

# A Comparative Usability Study on the Use of Auditory Icons to Support Virtual Lecturers in E-Learning Interfaces

Marwan Alseid

Department of Software Engineering  
Applied Science University  
Amman, Jordan

Mohammad Azzeh

Department of Software Engineering  
Applied Science University  
Amman, Jordan

Yousef El Sheikh

Department of Computer Science  
Applied Science University  
Amman, Jordan

**Abstract**—Prior conducted research revealed that the auditory icons could contribute in supporting the virtual lecturers in presence of full body animation while delivering the learning content in e-learning interfaces. This paper presents further empirical investigation into the use of these supportive auditory icons by comparing three different e-learning interfaces in terms of usability aspects; effectiveness, user satisfaction and memorability. The aim is to find out which combination of the tested multimodal metaphors is the best one in terms of utilizing the auditory icons to supplement the presentation of learning material by virtual lecturer. The first experimental e-learning interface incorporates a speaking virtual lecturer with full body gestures along with supportive auditory icons. The second experimental e-learning interface includes the use of virtual lecturer speech in the absence of his body and accompanied with the same auditory icons used in the first interface. However, the third interface is similar to the second one in terms of using the virtual lecturer's speech but without any additional auditory icons. The obtained results have shown that the inclusion of auditory icons could enhance the usability and learning performance of e-learning interfaces much better if combined along with speaking virtual lectures in the absence of any body animation.

**Keywords**—auditory icons; virtual lecturer; e-learning; usability; speech; avatar; multimodal interaction

## I. INTRODUCTION AND MOTIVATION

The world has witnessed and still, a tremendous and accelerating development in the field of information technology and computer networks, which resulted in a quick and easy access to a huge amount of information including educational content. E-Learning describes the learning process that utilizes information and communication technology in delivering and managing the learning material. Recently, the majority of e-learning interfaces concentrate on user's visual channel to communicate information whereas other multimodal interaction metaphors could be involved to make use of other human sense in the interaction process. Previous related work demonstrated that speech and non-speech sounds as well as virtual lecturers could be beneficial in enhancing the usability of e-learning interfaces. Even though, the incorporation of multimodal interaction metaphors in this domain needs to be investigated further. The aim of the presented experimental study is to reveal the best utilization of auditory icon as

supportive sounds to the virtual lecturer in e-learning interfaces. Therefore, three different e-learning interfaces have been developed and independently tested in terms of effectiveness, memorability and user satisfaction. Each of these interfaces involved different combination of speaking virtual lecturer and auditory icons to communicate the learning material about class diagram notation. The following sections present an overview of the relevant work in e-learning and multimodal interaction, the experimental e-learning interfaces, the experimental design, analysis and discussion of the obtained results.

## II. RELATED WORK

E-learning is the term that describes the learning process via information and communication technology [1, 2] in which a huge educational content could be easily and quickly accessed. As a result, utilizing this technology in delivering e-learning content has been and still investigated by researchers. Scheduled delivery is an example of the technology used in e-learning [3] where video broadcasting, remote libraries, and virtual classrooms have been used and constrained with time and place. This technology has been enhanced by the on-demand delivery platforms that facilitate anytime and anywhere learning in the forms of interactive training CD ROMs and web-based training. Compared to traditional learning, e-learning has many advantages such as offering more flexible learning in terms of time and location, enabling better adaptation to individual needs [4], facilitating online collaborative learning over the Internet [5] as well as increasing learners' motivation and interest about the presented material [6]. Nevertheless, it was found that students felt uncomfortable with computer-based learning and missed traditional face-to-face interaction with teacher. Therefore, users' accessibility and their attitude in regard to e-learning should be enhanced [7].

Multimodal interaction involves more than one human sense in human computer interaction and could be utilized to enhance the usability of user interfaces. It facilitates the use of different channels to communicate different information [8]. Also, it enables users to employ the most suitable communication metaphor to their abilities [9]. This multimodality in the interaction process was found to be helpful in enhancing the learning experience where visual, aural, haptic and other channels could be integrated in a

Acknowledgment: The authors are grateful to the Applied Science Private University, Amman, Jordan for the full financial support granted to this research.

multimodal approach to deliver the learning content in e-learning interfaces.

Avatar is a multimodal metaphor that represents a human-like or cartoon-like computer-generated character [10] and utilizes both auditory and visual human senses in human-computer interaction. It has been used in interactive interfaces to communicate verbal and non-verbal information through facial expressions and body gestures [11]. Also, it was found that users' satisfaction and their ability to understand and remember the provided knowledge has been enhanced by the incorporation of speaking avatar [12]. In addition, it has been demonstrated by several studies that the use of avatars contributed positively in terms of facilitating the learning process and enhancing users' satisfaction in e-learning [13-15] and in e-book assessment interfaces [16]. Even though, these studies did not investigate the use of avatars along with auditory icons in e-learning interfaces.

Speech and non-speech sounds could be used to complement the visual output; however, sound is more flexible as it can be heard without paying visual attention to the output device. It was found that speech sounds could contribute with graphics and non-speech sounds (earcons and auditory icons) to enhance the usability of search engines interfaces [17] and e-government interfaces [18]. A study by Alseid and Rigas [19] investigated three different e-learning interfaces and demonstrated that the interface incorporating speaking virtual lecturer with full body gestures was found to be the most efficient, effective and satisfactory in communicating the learning content as opposed to the other two interfaces incorporating either single or two talking heads of facially expressive virtual lecturers. Even though, that study did not explore the incorporation of supportive non-speech sounds such as auditory icons. Therefore, a further experimental study has [20] been carried out by the same authors to investigate the non-speech sounds when used along with the speech of full body animated virtual lecturer during the presentation of learning content and found that earcons and auditory icons could be used beneficially in communicating auditory messages related to important parts of the learning content while being delivered by speaking virtual lecturer. However, that study involved only one group of participants who tested only one interface and did not investigate other combinations of non-speech sounds with virtual lecturers in e-learning interfaces. Therefore, this study aimed to investigate the use of auditory icons further. More specifically, it compares three experimental e-learning interfaces in terms of usability. Two of these interfaces incorporate auditory icons to support speaking virtual lecturer but one of them include the presence of full body gesture of the virtual lecturer whereas the other do not. The third interface employs the speech of virtual lecturer in the absence of his body gestures and any supportive auditory icons.

### III. EXPERIMENTAL E-LEARNING INTERFACES

Three different e-learning interfaces have been designed and built to serve as a basis for this study. These interfaces provide a command button to start the presentation of the learning material and pause/play functionalities to facilitate more control on the learning. The first interface; VSNS (Virtual lecturer with Speech and Non-Speech) includes an avatar as a

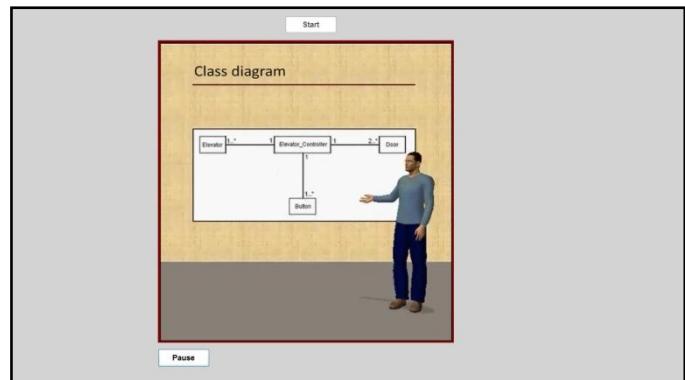


Fig. 1. Screenshot of VSNS (Virtual lecturer with Speech and Non-Speech) interface.

virtual lecturer with full body gestures and speaking naturally (recorded speech) with prosody to present the learning material in audio-visual format. Also, it includes the use of auditory icons to support the learning material presented by the virtual lecturer. The aim of including the avatar as mentioned earlier was to imitate the traditional interaction between the lecturer and the learner that usually take place in traditional classroom-based learning. In addition, VSNS interface offers textual and graphical representation of the communicated learning material placed in the background of the virtual lecturer who has been designed to simulate body gestures similar to those usually used by the human lecturer in the real classroom situation. Fig. 1 shows a screenshot of VSNS interface. In the second e-learning interface; RSNS (Recorded Speech with Non-Speech), the same textual and graphical information related to the presented learning material is placed in the middle of the interface (see Fig. 2) and explained similarly by the speech of the virtual lecturer but in the absence of his body. In addition, the same supportive auditory icons used in VSNS have been used similarly in RSNS. However, the third interface; RSON (Recorded Speech ONLY) is identical to RSNS but without using any auditory icons.

Auditory icons are non-speech sounds come from the surrounding environment such as opening a window, closing a window, dropping a can and shredding a paper sounds. Such sounds have been empirically investigated to communicate

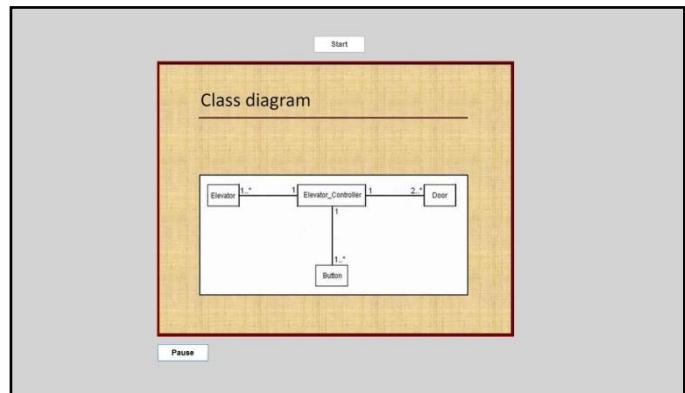


Fig. 2. Screenshot of RSNS (Recorded Speech with Non-Speech) interface.

different types of information and found to be beneficial in delivering such information and successfully enhanced the usability of user interfaces [21]. As explained earlier, two of the experimental e-learning interfaces (VSNS and RSNS) include the use of auditory icons. This inclusion aimed to bring users' attention to two key aspects of the learning material while being communicated by the virtual lecturer. These are the beginning and end of an important statement in the presented content where the sound of door opening has been used to indicate that the virtual lecturer is about to start mentioning an important statement, and the sound of door closing to indicate that this statement completed. Accordingly, representing these aspects by these auditory icons could provide natural mapping to facilitate remembering and interpreting its meaning successfully by users. The two sounds are played five times for five different statements in the presented content. Also, these sounds are played in pause intervals of virtual lecturer speech in order to avoid interference between both sounds. On overall, three types of multimodal interaction metaphors are incorporated in the experimental e-learning interfaces as shown in table 1.

TABLE I. INCORPORATED MULTIMODAL INTERACTION METAPHORS

Interfaces	Text and Graphics (Visual)	Speaking Virtual Lecturer with Full Body Gestures (Audio-Visual)	Speaking Virtual Lecturer without Body (Audio)	Auditory Icons (Audio)
VSNS	X	X		X
RSNS	X		X	X
RSON	X		X	

With respect to the learning material presented by the three tested e-learning interfaces, it is the same. It is a single lesson that contains introductory information about class diagrams including class diagram notations, what is meant by class and object, and how to differentiate between them. This content has been adapted from [22].

#### IV. EXPERIMENTAL DESIGN

The experimental work described in this paper aimed to examine three different usability aspects of the experimental e-learning interfaces. These aspects are the effectiveness, memorability, and user satisfaction. The effectiveness is evaluated in terms of users' learning performance by the correctness of their answers to the learning activities prior and post experimentation with the tested e-learning interfaces. The memorability is evaluated by users' ability to recognize the sound (i.e. auditory icon) used and its meaning. However, the user satisfaction is assessed by users' responses to the satisfaction questionnaire. More specifically, the experiment aims to answer the following questions:

1) Which of the experimental e-learning interfaces is more usable than the others in terms of effectiveness, memorability and user satisfaction?

2) Which of the experimental e-learning interfaces is the most effective one in terms of users' learning performance?

3) Would the participants be able to remember the meaning of the incorporated auditory icons successfully?

4) Which of the experimental e-learning interfaces is the most satisfactory to the participants?

5) Which is better to use in e-learning interfaces; a speaking and full body animated virtual lecturer with supportive auditory icons or without it?

6) Which is better to use in e-learning interfaces; the speech of the virtual lecturer along with supportive auditory icons or without it?

#### B. Hypotheses

The main hypotheses in relation to the presented study are:

1) There will be a difference in the usability of the tested e-learning interfaces in terms of effectiveness, memorability and user satisfaction.

2) There will be a difference in the users' learning performance before and after interacting with the experimental e-learning interfaces.

3) There will be a difference in users' evaluation of the involved auditory icons before and after experimentation.

4) The participants will be able to correctly remember the investigated auditory icons.

#### C. Participants

In total, 45 participants involved voluntarily in this experiment and randomly assigned in equal proportions ( $N=15$ ) to one of three independent groups each of which tested one of the experimental interfaces; VSNS, RSNS and RSON. All the participants used the experimental interfaces for the first time and were undergraduate students enrolled in information technology programs at the Applied Science University, Jordan. Fig. 3 shows that most of the participants (93%) in the three groups were 18-23 years old and first year undergraduate students (87%-93%). With respect to their gender, the number of females and males participants was approximately equal. Furthermore, their area of study was distributed as follows: Software Engineering (SE) with 27% of each group, Computer Information Systems (CIS) and Computer Networks Systems (CNS) with 27%, 20% and 27% for VSNS, RSNS and RSON respectively, and Computer Science (CS) with 20% of VSNS, 33% of RSNS, and 20% of RSON. The selection of participants was mainly based on their previous knowledge in the learning topic; class diagram notation.

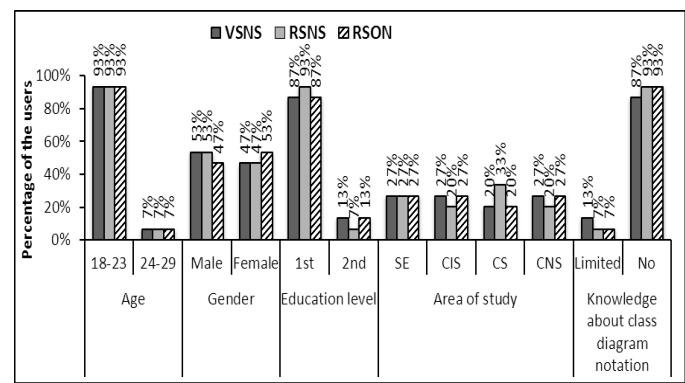


Fig. 3. Users' profiling.

In this regard, the majority of them (93%-100%) had no experience indicating that they relied on the communicated learning information to answer the learning performance questions after experimenting with the tested interfaces.

#### D. Procedure

The same procedure has been followed with each user throughout the course of the experiment and the participation was on an individual bases. The experiment started with user-profiling and pre-experimentation tasks. Also, users of VSNS and RSNS have been provided with a short training in which they had the opportunity to listen to the implemented auditory icons to insure their ability to understand and interpret each of these sounds when incorporated later in the experimental e-learning interfaces VSNS and RSNS. Next, the learning material about class diagram notation has been presented using one of the experimental e-learning interfaces. Once this presentation finished, the user has been instructed to carry out post-experimentation tasks as well as to provide any comments or suggestions.

#### E. Tasks

Prior to experimentation with the assigned interface, each user has been requested to provide personal data in relation to age, gender, educational level and major. In addition, they have been requested to declare their previous experience in Computers, Internet, class diagram notation and e-learning applications. Furthermore, users involved in testing both VSNS and RSNS interfaces have been asked to express their points of views in terms of annoying, focus and helpfulness regarding the use of auditory icons in electronic learning in the absence of any interactive context. Then, users have been asked to answer a set of four questions about the learning material adopted in this experiment (pre-test). Two of these questions are recall activities in which the user needs to retrieve some information to be able to answer. The other two questions are recognition activities that provide the user with multiple answers to the required question and he/she is required to recognize the correct one.

Once the experimentation finished, the user has been instructed to carry out learning performance (post-test), memorability, and satisfaction tasks as well as to provide any comments or suggestions. The learning performance tasks asked the user to answer the same question of the pre-test in relation to the learning content delivered by the tested interface in order to measure the effectiveness of that interface as well as to measure how much learning gained by users. The memorability task aimed to evaluate users' ability to identify the sounds used to communicate key aspects of the presented content. More specifically, three different sounds were played and the user had to recognize which one has been used to communicate the start or end of an important statement in the lesson. This task has been applied only with users of both VSNS and RSNS interfaces. The satisfaction task aimed to obtain users' attitude towards the tested interface. In this task, the user has been instructed to fill a satisfaction questionnaire of 14 different statements on a 5-point Likert scale. The first 10 statements were the SUS questionnaire [23] whereas the remaining 4 statements were related to learning experience. For

VSNS and RSNS users, the experiment finished with additional task to obtain their opinions with respect to the implemented auditory icons in terms of annoying, focus and helpfulness.

#### F. Variables

Three different types of variables have been considered which are: independent variables, dependent variables and controlled variables.

The independent variables represent the factors manipulated in the experiment and assumed to be the cause of the results. In this study, the presentation mode has been considered as the independent variable where the experimental e-learning interfaces offered three different modes for the presentation of the learning material. The first mode used text with graphics, a speaking virtual lecturer with full body gestures and auditory icons (VSNS interface). The second mode incorporated text with graphics along with the speech of virtual lecturer, and auditory icons (RSNS interface). However, the third mode included text with graphics and the speech of virtual lecturer (RSON interface).

The dependent variables are measured as a result of manipulating the independent variables. The dependent variables regarded in this study are as follows:

- Effectiveness: correctness of user's responses to the required learning performance activities; recall and recognition. In recall questions, partial or total correct answers have been considered whilst in the recognition questions, the answer had to be totally correct. The difference in users' learning performance between pre-test and post-test has been considered as well.
- Memorability: users' recognition of auditory icons has been measured by the number and percentage of users successfully recognized the non-speech sounds after being used in the experimental e-learning interfaces (VSNS and RSNS).
- User satisfaction: measured by users' responses to satisfaction questionnaire.

The controlled variables represent the external variables associated with the procedure of the experiment and could affect the obtained results. The controlled variables (known also as confounding variables) should be kept consistent throughout the experiment to avoid their influence on the dependent variables and so insure that the independent variables are the only cause of the experimental results. The controlled variables in this experiment are:

- Required tasks: the same tasks have been required from the participants.
- Presented learning material: the same learning content about class diagram notation has been presented by all experimental e-learning interfaces.
- Awareness of questions: none of the users were aware of the required learning performance questions in both pre and posttests.

- Procedure consistency: the same procedure has been followed during the execution of the experiment including measurement tools and used equipment.
- Familiarity with the interface: all participants were first-time users of the tested interfaces and provided with the same level of training prior to experimentation.

## V. RESULTS

The obtained experimental results have been analyzed in terms of different parameters including users' views regarding the auditory icons accompanied the virtual lecturer voice in both the absence and presence of interactive e-learning context. In addition, these parameters involved the effectiveness (learning performance), memorability and users' satisfaction. A more details are provided in the following subsections.

### A. Users' Evaluation of auditory icons

Before starting the interaction with the tested interface, users of both VSNS and RSNS have been requested to provide their opinions towards auditory icons when used to accompany the voice of virtual lecturer in e-learning interfaces. They have been asked to answer Yes or No if they think that it is annoying, could aid to focus, and helpful during the learning process. The same question has been repeated by the end of the experimentation. Fig. 4 shows that users' attitude positively changed towards auditory icons after being used in an interactive e-learning context within the tested VSNS and RSNS interfaces. The percentage of users who felt annoyed dropped from 67% to 27% and from 80% to 40% for VSNS and RSNS respectively which means that about half of them found it not annoying. This figure also demonstrates that the tested auditory icons did not substantially split users' attention away from the presented content where 73% of VSNS users and 67% of RSNS users believe that these sounds aided them to focus during the interaction compared to 20% and 7% respectively who think that it could enhance their concentration when it has been introduced in the absence of any interactive e-learning context. With respect to the helpfulness of the tested auditory icons in the learning process, it can be seen from Fig. 4 that users' opinion considerably changed after they tested the two interfaces. Prior to the experiment, 33% of VSNS users and smaller percentage (13%) of RSNS users thought that incorporating these non-speech sounds could help in enhancing their learning. This percentage increased remarkably to 80% (VSNS) and 73% (RSNS) after they have had the opportunity

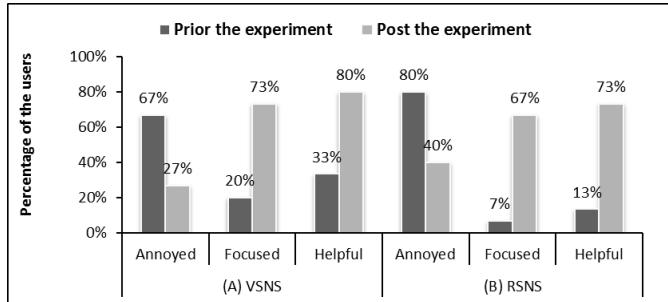


Fig. 4. Views of VSNS users (A) and RSNS users (B) about the tested auditory icons when used in both the absence and presence of interactive e-learning context.

to experience it interactively. In summary, the addition of auditory icons to the experimental e-learning interfaces was found to be neither annoying nor distracting and helpful to improve learning. These findings support the results of previous research [20].

### B. Learning Performance

One of the main concerns of the present study was the learning performance of the participants (effectiveness of the tested interfaces) which has been measured and compared in terms of the number of correctly answered questions related to the experimental content. Each user was required to answer the same 4 questions before (pretest) and after (posttest) interacting with the experimental interfaces, and therefore, the total number of questions in each case was 60. Fig. 5 shows the percentage of correctly answered questions by each group of users in both pretest and posttest. The obtained results demonstrated that the participants had a weak background about class diagram notation prior the experiment. It can be seen that the overall percentage of correct answers was 15% in VSNS and RSNS each compared to 10% in RSON. The participants achieved an average total score (i.e., the sum of correct answers out of 4) of 0.6 ( $SD = .63$ ) in condition VSNS, 0.6 ( $SD = .51$ ) in condition RSNS and 0.4 ( $SD = .63$ ) in condition RSON. As expected, Kruskal-Wallis test revealed that no significant differences have been found between the experimental conditions in the pretest ( $H(2) = 1.64$ ,  $p > .05$ ), which means that all participants were at the same low level of knowledge with respect to the content communicated later on by the experimental interfaces. Mann-Whitney tests were used to follow up this finding and revealed that no significant differences were found between RSNS and VSNS ( $U = 109.5$ ,  $r = -.26$ ), between RSNS and RSON ( $U = 87$ ,  $r = -.22$ ), and between VSNS and RSON ( $U = 91.5$ ,  $r = -.18$ ). These results were found to be consistent with users' profiles where most of them had no or limited previous knowledge about the experimental learning material (refer to Fig. 3).

Confirming what has been initially hypothesized, the posttest provided different results. The overall percentages of correct answers were 75%, 60% and 50% in RSNS, VSNS and RSON respectively. In other words, condition RSNS resulted in a higher average total score (3,  $SD = 0.65$ ) than condition VSNS (2.4,  $SD = 0.91$ ) which in turn achieved a higher average total score than the RSON condition (2,  $SD = 1.13$ ). The differences in posttest results were found to be significant according to Kruskal-Wallis test ( $H(2) = 8.42$ ,  $p < .05$ ) indicating that users' learning has been significantly affected by

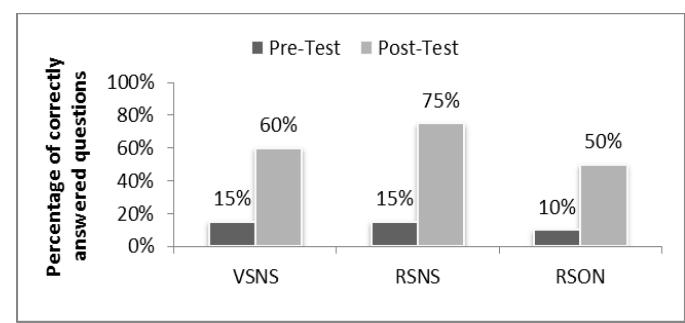


Fig. 5. Percentage of correct answers achieved by the users.

the way each of the tested interfaces used to deliver the learning content. Mann-Whitney follow up tests revealed that RSNS users significantly outperformed both VSNS users ( $U = 69$ ,  $r = -.36$ ) and RSON users ( $U = 51$ ,  $r = -.49$ ). However, no significant difference was found between VSNS users and RSON users ( $U = 82.5$ ,  $r = -.24$ ). Therefore, it can be concluded that if the learning material is presented using RSNS interface it will significantly increase users' learning performance (i.e., correctly answered questions) compared to presenting the same content using VSNS or RSON interfaces; however, users' learning performance will not be affected by VSNS or RSON interfaces.

For further analysis, the difference between posttest results and pretest results were computed to obtain a feedback on how much each of the tested interfaces affected on the learning of the users. It can be seen that the highest difference 60% was achieved by RSNS users followed by VSNS users (45%) and RSON users (40%).

Summarizing these findings, communicating the learning content using the speech of virtual lecturer in line with auditory icons (condition RSNS) could result in larger learning advantage in comparison to conditions VSNS and RSON.

### C. Remembrance of Auditory Icons

In order to make use of auditory icons in communicating auditory notifications related to key features of the learning content while being presented by the virtual lecturer, users should be able to successfully recognize it and interpret its meaning correctly. Therefore, by the end of interacting with the tested interfaces, users of both RSNS and VSNS were requested to carry out the memorability task to measure their remembrance of the tested auditory icons. Three sounds (auditory icons) have been played for each of the two key features and the user had to recognize which sound has been used to communicate which feature. This task has not been requested from RSON users because this interface did not incorporate any use of auditory icons. Fig. 6 shows the correctness rate of users' responses to this task. Although users of RSNS performed better, it can be seen that users of both interfaces achieved a high percentage of correct recognition. On overall, 97% of the tested auditory icons were correctly recognized by users of RSNS compared to 83% by users of VSNS. Also, users of RSNS were more able to correctly recognize the start of statement sound "opening a door" (100%) and the end of statement sound "closing a door" (93%) than users of VSNS (93% and 7% for "opening a door" and "closing a door" respectively). In brief, these results demonstrate that the tested auditory icons could be successfully interpreted and remembered by the users when used to indicate the importance of specific content delivered by a virtual lecturer in e-learning interfaces. This was found to be consistent with the findings of previous experiment [20] were similar results have been attained but with different experimental design in which the same auditory icons were used and tested in a single interface in addition to other auditory icons.

### D. User Satisfaction

At the end of experiment, users were required to respond to the satisfaction questionnaire composed of 14 statements each

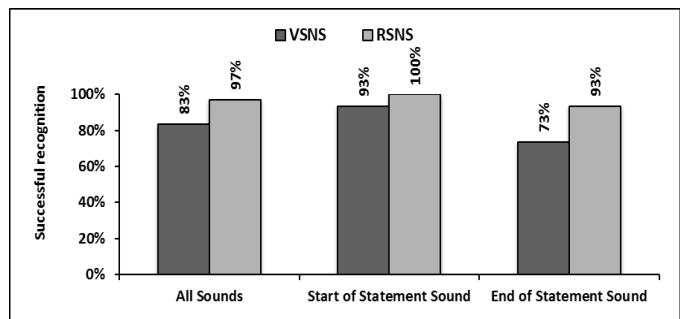


Fig. 6. Users' successful recognition of the tested auditory icons.

of which had a 5-point Likert scale ranging from 1 representing strong disagreement to 5 representing strong agreement. The first 10 statements were adopted from SUS questionnaire [23] to obtain users' attitude towards the different aspects of the tested interfaces. For the analysis of results, the SUS scoring method has been used for the SUS statements, whereas the mode and median were calculated for the other 4 statements.

Findings showed that the RSNS interface scored the highest SUS satisfaction score ( $M = 84.17$ ,  $SD = 21.79$ ) compared to VSNS ( $M = 79.5$ ,  $SD = 25.88$ ) and RSON ( $M = 60.5$ ,  $SD = 13.63$ ). An analysis of variance (ANOVA) yielded significant difference in users' attitude towards the three interfaces ( $F(2,1) = 5.32$ ,  $p < .05$ ). Also, the results of follow up pairwise comparisons found significant satisfaction difference between both RSNS and RSON ( $p < .05$ ) and between VSNS and RSON ( $p < .05$ ), however, not between RSNS and VSNS ( $p > .05$ ). In other words, the interfaces that incorporated auditory icons were more satisfactory to the users than the interface which didn't use this kind of non-speech sounds.

In addition to the SUS statements, another 4 statements were included to obtain feedback from users regarding their learning experience attained during the interaction with the tested interfaces. More specifically, these statements investigated users' excitement and interest about the presented content (*S11- I was excited and interested about what has been presented in the lesson*), ease of identifying important parts of this content (*S12- It was easy to identify the important parts of the presented lesson*), and user's willing to use e-learning if presented similar to the tested interface (*S13- I would like to use e-learning once more if presented this way*). The last statement aimed to evaluate overall users' satisfaction (*S14- On overall, I am satisfied with the interface*). Users' responses to the additional four statements are shown in Fig. 7. The same level of users' ratings for S11, S12, and S13 statements could be observed in RSNS and VSNS with mode and median valued four where users of both interfaces expressed their agreement about these statements. However, users of RSON were neutral about the same statements. In other words, users of both RSNS and VSNS were more excited and interested, more capable to capture important parts of the content, and more willing to reuse these interfaces for e-learning compared to users of RSON. On overall (S14), users of RSNS were more satisfied than users of VSNS and users of RSON who were generally satisfied in spite of their neutral impressions regarding S11, S12, and S13. To summarize, both RSNS and VSNS provided the users with more enriching learning experience in comparison to RSON.

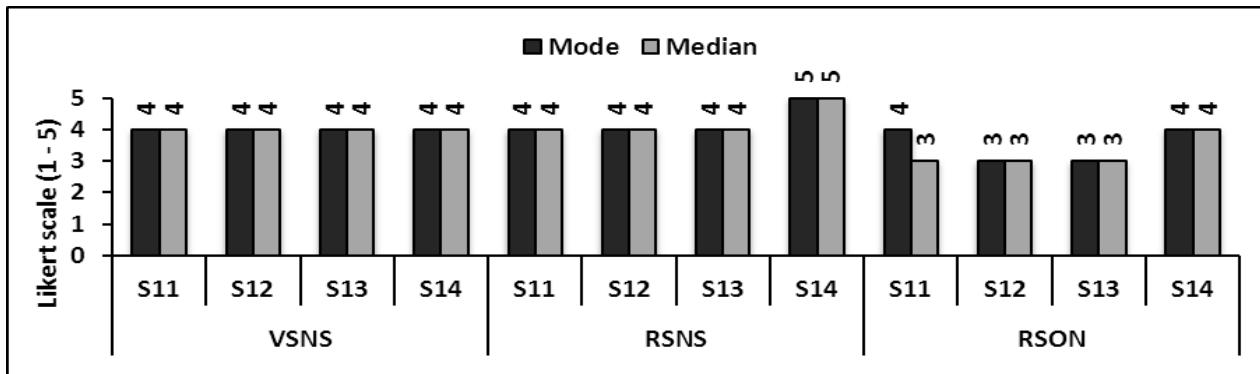


Fig. 7. Users' rating of the additional four statements S11, S12, S13, and S14 in the satisfaction questionnaire.

## VI. DESCUSSION

The experimental study reported in this paper investigated three different e-learning interfaces each of which incorporated different multimodal approach to communicate the learning content to the participants. The first interface (VSNS) involved the use of virtual lecturer's speech with the presence of animated body gesture and accompanied by the sounds of auditory icons. The second interface (RSNS) was similar to VSNS but with the absence of virtual lecture's body. However, the third interface (RSON) included only the speech of the virtual lecturer. Otherwise, the three interfaces were similar to each other in regards to visual appearance. Our goal was to identify among these interfaces which one provides better use of auditory icons in accompanying the virtual lecturer speech. The obtained results have been used to compare these ways of presentation in terms of effectiveness (learning performance), memorability and user satisfaction where the difference among the three experimental interfaces with respect to these usability attributes has been predicted in the research hypotheses.

The experimental results revealed that the tested interfaces were significantly different in terms of users' learning performance as well as users' satisfaction. In addition, the results of multiple comparisons among the three interfaces demonstrated that the RSNS was the most effective in communicating the learning content to the participants and this has been reflected on their ability to achieve the highest learning performance in terms of correctly answered questions. Also, RSNS was found to be the most satisfactory presentation to the participants. Presenting the learning content using the RSNS interface enabled the users to fully concentrate on the delivered content shown on the screen and explained by the speech of virtual lecturer and as a result enhanced their understanding of the presented information. At the same time, using auditory icon sounds contributed to capture users' attention to the important statements spoken by the lecturer and helped them to identify the most important parts of the learning content particularly if we know that most of the participants who tested RSNS stated that these sounds did not annoy them, helped them to focus and were helpful in their learning (see Fig. 4). This can be attributed to the fact that using "opening a door" and "closing a door" sounds helped the users to establish natural mapping between the communicated information and familiar sounds from everyday life and each of these sounds

transmit only one meaning and used consistently throughout the tested interfaces. As well, users of RSNS were more able to remember these sounds compared to the users of VSNS (see Fig. 6). This has been supported by users responses to the satisfaction questionnaire where RSNS users found themselves excited and interested about what has been presented, capable to easily identify the important parts of content, and like to learn from similar e-learning interfaces (see Fig. 7). As a result, they were significantly more satisfied compared to users of VSNS and RSON.

On the other hand, the VSNS interface came in the second place in terms of the investigated usability attributes. Previous research [18] has proven that e-learning interface which include virtual lecturer speaking the learning material with the presence of full body gesture along with supportive auditory non-speech messages (similar to VSNS) could significantly enhance the usability of e-learning interfaces. That research, however, tested only that interface by one group of users without comparing it with another interface that make use of virtual lecturer and non-speech sounds such as RSNS. The findings of the current study demonstrated that VSNS-like interfaces performed lower in usability evaluation when compared to RSNS. Although the VSNS interface enabled the users to be engaged in learning environment similar to the real face-to-face interaction take place in the traditional class rooms, it seems that the presence of full body animation contributed to split users' attention away from the presented learning content and overloaded their visual channel moving their eyes between two visual metaphors; virtual lecturer and his background content, and as a result they achieved lower learning performance (see Fig. 5) giving that they expressed positive impressions about the used auditory icons (see Fig. 4) and were able to remember it successfully (see Fig. 6) as well as were satisfied with the interface and attained learning experience (see Fig. 7) like RSNS users.

Similar to RSNS, the RSON interface enabled the users to hear the spoken explanations of the presented content and watching that content at the same time which contributed to reduce users' visual overload and keeping them involved better in cognitive processing of the communicated content compared to VSNS. Even though, users of RSON attained the lowest number of correctly answered questions. This can be attributed to the contribution made by the auditory icons used in RSNS

and VSNS which improved users' concentration and attention towards the presented content. Also, users of RSON were found to be significantly less satisfied comparable to VSNS and RSNS users. On overall, the obtained results confirmed the experimental hypotheses and suggest that combining auditory icons with virtual lecturer speech in the absence of body gestures is much better in enhancing the usability and learning performance of e-learning interfaces than combining auditory icons with virtual lecturer speech in the presence of body gesture.

## VII. CONCLUSION

This paper described further empirical investigation into the use of auditory icon to communicate supportive auditory messages related to the learning content while being delivered by the virtual lecturer speech. The main aim of this investigation was to identify the best combination of these two metaphors when incorporated in e-learning interfaces in addition to other visual metaphors. In order to achieve this aim, three different e-learning interfaces have been built and experimentally tested by three independent groups of users each of which examined one of the experimental e-learning interfaces in terms of usability attributes; effectiveness (learning performance), memorability of auditory icons, and user satisfaction. The obtained results revealed that incorporating auditory icons with the speech of full-body animated virtual lecturer could enhance the usability of e-learning interfaces much better compared to the remaining tested combinations of multimodal interaction metaphors.

## REFERENCES

- [1] S. Alexander, "E-learning developments and experiences," *Education and Training*, vol. 43, pp. 240-248, 2001.
- [2] D. Yu, 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.
- [3] R. Hamilton, C. Richards, and C. Sharp, "An examination of e-learning and e-books," 2001.
- [4] F. Mikic 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*, pp. 217-222, 2006.
- [5] A. P. Correia 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*, pp. 521-528, 2001.
- [6] G. Theonas, D. Hobbs, and D. Rigas, "Employing Virtual Lecturers' Facial Expressions in Virtual Educational Environments," *International Journal of Virtual Reality*, vol. 7, pp. 31-44, 2008.
- [7] W. L. Johnson, S. Kole, E. Shaw, and H. Pain, "Socially intelligent learner-agent interaction tactics," *Proceedings of the International Conference on Artificial Intelligence in Education*, 2003.
- [8] N. B. Sarter, "Multimodal information presentation: Design guidance and research challenges," *International Journal of Industrial Ergonomics*, vol. 36, pp. 439-445, 2006
- [9] A. Dix, G. Abowd, J. Finlay, and R. Beale, *Human-Computer Interaction (3rd Edition)*. Prentice Hall, 2004
- [10] R. Sheth, "Avatar Technology: Giving a Face to the e-Learning Interface," *The eLearning Developers' Journal*, 2003.
- [11] J. Beskow, "Animation of Talking Agents," *Proceedings of AVSP*, vol. 97, pp. 149-152, 1997.
- [12] M. Alotaibi and D. Rigas, "The Role of Avatars with Facial Expressions to Communicate Customer Knowledge," *International Journal of Computers, NAUN*, vol. 3, pp. 1-10, 2009.
- [13] J. Holmes, "Designing agents to support learning by explaining," *Computers & Education*, vol. 48, pp. 523-547, 2007.
- [14] R. Moreno and R. Mayer, "Interactive Multimodal Learning Environments," *Educational Psychology Review*, vol. 19, pp. 309-326, 2007.
- [15] L. A. Annetta and S. Holmes, "Creating Presence and Community in a Synchronous Virtual Learning Environment Using Avatars," *International journal of instructional technology and distance learning*, vol. 3, pp. 27-43, 2006.
- [16] D. Rigas and A. Algahtani, "An Investigation of Multimodal Metaphors in E-Book Assessment Interfaces", *HCI International 2013-Posters' Extended Abstracts*, 567-571, 2013.
- [17] A. Ciuffreda and D. Rigas, "A usability Study of multimodal interfaces for the presentation of Internet Search Results," *International Journal of Computers, NAUN*, vol. 2, pp. 120-125, 2008.
- [18] D. Rigas and B. Almutairi, "Investigating the impact of combining speech and earcons to communicate information in e-government interfaces", *Human-Computer Interaction. Interaction Modalities and Techniques*, 23-31, 2013.
- [19] M. Alseid and D. Rigas, "Three Different Modes of Avatars as Virtual Lecturers in E-learning Interfaces: A Comparative Usability Study", *Open Virtual Reality Journal, Bentham Open, ISSN: 1875-323X*, volume 2: p.p. 8-17, 2010.
- [20] M. Alseid and D. Rigas, "The Role of Earcons and Auditory Icons in the Usability of Avatar-Based E-Learning Interfaces", *Proceedings of the 4th International Conference on Developments in E-Systems Engineering*, p.p. 276-281, 2011
- [21] W. W. Gaver, "The SonicFinder: An Interface That Uses Auditory Icons," *Human-Computer Interaction*, vol. 4, pp. 67-94, 1989.
- [22] T. C. Lethbridge and R. Laganiere, *Object-oriented software engineering*: McGraw-Hill Education, UK, 2001.
- [23] J. Brooke, "SUS: a " quick and dirty" usability scale," *Usability evaluation in industry*, pp. 189-194, 1996