25] A. Dix, G. Abowd, J. Finlay, and R. Beale, Human-Computer Interaction (3rd Edition). Prentice Hall, 2004.
[26] A. J. Dix, Human-computer interaction. New York; London: Prentice-Hall, 1993
[27] S. A. Brewster, P. C. Wright, A. J. Dix, and A. D. N. Edwards, "The sonic enhancement of graphical buttons," Proceedings of Interact, vol. 95, pp. 43-48, 1995.
[28] Rashid, S., and Rigas, D. (2007). An Initial Empirical Study into E-Note-Taking, In Proceedings of IADIS International Conference Interfaces and Human Computer Interaction (MCCSIS ’07), Lisbon, Portugal, (3-8 July, 2007), pp. 161-165.
[29] D. Rigas and D. Hopwood, "The Role of Multimedia in Interfaces for On-Line Learning," 9th Panhellenic Conference on Informatics (PCI'2003). , Thessaloniki, Greece, 2003.
[30] M. M. Blattner, D. A. Sumikawa, and R. M. Greenberg, "Earcons and Icons: Their Structure and Common Design Principles," Human-Computer Interaction, vol. 4, pp. 11-44, 1989.
[31] J. Beskow, "Animation of Talking Agents," Proceedings of AVSP,
vol. 97, pp. 149-152, 1997.