Glosses, Spatial Intelligence, and L2 Vocabulary Learning in Multimedia Context

AFSAR ROUHI
English Department, Payame Noor University, Iran
afsar.rouhi@pnu.ac.ir

HASSAN MOHEBBI
English Department, Payame Noor University, Iran

ABSTRACT

L2 research has consistently shown that providing glosses enhances L2 vocabulary learning. This paper reports on a study that examined the effect of pictorial, pictorial + sound, and video glosses on L2 vocabulary learning in multimedia context. More specifically, we investigated whether learners' high spatial intelligence ability has any effect on their ability to take advantage of multimedia glosses in L2 vocabulary learning. Sixty-two Azari-Turkish pre-university students were randomly assigned to three experimental groups and a control group. The experimental groups read the reading sections of three units of their textbook for six sessions using Scaffoglossing software and consulted one of the glosses, i.e., pictorial, pictorial + sound, or video glosses. All groups took a Persian equivalent test, a multiple-choice test, and a sentence completion test once as the immediate posttest and 25 days later as the delayed posttest. Data analyses revealed the positive effect of multimedia glosses on L2 vocabulary learning. However, there was no significant difference among the participating experimental groups. Contrary to our expectations, no significant difference was observed between the high and low spatial ability groups. The findings and their implications are discussed in detail.

Keywords: multimedia glosses; spatial intelligence; cognitive theory of multimedia learning; involvement load hypothesis; scaffoglossing software

INTRODUCTION

Recently the issue of using glosses for enhancing second language (L2) vocabulary learning is receiving a great amount of attention and findings of emerging body of L2 research have provided robust evidence for the positive effect of glosses on prompting L2 vocabulary learning (Kim & Gilman 2008, Nagata 1999, Yanguas 2009, Yoshii 2006).

A gloss used to be limited to a definition or meaning of a key word provided in the first or second language which was given in the margin of a text. Nowadays, it can be presented as pictorial, video, or interactive multimedia elements. Providing glosses is argued to be a bottom-up lexical help, which provides direct support for L2 vocabulary learning (Gettys, Imhof, & Kautz 2001). More specifically, glosses can be viewed as focus on form since they possess the feature of drawing learners' attention to lexical items within a communicative task environment (Laufer & Girsai 2008).

Lenders (2008) identified three types of glosses, namely dictionary-type glosses, ready-made glosses, and special types of glosses. Dictionary-type glosses contain information about the meaning/s of a word in the form of a definition, synonym, antonym, L1 equivalent, phonetic script, or example sentences. Ready-made glosses are especially prepared for the particular needs of learners in a given course and they may contain a spoken or written L2 definition, an L1 translation, or a still or moving image depicting the glossed word. And, the third set of glosses is called special types, which involves a task for the learner as well as elaborating on the targeted words like multiple-choice glosses.

Nagata (1999) attributed the effectiveness of glosses on fostering L2 vocabulary learning and reading to four factors: they, as she asserts, are easier to use and less time-consuming than dictionaries, they draw learners' attention to targeted words which supports the notion of 'consciousness-raising' and 'input enhancement', they help connect words to meanings immediately, contributing to the 'meaning-form connection' approach. In addition, they encourage learners to move back and forth between the targeted words and glosses, stimulating them to perform lexical processing. In other words, modifying authentic L2 readings by employing glosses seems to be effective for enhancing L2 reading comprehension and vocabulary learning (Abraham 2008). Additionally, providing glosses might lead to L2 learners' greater use of authentic materials (O'Donnell 2013). However, there is a gap in second language acquisition (SLA) literature with respect to the effect of glosses on L2 vocabulary learning in multimedia context and the mediating factors which might enhance/impede the effectiveness of multimedia glosses. The present study extends the research on glosses and L2 vocabulary learning by examining the effect of multimedia glosses (i.e., pictorial, pictorial + sound, and video glosses) on triggering L2 vocabulary learning in English as a foreign language (EFL) context. Furthermore, as spatial intelligence of learners might influence their ability to take advantage of multimedia learning this study sets out to investigate the effect of learners' spatial intelligence on L2 vocabulary learning in multimedia context.

MULTIMEDIA GLOSSES AND L2 VOCABULARY LEARNING

To date, a majority of studies investigating the effect of glosses on enhancing L2 vocabulary learning and reading comprehension (Ko 2012, Kost, Foss, & Lenzini 1999, Rouhi & Mohebbi 2012, Yang, Yu, & Sun 2013) have confirmed the positive effect of glosses–operationalized in different forms–on enhancing L2 vocabulary learning and reading comprehension. There are also a few empirical studies which found no significant effect of glosses on L2 vocabulary learning and reading comprehension (Ariew & Ercetin 2004, Sakar & Ercetin 2005). They argued that glosses might be distracting for intermediate and advanced L2 learners. In a recent study, O'Donnell (2013) examined L2 learners' use of glosses through a quantitative and qualitative analysis of audiotaped think-aloud protocols. Interestingly, she observed unanticipated use of glosses, namely employing wrong glosses, misreading correct glosses, and not being able to understand the glosses.

Glosses can be provided as simultaneous representation of meaning of a targeted lexical item with a picture depicting it. There is convincing evidence to suggest that people learn more from words and pictures than from words alone (Guichon & McLornan 2008). Paivio's (1986, as cited in Lin & Chen 2007) dual-coding theory provides support for the use of visuals learning. According to this theory, words and sentences are processed and encoded in the verbal system, while images are processed and encoded in the nonverbal system. It is claimed that the simultaneous use of two channels of information delivery aids the retrieval of information and results in better learning (Gyselinck, Jamet, & Dubois 2008).

Cognitive theory of multimedia learning (CTML) is one of the theoretical underpinnings of this study. As it is clear in Figure 1, this model specifies five cognitive processes in multimedia learning: selecting relevant words from the presented text or narration, selecting relevant images from the presented illustrations, organizing the selected words into a coherent verbal representation, organizing selected images into a coherent pictorial representation, and integrating the pictorial and verbal representations and prior knowledge (Mayer 2005). For meaningful learning to occur, the learner must carry out all these cognitive processes (Mayer & Moreno 2002a). Mayer and Moreno (2002b) claimed that multimedia presentations are more likely to result in meaningful learning than single-medium presentations.

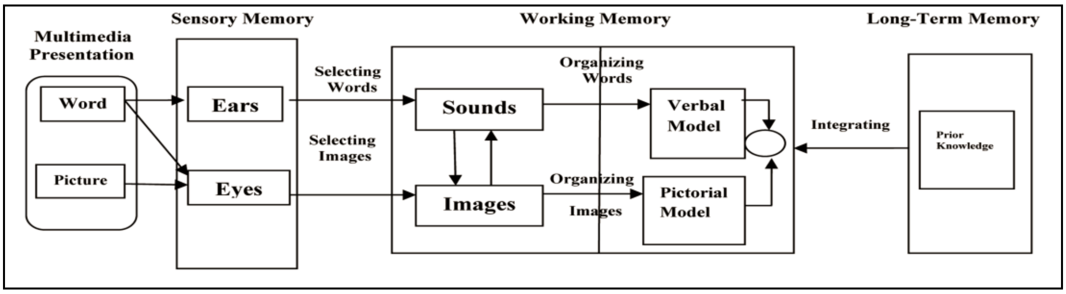

FIGURE 1. Cognitive theory of multimedia learning (CTML). (Mayer 2005, p. 37)

In their study, Chun and Plass (1996a) found significantly higher scores for words that were glossed with text coupled with pictures than for those with text + video or text only. In a similar vein, findings of studies conducted by Kim and Gilman (2008), Yeh and Wang (2003), Yoshii and Flaitz (2002), Akbulut (2007), and Shahrokni (2009) provide robust evidence for the positive effect of textual + pictorial glosses on enhancing L2 vocabulary learning.

In a similar line of research, Yanguas (2009) investigated the effect of textual, pictorial, and textual + pictorial glosses on L2 text comprehension and vocabulary learning. Results of quantitative and qualitative analyses of the data demonstrated that all multimedia glosses groups performed better than the control group and the textual + pictorial glosses group outscored all the other groups which supported the CTML.

In a more recent research, Turk and Ercetin (in press) examined the effects of interactive versus simultaneous display of visual and verbal multimedia information on L2 vocabulary learning and reading comprehension of low-proficient learners of English. They found that the simultaneous display of visual and verbal information, i.e., definitions + pictorial glosses, led to better performance on reading and vocabulary tests. In conclusion, findings of L2 research in this field confirm the validity of CTML.

SPATIAL INTELLIGENCE

In the early 1980s, Gardner proposed the theory of multiple intelligences (MI) that postulates intelligence to consist of various specific modalities rather just one single ability. Each individual reveals varying levels of these intelligences. Intelligence is argued to include linguistic intelligence, logical-mathematical intelligence, musical intelligence, intrapersonal intelligence, interpersonal intelligence, bodily-kinesthetic intelligence, spatial intelligence, natural intelligence, and existential intelligence. No single type of intelligence is viewed as being superior to the others (Akbari & Hosseini 2008).

Spatial intelligence involves visualization of things or ideas, through which we can retain memories for a longer period of time. It enables us to grasp meanings better when they are coupled with visual images (Kim 2009). Learners with a keen spatial intelligence are assumed to learn well with visual input such as multimedia, art, videos, movies, slides, and imaginative games (Armstrong 2009).

For learners with high spatial intelligence, Kim (2009) proposed computer-assisted language learning (CALL) activities such as playing card games, using graphic programs, learning with pictures on CD or DVD or with video clips on the Web, using presentation software, and creating videos or digital storytelling products. In sum, L2 learners' learning preference such as visualizers, i.e., high-spatial ability learners, or verbalizers seems to be a critical variable in the effectiveness of multimedia glosses (Yun 2011).

Liu et al. (2012) investigated whether verbalizers can learn as well as visualizers in simulation-based computer-assisted learning with predominantly visual presentations. In marked contrast to the above-mentioned premise, the data analysis revealed no significant difference between visualizers and verbalizers in terms of using dynamically linked multiple representations, their methods of reading, and their learning strategies.

Cognitive load of different kinds of glosses is another factor which plays a key role in the effectiveness of multimedia glosses on L2 vocabulary learning. Due to limited capacity of human working memory which imposes restriction in learning novel information, L2 learners might face challenges in simultaneous processing of different modalities of input, namely textual glosses presented with pictorial, pictorial + sound, or video glosses.

Wallen, Plass, and Brunken (2005) found that multiple types of glosses resulted in higher cognitive load which caused lower performance; this detrimental effect appeared to be stronger for low-verbal ability learners than high-verbal ability learners.

THE STUDY

Building upon the studies conducted on L2 vocabulary learning to date, the present study investigated the effect of multimedia glosses on L2 vocabulary learning. Moreover, we aimed to investigate whether spatial intelligence exerts any influence on the effectiveness of multimedia glosses. Studies conducted to date have shown positive effect of glosses on L2 vocabulary learning but limited research regarding multimedia glosses and the mediating effect of individual differences on L2 vocabulary learning in multimedia context has been carried out so far. SLA research has revealed that spatial ability exerts mediating effect on learning through multimedia glosses but there is inconclusive evidence with controversial results regarding this issue which calls for further empirical L2 research (Turk & Ercetin, in press).

RESEARCH QUESTIONS

With respect to multimedia glosses, spatial intelligence, and L2 vocabulary learning, we formulated the following two research questions for our study:

1. Do different types of multimedia glosses, including pictorial, pictorial + sound, and video, differ in their effectiveness on L2 vocabulary learning and retention?
2. Is there any difference between high and low-spatial ability learners' L2 vocabulary learning and retention gained from multimedia glosses?

METHOD

PARTICIPANTS

A total of 62 male pre-university students from one of the pre-university centers in Ardabil, Iran, participated in this study. They were 18–20 years old, bilingual in Azari-Turkish and Persian. They had studied in the same school and had received 6 years of formal English instruction. Participants reported that they were familiar with the basics of computer. English classes were held two sessions per week. The grammar translation method is usually employed in English classes in Iranian state schools. There is little focus on students' speaking skill and the main focus is on grammar, reading, and a limited number of lexical items. Participants in this study were randomly assigned to four groups: pictorial glosses (n = 17), pictorial + sound glosses (n = 17), video glosses (n = 16), and no glosses (n = 12).

MATERIALS

READING MATERIALS AND TARGETED WORDS

The texts were adapted from units 6, 7, and 8 of the pre-university textbook (Birjandi et al., 2010) simply because the study was conducted in the second semester of academic year, during which the students were supposed to study these units. Thirty glossed words were chosen after consulting several high school and pre-university English teachers and students. The targeted lexical items appeared only once in the text. The targeted lexical items were verbs (e.g., launch, forward, marry) and nouns (e.g., astronauts, slum, convent, planet, mountains).

MULTIMEDIA GLOSSES SOFTWARE (SCAFFOGLOSSING)

We designed a multimedia software, *Scaffoglossing* with the help of three computer experts. In designing this software, the instructional and multimedia principles such as coherence, modality, spatial contiguity, and temporal contiguity (Mayer 2005) and technological principles such as screen design, display location, navigation, learner control, colour, graphics, simplicity, functionality, objective effectiveness, and technological efficiency (Al-Seghayer 2003, Bloch 2009) were intimately observed.

In an attempt not to interrupt the classes during the research, to move toward CALL normalization, and to attract students' attention, the students' textbook was used. The software included three units of participants' textbook along with three kinds of glosses: pictorial, pictorial + sound, and video glosses. The video glosses were short silent clips. By clicking on the targeted words (typed **bold face,** in different colour, and underlined) participants could use the glosses. We conducted a pilot study with learners from different age groups who were studying English at a private language school. Most of the participants in the pilot study were high school and pre-university students and some of them were studying in the same school where the study was conducted. This pilot study was conducted in an informal situation, where participants were asked to work with the software, share their opinions about the software, and infer the words which the pictures and videos implied. The pilot study revealed that most of the pictorial glosses and video glosses connoted the same meaning for the participants of the pilot study, but two of the pictures and one of the videos were not clear enough. So, we changed them and again we asked them in another session to evaluate them.

SPATIAL INTELLIGENCE QUESTIONNAIRE

A 10-item questionnaire in the form of five-point Likert scale measuring the spatial intelligence (McKenzie 1999) was used as the first instrument of the study. It was supposed to provide an index of participants' spatial intelligence. The overall internal consistency reported for the questionnaire was .89.

PRE-TEST

We compiled a pre-test to check the homogeneity of participants in terms of their reading comprehension and lexical and grammatical knowledge. It included vocabulary, grammar, and reading comprehension questions extracted from their textbook. The Cronbach alpha coefficient computed for the pre-test was .52.

IMMEDIATE AND DELAYED POSTTESTS

To assess the effectiveness of different glosses, namely pictorial, pictorial + sound, and video glosses participants took the immediate posttest at the end of the last session. The immediate and delayed posttests were the same and participants were required to take the delayed posttest four weeks after the immediate posttest. The package included two production tests and one recognition test. The 30 targeted lexical items were randomly assigned to production and recognition tests. The first test required participants to write Persian equivalents (i.e., L1 translation) to 10 of the targeted words; in the sentence completion test, they were asked to complete sentences with English words. In this test the first letter of each targeted word was given; and the third one was a multiple-choice test. The reliability of the multiple-choice test computed as Cronbach alpha coefficient was .54. The 30 targeted words were randomly divided to be assessed in the production and recognition posttests.

PROCEDURE

On the day before the treatment sessions, the four participating groups took the pretest and completed the spatial intelligence questionnaire. They were homogeneous in terms of their reading comprehension, lexical, and grammatical knowledge, $F(3, 58) = 2.014$, $p = .122$. The three experimental groups attended the computer site of the school to use the software designed for them. The treatment lasted eight sessions. Every session lasted 45 minutes; session 7 was an exception and lasted 90 minutes. We conducted the study during participants' regular class period. In the beginning of the first session, the way the software worked was demonstrated in full detail. Participants read the texts by using Scaffoglossing software and consulted specific glosses provided for them. Group 1 took advantage of pictorial glosses; group 2 were provided with picture + sound glosses, that is, they received the pronunciation of words besides getting the pictorial glosses; and group 3 received videos of the targeted words. Participants of the control group attended their regular class and were taught the reading comprehension sections of the textbook by their instructor. In the end, all participating groups took the immediate posttest. After a four-week interval they took the delayed posttest. Figure 2 depicts the design and timeline of the study.

Session →	1	2	3	4	5	6	7	(After 4 Weeks) 8
Treatment → **&** **Testing**	Pre-test + SIQ	Unit 6	Units 6 and 7	Unit 7	Units 7 and 8	Unit 8	Review of Units 6, 7, and 8 + Immediate posttest	Delayed posttest

FIGURE 2. Design and timeline of the study

RESULTS

The scores obtained from the immediate and delayed posttests were put into SPSS and 3 repeated measures ANOVAs were run to assess the differences across the four participating groups, the trend of change from the immediate posttest to the delayed posttest, and any possible interaction effect between the treatment and time.

PERSIAN EQUIVALENT TEST

The first ANOVA examined the effect of multimedia glosses on L2 vocabulary learning in the Persian equivalent test. Table 1 presents the means and standard deviations for this test and Figure 3 displays graphically the means of the participating groups.

TABLE1. Descriptive statistics for the Persian equivalent test

Treatment		Immediate Posttest		Delayed Posttest	
	n	*M*	*SD*	*M*	*SD*
Pictorial+ sound	17	9.47	1.50	8.82	1.46
Pictorial	17	7.47	2.47	8.58	1.97
Video	16	9.37	.88	9.06	1.12
Control	12	4.75	2.89	5.58	3.17
Total	62	7.98	2.67	8.19	2.32

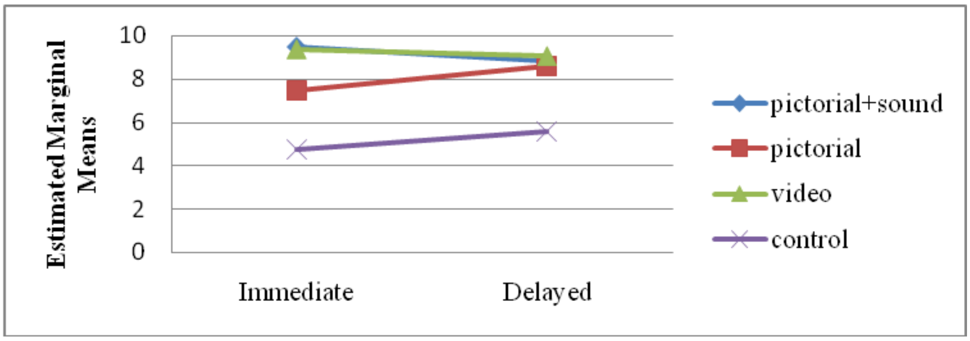

FIGURE 3. Performance of the participating groups in the Persian equivalent test

The output of the omnibus analysis showed that the main effect for the multimedia glosses condition was statistically significant with a large effect size, $F(3) = 14.24$, $p < .001$, $\eta_p^2 = .42$. A significant interaction effect was also observed between the treatment and time with a large effect size, $F = 4.32$, $p = .008$, $\eta_p^2 = .18$. However, the main effect for time was not significant, $F = 1.36$, $p = 024$, $\eta_p^{2 =}$ .02.

To statistically determine where the significant differences lay between the participating groups, a post-hoc (Tukey HSD) test was also run. In Persian equivalent test, the difference between the video glosses group and pictorial + sound glosses group reached statistical significance, $p < .001$. As it is clear in Table 1, in the Persian equivalent test the video glosses group performed better than the other groups.

MULTIPLE-CHOICE TEST

The second ANOVA examined the effect of multimedia glosses on L2 vocabulary learning in the multiple-choice test. Table 2 demonstrates the descriptive statistics, namely the means and standard deviations for the multiple-choice test. Figure 4 depicts the means of the participating groups.

TABLE 2. Descriptive statistics for the multiple-choice test

Treatment		Immediate Posttest		Delayed Posttest	
	n	*M*	*SD*	*M*	*SD*
Pictorial+ sound	17	8.05	1.81	6.58	2.42
Pictorial	17	7.88	1.93	6.29	2.33
Video	16	7.93	1.84	8.50	1.31
Control	12	6.25	2.59	5.66	2.26
Total	62	7.62	2.09	6.82	2.32

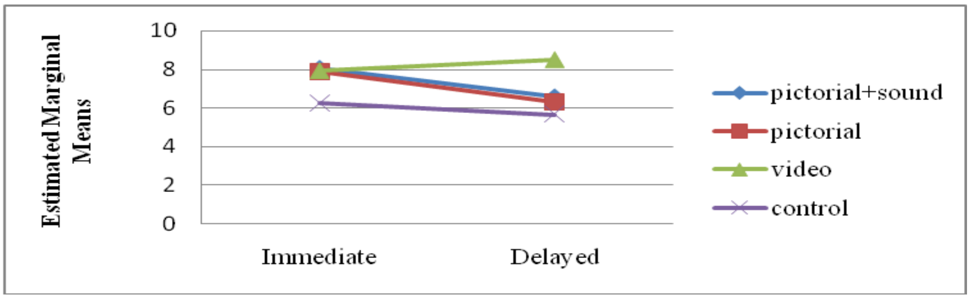

FIGURE 4. Performance of the participating groups in the multiple-choice test

The ANOVA test run showed that the effect of multimedia glosses was statistically significant with a moderate effect size, $F(3) = 3.63$, $p < .05$, $\eta_p^2 = .14$. There was a significant interaction effect between the treatment and time with a large effect size, $F = 5.006$, $p < .01$, $\eta_p^2 = .20$. Also the effect for time was significant with a moderate effect size, $F = 11.15$, $p < .01$, $\eta_p^2 = .16$.

To pinpoint the exact location of the difference observed, a post-hoc (Tukey HSD) test was also run. The difference between the video glosses group and the control group was statistically significant. The difference between the participating experimental groups did not reach statistical significance. Although their difference was not statistically significant, as Table 2 shows, in the multiple-choice test the participants in the video glosses group performed better than the participants of the other groups involved.

SENTENCE COMPLETION TEST

The third ANOVA was run on the data obtained from the second production test in which participants were asked to complete the sentences with appropriate targeted words. Table 3 shows the descriptive statistics for the sentence completion test. The means of participating groups have been presented graphically in Figure 5.

TABLE 3. Descriptive statistics for the sentence completion test

Treatment		Immediate Posttest		Delayed Posttest	
	n	*M*	*SD*	*M*	*SD*
Pictorial+ sound	17	6.64	1.41	4.70	1.21
Pictorial	17	7.88	1.53	5.29	1.57
Video	16	6.37	1.62	7.12	2.02
Control	12	2.66	1.43	3.75	1.81
Total	62	6.14	2.33	5.30	2.02

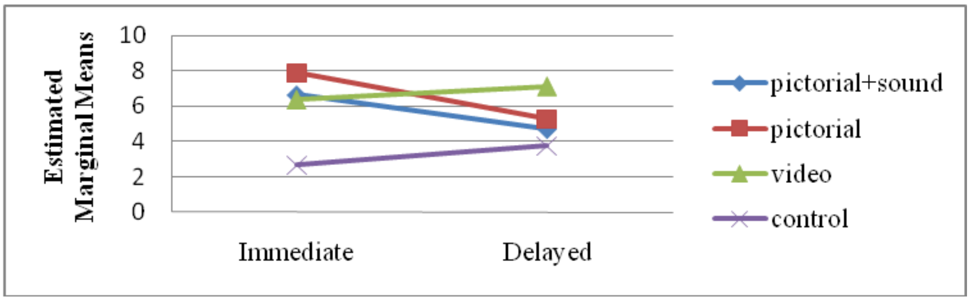

FIGURE 5. Performance of the participating groups in the sentence completion test

The output of the omnibus analysis of the scores of the sentence completion test showed that the main effect for the multimedia glosses was statistically significant with a sizeable effect size, F (3) = 18.40, $p < .001$, $\eta_p^2 = .48$. There was a significant interaction effect between the treatment and time with a substantial effect size, $F = 21.28$, $p < .001$, $\eta_p^2 = .52$. Also, the effect for time was significant, $F = 11.19$, $p < .001$, $\eta_p^2 = .16$.

A post-hoc (Tukey HSD) test was also run to reveal the exact location of the difference observed. Although the difference between the participating experimental groups was not statistically significant, in the sentence completion test the video glosses group showed better performance than the other participating groups in terms of learning the targeted lexical items.

RESULTS OF SPATIAL INTELLIGENCE

In the next stage, the participants of the pictorial, pictorial + sound, and video glosses groups were ranked according to their spatial intelligence indices and were divided into high and low spatial ability learners. A *t*-test was run to check the difference between the two groups in terms of their spatial intelligence index. There was statistically significant difference between the high and low spatial ability groups with a large effect size, $F(48) = 9.97$, $p < .001$, $\eta_p^2 = .67$.

Three repeated measures ANOVAs were run to assess the differences between the high and low spatial ability groups in the immediate and delayed posttests. There was no statistically significant difference between the two groups in the immediate and delayed Persian equivalent test, $F(1) = .63$, $p = .42$, $\eta_p^2 = .01$, the multiple-choice test, $F(1)= .25$, $p = .61$, $\eta_p^2 = .005$, and the sentence completion test, $F(1) = .20$, $p = .65$, $\eta_p^2 = .004$.

In the Persian equivalent and multiple-choice tests, the effect for time was not significant. Neither was there any significant interaction effect between the treatment and time. In the sentence completion test, the effect for time was significant with a large effect size, $F = 18.46$, $p < .001$, $\eta_p^2 = .27$. There was no significant interaction effect between the treatment and time.

DISCUSSION

Before turning to the research questions formulated, it is worth mentioning that the findings of this study, together with the studies reviewed (Chun & Plass 1996a, 1996b, Yanguas 2009, Yoshii 2006, Yoshii & Flaitz 2002), provide support to the positive effect of glosses on L2 vocabulary learning.

Research question 1: Do different types of multimedia glosses, including pictorial, pictorial + sound, and video, differ in their effectiveness on L2 vocabulary learning and retention?

The pictorial + sound group performed better than the pictorial group in the Persian equivalent and multiple-choice tests; on the theoretical side, this finding provides support for CTML, suggesting that providing verbal and visual input leads to meaningful learning. This finding is in line with the previous studies supporting the dual coding theory and CTML. Nevertheless, unlike the previous research which studied the pictorial coupled with textual glosses, we provided the audio of the words. Our finding is in contrast with Yeh and Wang (2003). They did not find any significant effect of the addition of sound on glosses for vocabulary learning. In a similar vein, Kim and Gilman (2008) argued that the inclusion of sound is distracting. We cannot, however, compare our finding with theirs because in their study the sound was acting as a third input besides textual and pictorial glosses.

Contrary to our expectations, in the sentence completion test the pictorial group performed better than the pictorial + sound group. This contradictory result might be due to the tests given, not the treatment, i.e., glosses. It can be argued that our testing package did not match the treatment given.

In general, there can be three reasons for the failure of dual presentation (i.e., pictorial + sound glosses) to work better than single presentation (i.e., pictorial glosses) in this study; they include limited capacity of working memory assumption, split-attention effect, and the redundancy principle of CTML and cognitive load theory.

The capacity of working memory is severely limited which restricts the learner in the amount of novel information he or she can process in each channel simultaneously. Our participants' working memory might be overburdened by simultaneous processing of pictorial glosses coupled with sound, leading to their failure in processing them properly.

Secondly, it might be related to the split-attention effect; it occurs when two or more sources have to be processed simultaneously. The working memory load imposed by the need to mentally integrate the disparate sources of information splits attention and overburdens limited working memory capacity which is likely to interfere with learning (Low & Sweller 2005).

The third reason might be associated with the redundancy principle of CTML. According to this principle, redundant input interferes with learning. Redundancy occurs when the same piece of information is presented in more than one form which is in fact unnecessary (Sweller 2005). It stands to logic to assume that in our scaffoglossing software the sound of targeted lexical items was redundant.

Another noticeable finding of the present study is the effect of the video glosses. By and large, in the posttests given the video glosses group performed better than the other groups. This is in line with Al-Seghayer (2001) and in contrast to Akbulut (2007). Al-Seghayer (2001) argued that the video builds a better mental image, and creates more curiosity which leads to a great deal of concentration.

Viewed in light of the involvement load hypothesis (ILH), the reason for this result can be attributed to the higher rate of the involvement load imposed by video glosses. ILH holds that word learning and retention are dependent on the amount of mental effort or involvement that a task imposes (Hulstijn & Laufer 2001). Task-induced involvement is a motivational-cognitive construct with three task factors: need, search, and evaluation. '*Need'* concerns the necessity to comply with the task. '*Search'* entails attempts to determine the meaning of an unknown word, and '*evaluation'* involves passing judgment about the semantic and formal appropriateness of the word and context.

It can be argued that learners who were provided with video glosses invested greater amounts of time and processing effort which pushed them for deeper or more elaborate processing of video glosses. Video glosses, also, are supposed to be more motivational by prompting authentic, real-life, and social aspects to L2 learning classroom context.

Research question 2: Is there any difference between high and low-spatial ability learners' L2 vocabulary learning and retention gained from multimedia glosses?

One of the main advantages of multimedia learning is that one can tailor it to meet the diverse needs of learners in accordance with their different learning/cognitive styles. It was assumed that individual learners in the high spatial ability group would outperform those involved in the low spatial ability group in the multimedia context. Contrary to our expectations, no significant difference was observed between the high and low spatial ability groups in the posttests given.

Firstly, such a non-significant result can be related to the cognitive architecture of learners and in particular to the constraints associated with working memory. Due to the limitations inherent in working memory, we need to make a balance between participants' working memory capacity and the imposed cognitive load of tasks. In multimedia environments we have to take the cognitive load which different kinds of glosses impose on learners. The lack of any observed significant difference between our high and low spatial ability groups might be partially related to the design of our software and the glosses presented. Instructional design should reduce the extraneous cognitive load caused by inappropriate instructional procedures. Reducing the extraneous cognitive load is likely to free working memory capacity and enhance the germane cognitive load and facilitate learning (Sweller 2005). Demonstrations, animations, simulations, exploratory environments, applying various modes, and modalities can impose extraneous cognitive load (Kalyuga 2009). It can be assumed that in the present study the cognitive load imposed by the pictorial, pictorial + sound, and video glosses exceeded the limited capacity of the participants and individual learners in the high spatial ability group were not able to manage the cognitive load imposed, which resulted in their failure in surpassing those in the low spatial ability group in terms of L2 vocabulary learning. For getting assurance as to the efficiency of visual presentations we should measure cognitive load to ensure that the cognitive load imposed by tasks does not exceed learners' working memory capacity.

Secondly, it can be associated with this assumption that high spatial ability might lead to outperformance when instruction induces high levels of cognitive load, for example, when it presents complex visio-spatial materials. While the low spatial ability learners may not be able to process such high-load materials deeply, learners with higher spatial ability are assumed to have the cognitive capacity to benefit from them (Plass, Kalyuga, & Leutner 2010). Given this issue, our finding might be justified by the fact that the pictorial, pictorial + sound, and video glosses of our software did not warrant high levels of cognitive processing, in which the high spatial ability group might have been able to perform better than the low spatial ability group.

CONCLUSIONS

The current study set out to continue the line of L2 research attempting to attain a better understanding of the role of multimedia glosses in L2 vocabulary learning. The findings of this study, together with those of the studies reviewed, lend support to the positive effect of providing multimedia glosses on enhancing L2 vocabulary learning. We also observed that there was no significant difference between pictorial, pictorial + sound, and video glosses in terms of fostering L2 vocabulary learning in multimedia context. Additionally, contrary to our expectations, the high spatial ability group failed to outscore the low spatial ability group in taking advantage of multimedia glosses.

Surprisingly, our findings with respect to spatial intelligence did not provide support for adapting multimedia software, in particular multimedia glosses in accordance with learners' different learning/cognitive styles. This result is in line with Tight (2010). He argued that "specific learning style preferences, in and of themselves, are neither a boon nor a hindrance to L2 vocabulary learning. Rather, participants of various perceptual learning style preferences appear to be equally capable of lexical learning" (p.817). Therefore, future research is needed to shed more light on the role of individual differences in multimedia environments.

REFERENCES

Abraham, L. B. (2008). Computer-mediated glosses in second language reading comprehension and vocabulary learning: A meta-analysis. *Computer Assisted Language Learning, 21*(3), 199–226.

Akbari, R., & Hosseini, K. (2008). Multiple intelligences and language learning strategies: Investigating possible relations. *System, 36*, 141–155.

Akbulut, Y. (2007). Effects of multimedia annotations on incidental vocabulary learning and reading comprehension of advanced learners of English as a foreign language. *Instructional Science, 35*, 499–517.

Al-Seghayer, K. (2001). The effect of multimedia annotation modes on L2 vocabulary acquisition: A comparative study. *Language Learning & Technology, 5*(1), 202–232.

Al-Seghayer, K. (2003). Technological and pedagogical considerations for a more effective electronic glossary. *The Reading Matrix, 3*(1), 1–15.

Ariew, R., & Ercetin, G. (2004). Exploring the potential of hypermedia annotations for second language reading. *Computer Assisted Language Learning, 17*(2), 237–259.

Armstrong, T. (2009). *Multiple intelligences in the classroom.* Alexandria, Virginia: ASCD.

Birjandi, P., Mir Hassani, A., Anani Sarab, M. R., & Samimi, D. (2010). *Learning to read English for pre-university students.* Tehran: Publishing Company of Iranian Textbooks.

Bloch, J. (2009). The design of an online concordancing program for teaching about reporting verbs. *Language Learning & Technology, 13*(1), 59–78.

Chun, D. M., & Plass, J. L. (1996a). Effects of multimedia annotations on vocabulary acquisition. *The Modern Language Journal, 80*(2), 183–198.

Chun, D. M., & Plass, J. L. (1996b). Facilitating reading comprehension with multimedia. *System, 24*(4), 503–519.

Gettys, S., Imhof, L. A., & Kautz, J. O. (2001). Computer-assisted reading: The effects of glossing format on comprehension and vocabulary retention. *Foreign Language Annals, 34*, 91–106.

Guichon, N., & McLornan, S. (2008). The effects of multimodality on L2 learners: Implications for CALL resource design. *System, 36*, 85–93.

Gyselinck, V., Jamet, E., & Dubois, V. (2008). The role of working memory components in multimedia comprehension. *Applied Cognitive Psychology, 22*, 353–374.

Hulstijn, J. H., & Laufer, B. (2001). Some empirical evidence for the involvement load hypothesis in vocabulary acquisition. *Language Learning*, 51, 539–558.

Kalyuga, S. (2009). *Managing cognitive load in adaptive multimedia learning.* New York: Information Science Reference.

Kim, D., & Gilman, D. A. (2008). Effects of text, audio, and graphic aids in multimedia instruction for vocabulary learning. *Educational Technology & Society, 11*(3), 114–126.

Kim, I. S. (2009). The relevance of multiple intelligences to CALL instruction. *The Reading Matrix, 9*(1), 1–21.

Ko, M. H. (2012). Glossing and second language vocabulary learning. *TESOL Quarterly, 46*, 56–79.

Kost, C. R., Foss, P., & Lenzini, J. J. (1999). Textual and pictorial glosses: Effectiveness on incidental vocabulary growth when reading in a foreign language. *Foreign Language Annals, 32*(1), 89–97.

Laufer, B., & Girsai, N. (2008). Form-focused instruction in second language vocabulary learning: A case for contrastive analysis. *Applied Linguistics, 29*(4), 697–716.

Lenders, O. (2008). Electronic glossing-is it worth the effort? *Computer Assisted Language Learning, 21*(5), 457–481.

Lin, H., & Chen, T. (2007). Reading authentic EFL text using visualization and advance organizers in a multimedia learning environment. *Language Learning & Technology, 11*(3), 83–106.

Liu, T., Kinshuk, Lin, Y., & Wang, S. (2012). Can verbalizers learn as well as visualizers in simulation-based CAL with predominantly visual representations? Preliminary evidence from a pilot study. *British Journal of Educational Technology*. *43*(6), 965–980.

Low, R., & Sweller, J. (2005). The modality principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 147–158). New York: Cambridge University Press.

McKenzie, W. (1999). Multiple intelligences inventory. Retrieved from http://surfaquarium.com/MI/inventory.htm/

Mayer, R. E. (2005). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 31–48). New York: Cambridge University Press.

Mayer, R. E., & Moreno, R. (2002a). Aids to computer-based multimedia learning. *Learning and Instruction, 12*, 107–119.

Mayer, R. E., & Moreno, R. (2002b). Animation as an aid to multimedia learning. *Educational Psychology Review, 14*(1), 87–99.

Nagata, N. (1999). The effectiveness of computer-assisted interactive glosses. *Foreign Language Annals, 32*, 469–479.

O'Donnell, M. E. (2013). Second language learners' use of marginal glosses. *Foreign Language Annals, 45*(4), 543–563.

Plass, J. L., Kalyuga, S., & Leutner, D. (2010). Individual differences and cognitive load theory. In J. L. Plass, R. Moreno, & R. Brunken (Eds.), *Cognitive load theory* (pp. 65–87). New York: Cambridge University Press.

Rouhi, A., & Mohebbi, H. (2012). The effect of computer assisted L1 and L2 glosses on L2 vocabulary learning. *The Journal of Asia TEFL, 9*(2), 1–19.

Sakar, A., & Ercetin, G. (2005). Effectiveness of hypermedia annotations for foreign language reading. *Journal of Computer Assisted Learning, 21*, 28–38.

Shahrokni, S. A. (2009). Second language incidental vocabulary learning: The effect of online textual, pictorial, and textual pictorial glosses. *TESL-EJ, 13*(3), 1–17.

Sweller, J. (2005). The redundancy principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 159–168). New York: Cambridge University Press.

Tight, D. G. (2010). Perceptual learning style matching and L2 vocabulary acquisition. *Language Learning, 60*(4), 792–833.

Turk, E., & Ercetin, G. (in press). Effects of interactive versus simultaneous display of multimedia glosses on L2 reading comprehension and incidental vocabulary learning. *Computer Assisted Language Learning*. Retrieved from http:// www.tandfonline.com/action/showAxaArticles.

Wallen, E., Plass, J. L., & Brunken, R. (2005). The function of annotations in the comprehension of scientific texts: Cognitive load effects and the impact of verbal ability. *Educational Technology Research and Development, 53*(3), 59–72.

Yang, X., Yu, S., & Sun, Z. (2013). The effect of collaborative annotation on Chinese reading level in primary schools in China. *British Journal of Educational Technology, 44*(1), 95–111.

Yanguas, I. (2009). Multimedia glosses and their effect on L2 text comprehension and vocabulary learning. *Language Learning & Technology, 13*(2), 48–67.

Yeh, Y., & Wang, C. W. (2003). Effects of multimedia vocabulary annotations and learning styles on vocabulary learning. *CALICO Journal, 21*(1), 131–144.

Yoshii, M. (2006). L1 and L2 glosses: Their effects on incidental vocabulary learning. *Language Learning & Technology, 10*, 85–101.

Yoshii, M., & Flaitz, J. (2002). Second language incidental vocabulary retention: The effect of text and picture annotation types. *CALICO Journal, 20*(1), 33–58.

Yun, J. (2011). The effects of hypertext glosses on L2 vocabulary acquisition: A meta-analysis. *Computer Assisted Language Learning, 24*(1), 39–58.