# The effect of multimedia modes on L2 vocabulary learning

**Abbas Ali Zarei, and Mahboubeh Gilanian***

1- Assistant professor, Imam Khomeini International University, Qazvin
2- MA, Islamic Azad University, Takestan, Iran

***Corresponding author E-mail:** m.gilanian@yahoo.com

## Abstract

**The present study sought to investigate the effect of various multimedia modes on L2 vocabulary comprehension and production. To this end, 52 male and female B.A level students at Imam Khomeini International University in Qazvin majoring in English translation and English language teaching in three groups were selected. A general proficiency test was administered to homogenize the participants. The same movie was presented to each group with different combinations of multimedia: the first group received video, audio, and captions (VAC), the second group were presented with video and captions (VC), and the third group received video and audio (VA). At the end of the experimental period, the participants received vocabulary comprehension and production post tests. Two separate one-way ANOVA procedures were used to analyze the obtained data. The results indicated no significant differences among the effects of various multimedia combinations on L2 vocabulary comprehension and production. The findings of the present study may have implications for L2 learners and teachers.**

***Key words:*** audio, captions, multimedia, video, vocabulary learning

## Introduction

It is evident that input has great importance in second language learning. This implies that a greater level of attention needs to be paid to the modality of input in language learning (Sydorenko, 2010). In recent years, many researchers have considered the effect of multimedia materials on second language learning. With the increasing popularity of multimedia sources among the younger generation, one can hardly deny the effect of various aspects of multimedia on the learning of various language components and skills.

This claim is corroborated by several researchers who have shown the positive effect of using different kinds of multimedia on language learning. For instance, Danan (2004) and Wang (2012) state that captioned movies can assist L2 learners to recall second language vocabulary through combining images, spoken words, and written text.

Nonetheless, there are others who are against the use of multimedia in learning a second language because they think multimedia use may be accompanied by a number of difficulties. For example, Moreno and Mayer (2007) claim that using captioned movies is not useful for L2 language classroom because they can overload the visual-pictorial channel by presenting animations, verbal information, and words on-screen text at the same time.

The present study aims to partially resolve the controversy surrounding this issue by addressing the question of the extent to which various aspects of multimedia can affect L2 vocabulary learning. The effect of each aspect of a film on vocabulary comprehension and production is investigated under three treatment conditions: first, video, audio, and captions (VAC), second, video and captions (VC), and finally video and audio (VA). In simple terms, this study is an attempt to answer the following research questions:

- Are there any significant differences among the effects of various multimedia modes on L2 vocabulary comprehension?
- Are there any significant differences among the effects of various multimedia modes on L2 vocabulary production?

According to Plass and Jones (2005), multimedia is a combination of print, audio, and imagery that enhance input by making it more comprehensible. They claim that pictures and video can improve reading

and listening comprehension. Similarly, Paivio (1986, 1991, and 2007) holds that based on Dual Coding Theory, a combination of imagery and verbal information enhance information processing. Purhosein Gilakjani (2012) states that multimedia include text, color, graphical images, animation, audio sound, and full motion video in single application that can be useful in improving students' understanding of language.

Haas (1989) has defined media by their technology, their symbol system, and their processing capabilities. Technology is the most obvious characteristic of a medium. Its function, its shape and other physical features are determined by mechanical and electronic aspects. Goodman (1976) refers to symbol systems as modes of appearance, or sets of elements such as words, picture components, and so on.

Kozma (1991) claims that the processing capabilities of a medium may facilitate performances the learner is able to operate, or perform those that the learner cannot. In other words, Salomon (1988) explains that the learner may incorporate such a process into his or her own repertoire of cognitive processes if they are explicit and fall within Vygotsky's (1978) "zone of proximal development". Nevertheless, Salomon and Clark (1977) point out that some of a medium's capabilities may not be used in a particular learning episode. However, Snow (1989) states that the use of a particular medium with certain capabilities will be useful for learners if the capabilities are used by instructional methods to supply certain representations or perform or model certain cognitive operations that are remarkable to the task and situation, and which the learners cannot perform or supply for themselves.

Among different kinds of media, one can refer to books, television, computer, and multimedia environments. A book is the most common medium that is used in school learning. A book can be characterized by the symbol system, and it can use text and pictures. According to Kintsch (1988), learning with text consists of the construction of two interconnected mental representations (i.e., a textbase and a situation model). Kintsch and van Dijk (1978) state that the textbase is a mental representation springing directly from the text, and it is a propositional representation of the meaning of the text. Anderson, Spiro, and Anderson (1978) call these structures schemata, but Minsky (1975) refers to them as frames, and Schank and Abelson (1977) label them as scripts.

On the other hand, orthographic symbols are not the only ones available in books. Many books, from primers to college books, consist of pictures and diagrams. Pressley (1977), and Levie and Lentz (1982) state that the use of pictures with text improves remembrance, especially for poor readers, if the pictures explain information central to the text, and when they evoke new content that is significant to the overall message. Rusted and Coltheart (1979) showed that texts with pictures of animals were useful for holding the information by both good and poor students.

Another medium is Television. A major role of video is its use of both auditory and visual symbol systems. Research has shown that visual attention is affected by several factors. While moment-to-moment visual attention of children may wander from the set, evidence suggests that they constantly monitor the presentation at a surface level, such that their visual attention is recaptured by certain audio cues (Anderson, Alwitt, Lorch, and Levin, 1979).

Many studies such as Meringoff (1982), Nugent (1982), Baggett and Ehrenfeucht (1982, 1983), Pezdek and Hartman (1983), Pezdek, Lehrer, and Simon (1984), and Pezdek and Stevens (1984) have been conducted to investigate a video program with its decomposed audio and visual presentations to compare the role of these two sources of information, individually and together. In most of these studies, the combination of visual and auditory symbol systems has led to more remembrance than visual-only and audio-only presentations.

So far, media have been characterized and differentiated from each other by their characteristic symbol systems, but some media such as computers are more differentiated by what they can do with information, that is, their capability to process symbols. For instance, computers can transform information in one symbol system to that in another one (Dickson, 1985). As a role of computer in L2 vocabulary learning, Pavicic (2008) refers to Computer Assisted Vocabulary Learning (CAVL) through which learners can successfully learn words by employing specialized programs available on CD-ROMs, the Internet, and popular computer games.

Learning may also happen through a combination of media or multimedia. According to Mayer (2005), multimedia is defined as presenting words as spoken and/or printed text, and pictures such as illustration, photos, animation, or video. Lexical material is presented in verbal form such as printed text or spoken text. Picture material is presented in pictorial form such as static graphics that involve illustrations, graphs, diagrams, maps, or using dynamic graphics which consist of animation or video. Kozma (1991) believes that multimedia can facilitate learning by putting together the different advantages of the individual media in a single instructional environment. Therefore, technology has combined several media devices and provided information which was previously achieved by several devices.

Multimedia learning is related to a number of theories. One such theory is Krashen's input hypothesis, based on which multimedia tools provide deeper comprehension through the combination of audio and visual input in the process of a second language learning (Wang, 2012). Another theory is the dual coding theory,

which has two assumptions: basic mental structures and processes. The structures are concerned with the auditory and visual representations. The processing assumptions govern the development and activation of the mental structures (Clark and Paivio, 1991). Baggett (1984) claims that referential connection can cause the representation of information in one channel to activate those in the other.

The third theory is the Cognitive Theory of Multimedia Learning. According to Mayer (2001), the cognitive theory is based on three theory-based assumptions which include the dual channel assumption, the limited capacity assumption, and the active processing assumption. Mayer (2002) asserts that the active learning processes happen when corresponding verbal and pictorial representations are in each channel of human cognitive system at one time.

Many researchers have investigated the effects of various multimedia environments on L2 vocabulary learning. Kellogg and Howe (1971) claimed that L2 vocabulary items are learned more effectively when they are along with images or actual objects related to their meaning. Terrell (1986) confirmed that a combination of unknown words with visual aids could facilitate vocabulary learning. Moreover, Underwood (1989) showed that images are remembered better than words, and words are remembered better if they are strongly associated with images. Similarly, Oxford and Crookall (1990) found that when pictures and texts are combined, the engagement of larger parts of the brain lead to greater depth of processing. And Baltova (1994) found that when L2 learners viewed the audio-visual material with subtitle, they learned more vocabulary.

Several researchers such as Ogasawara (1994) and Vanderplank (1993) believe that displaying captioned movies in language classrooms motivates L2 learners in second language learning because, by decreasing the affective filter effect during the learning process, it provides a relaxing and interesting environment for students.

On the other hand, some studies have shown that captioned movies are not useful for all L2 learners at any proficiency level. They can only be used for advanced and intermediate learners. If they are used for beginners, they should be matched to the learners' proficiency level (Baltova, 1999, Danan, 2004). Bird and Williams (2002), and Koolstra and Beentjes (1999) believe that captioned movies are helpful for learners with high reading ability, but students who have a poor reading skill cannot understand the film well. Other researchers such as Neuman and Koskinen (1992) also found that the modality of the captioned film may cause misunderstanding in the learning process. The fact that L2 learners try to watch the pictures of the movie and read the written texts on the screen simultaneously, causes difficulty in their comprehension.

Yeh and Wang (2003) tried to show the effect of multimedia vocabulary annotation and learning style on vocabulary learning. The results indicated that the most effective type of vocabulary annotation was text plus picture.

Jones (2004) investigated the effects of pictorial and written annotations on L2 vocabulary learning. The results showed that the written annotation and the pictorial and written annotation groups had significantly higher scores than the comparison group, but the difference between the pictorial annotation group and the comparison group was not meaningful.

Another study was done by Yoshii (2006) on the effects of various glossing on incidental vocabulary learning in a multimedia environment. Results showed that the textual-pictorial glosses group outperformed the textual glosses group on the definition-supply test.

In another experiment, Kim and Gilman (2008) tested the effects of multimedia components such as visual text, spoken text, and graphics on L2 vocabulary learning. The results showed that participants who received visual text and added graphics instruction or those who received visual text, added spoken text, and added graphic instruction outperformed the other groups.

Yanguas (2009) examined the effects of various multimedia glosses (textual, pictorial, and textual-pictorial on L2 vocabulary learning. The results showed that textual-pictorial gloss group performed better that all other groups.

Sydorenko (2010) studied the effect of input modality in three stimulus conditions (video, audio, and captions) on 1) the learning of written an aural word forms, 2) overall vocabulary gains, 3) attention to input, and 4) vocabulary learning strategies. He divided the learners into three groups. The first group was VAC that received video, audio, and captions, the second group that received video and audio was named VA, and the third group was VC that watched video and captions. The results showed that the scores of the VAC and VC groups on written recognition of words were higher than on aural recognition of the words, while the VA group scored higher on aural recognition of word forms than on written recognition. The findings also indicated that the VAC group learned more word meaning than the VA group.

In another study, Zarei and Rashvand (2011) considered the effect of multimedia on L2 vocabulary learning in different captioning conditions. They investigated the effect of verbatim and nonverbatim interligual and intralingual subtitles on L2 vocabulary comprehension and production. The results indicated that nonverbatim subtitles had positive effect on vocabulary comprehension irrespective of whether they

were interlingual or intralingual. The findings also showed that regardless of whether captions were verbatim or nonverbatim, intralingual subtitles affected vocabulary production positively.

A similar study was conducted with 120 first-year B.A students by Zarei and Sadeghi (2011), which examined the effectiveness of synchronous and asynchronous interlingual and intralingual captions on L2 learners' vocabulary comprehension and production. Participants were divided randomly into four groups. Each group saw the same film in different caption conditions. The results showed no significant differences among the four groups in L2 learners' vocabulary comprehension and production.

In another study, Tabatabaei and Shams (2011) investigated the effects of various multimedia glosses, namely text, picture, and text plus picture on online computerized L2 vocabulary learning of Iranian EFL learners. Based on the findings, they concluded that using multimedia gloss could have positive effect on online computerized L2 vocabulary learning.

Rezaee and Sharbaf Shoar (2011) investigated the impact of using multimedia, images and movies on learning vocabulary items included in a reading comprehension text. The findings indicated that students who were exposed to reading comprehension passages with movie clips outperformed the two other groups in learning and recalling of vocabulary.

Investigating the effect of mobile learning as a kind of multimedia environment on language learning, Khazaie and Ketabi (2011) concluded that L2 learners with high-visual and high-verbal abilities could learn more vocabulary when they were exposed to pictorial and written annotation. However, exposure to learning materials without annotations for L2 learners with low-visual and low-verbal abilities resulted in better vocabulary learning.

Another study by Zarei and Oshnouie Mahmoudzadeh (2013) investigated the effects of various multimedia glosses on L2 vocabulary learning and reading comprehension. They used 65 female students at a low-intermediate proficiency level. Randomly, one group was selected as the control group and the other three acted as the experimental groups. The three experimental groups received multimedia glosses in different conditions which involved a) textual glosses, b) pictorial glosses, and c) textual-pictorial glosses. Analyses revealed that differences between each of these three experimental groups (textual, pictorial, and textual-pictorial) and the control group in vocabulary production was statistically significant, but the differences among the three groups was not statistically meaningful.

In short, the various aspects of multimedia and L2 vocabulary learning have already been extensively studied in isolation. However, there seems to be a paucity of research on the effect of specific combinations of multimedia (different input modalities) on EFL vocabulary comprehension and production. The present study, therefore, aims to fill part of the existing gap in this area and shed light on some of the issues surrounding this little explored area.

## Method

### *Participants*

The participants of the study were 52 B.A level students of Imam Khomeini International University in Qazvin majoring in English translation and English teaching. The participants' age ranged from 18 to 30. The participants were divided into three treatment conditions. The first group was video, audio, and captions (VAC), which had 18 participants. The second, video and captions (VC), and the third, video and audio (VA) groups had 17 participants each.

### *Materials and Instruments*

The materials and data collection instruments utilized in this study included the following:

- In order to homogenize the participants in terms of their vocabulary knowledge, a standard language proficiency test was needed. Therefore, a multiple-choice vocabulary subtest of the Michigan Test was used for determining the students' proficiency level. It consisted of 40 vocabulary items.
- 'Challenges of Life' was the name of a documentary film which was played for students in three separate sessions. The total time duration of the film was 48 minutes. Therefore, the time duration of the film in each session was almost 16 minutes.
- A vocabulary pre-test consisting of 103 words was also administered. All the words that were included in the pre-test were selected from among the words that appeared in the content of the film. The participants were required to write the Persian equivalent of the highlighted words in each sentence. The purpose of this test was to remove from the post-tests those words that were familiar for the students.
- The post-test, which was given to the participants at the end of the experimental period in a separate session, consisted of 60 items in two parts. The first part included the 30 items of vocabulary comprehension in multiple-choice format. The second part involved 30 items of vocabulary production in 'fill-in-the-blanks' format.

*Procedures*

In order to achieve the purpose of the study, the following procedures were gone through.

First, 82 participants with the afore-mentioned characteristics were selected, of whom only 56 students attended all sessions. Second, the vocabulary subtest of the Michigan language proficiency test which consisted of 40 vocabulary items in multiple-choice format was administered to the students in the first session. As mentioned before, the purpose of this test was to homogenize the students. The time duration of this test was 30 minutes. After homogenization, 52 students who scored between one standard deviation above and below the mean were selected to constitute the participants of the study.

In the next step, a vocabulary pre-test was prepared. Attempt was made to choose difficult words that appeared in the content of the film. It included 103 English sentences with highlighted words requiring the students to write the meaning of the target words in Persian. The goal of this test was to remove the items that the students knew from the post-test. This test was given to the students in the second session, and it took 60 minutes.

Next, the participants were divided into three experimental groups. The first group was VAC, which received video with audio and captions. The participants watched the film during three sessions. In the first session, they watched 17:10 minutes of the film. In the second session, they watched another 16:20 minutes, and in the third session, the rest of the film was presented. After watching the film in each session, the students were required to write a summary of the film in 10 minutes. The purpose of this activity was to make sure that the students paid attention to the film. The second group was VA. This group watched the film without any captions. But all other procedures were the same as the first group. The third group was VC. This group went through the same procedure as the two previous groups, but they did not receive audio, they watched just video with captions.

At the end of the experimental period, the post-test, which consisted of 60 questions, was administered to the students in a separate session. This test consisted of two parts. Part A measured vocabulary comprehension; it had 30 questions in multiple-choice format in which each sentence had to be completed with one of the four alternatives. Part B included 30 vocabulary production items. This part was in 'fill-in-the-blanks' format in which sentences containing a blank were given to the students to fill in the blanks. To avoid the possibility of the participants filling the blanks with words other than the target words, the Persian equivalents of the target words were given in parentheses in front of each sentence. In addition, the first letters of the target words were presented in the blanks.

## Data Analysis

The gathered data were analyzed using two separate one-way ANOVA procedures. One ANOVA was used to investigate the effect of various multimedia on L2 vocabulary comprehension, and another ANOVA was used to see the effect of various multimedia on L2 vocabulary production.

## Results

*Investigation of the first research question*

The first research question aimed to investigate the effects of various multimedia modes on L2 vocabulary comprehension. To answer this question, a one-way ANOVA procedure was used. Descriptive statistics is presented in Table 1.

Table 1. Descriptive statistics for the one-way ANOVA on vocabulary comprehension

| Group | N | Mean | Std. Deviation | Std. Error | 95% Confidence Interval for Mean | |
|-------|---|------|----------------|------------|----------------------------------|---|
| | | | | | Lower Bound | Upper Bound |
| VAC | 18 | 10.5000 | 5.02055 | 1.18335 | 8.0033 | 12.9967 |
| VC | 17 | 14.1765 | 4.01926 | .97481 | 12.1100 | 16.2430 |
| VA | 17 | 13.8235 | 5.87617 | 1.42518 | 10.8023 | 16.8448 |
| Total | 52 | 12.7885 | 5.21061 | .72258 | 11.3378 | 14.2391 |

Based on the results of Table 1, it can be observed that the VC group has the highest mean ($\overline{X}$ = 14.17), followed closely by the VA group ($\overline{X}$ = 13.82). The VAC group has the lowest mean ($\overline{X}$ = 10.50). The implication is that the VC and VA conditions are more effective than VAC on learners' L2 vocabulary comprehension. To see whether or not the observed differences among the groups are statistically significant, the one-way ANOVA procedure was run. The result of the ANOVA procedure is presented in Table 2.

Table 2. ANOVA on vocabulary comprehension

|  | Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|
| Between Groups | 145.232 | 2 | 72.616 | 2.87 | .066 |
| Within Groups | 1239.441 | 49 | 25.295 |  |  |
| Total | 1384.673 | 51 |  |  |  |

Based on Table 2, the observed F value and the significance level ($F_{(2,49)}$ =2.87, P >.05) indicate that there are no statistically significant differences among the three groups.. Therefore, it can be concluded that although the VC and VA groups performed better than the VAC group on the vocabulary comprehension test, the differences among the three groups were not statistically meaningful.

### *Investigation of the second research question*

The aim of the second research question was to investigate whether or not there are any significant differences among the effects of various multimedia on L2 vocabulary production. To do so, another one-way ANOVA procedure was used, descriptive statistics, including the mean, standard deviation, etc. are shown in Table 3.

Table 3. Descriptive statistics for the one-way ANOVA on vocabulary production

|  | N | Mean | Std. Deviation | Std. Error | 95% Confidence Interval for Mean | |
|---|---|---|---|---|---|---|
|  |  |  |  |  | Lower Bound | Upper Bound |
| VAC | 18 | 4.8889 | 2.7629 | .6512 | 3.5149 | 6.2629 |
| VC | 17 | 7.1176 | 2.8036 | .6799 | 5.6762 | 8.5591 |
| VA | 17 | 5.5882 | 4.8997 | 1.1883 | 3.0690 | 8.1074 |
| Total | 52 | 5.8462 | 3.6642 | .5081 | 4.8260 | 6.8663 |

It can be seen from Table 3 that the group that received video and captions (VC) has the highest mean ($\overline{X}$ = 7.11), followed by the group that received video and audio (VA) ($\overline{X}$ = 5.58). The group receiving video, audio and captions has the lowest mean ($\overline{X}$ = 4.88). Another one-way ANOVA procedure was used to see whether or not the differences among the groups are statistically significant. Table 4 presents the results of the one-way ANOVA procedure.

Table 4. ANOVA on vocabulary production

|  | Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|
| Between Groups | 45.109 | 2 | 22.555 | 1.728 | .188 |
| Within Groups | 639.660 | 49 | 13.054 |  |  |
| Total | 684.769 | 51 |  |  |  |

Based on Table 4, the observed F value and the significance level ($F_{(2,49)}$= 1.72, P >.05) show that there are no statistically significant differences among the three groups. Therefore, it can be concluded that although the VC group performed better than the VAC and VA groups on vocabulary production test, the differences were not statistically significant.

### Discussion

The results of the present study indicated that there were no significant differences among the three experimental groups. This means that these three stimulus conditions had no differential effect on learners' L2 vocabulary comprehension and production. Although, there was a strong trend that students in the VC and VA groups got better vocabulary comprehension scores than those in the VAC group, and with respect to vocabulary production, the VC group was better than the two other groups, the differences were not statistically significant.

The findings of the present study are different from a number of studies some of which were reviewed in chapter two, and share certain aspects with others. For example, the result of this study contradicts Syndorenko's (2010) findings, which indicated that the VAC group learned more word meanings than the VA group. In his study, the performance of the groups that received captions (VAC and VC) were better than the VA group on the written recognition of word forms while the VA group was better than the two mentioned groups on the aural recognition of words.

In this study, the mean score of the VAC group was lower than that of the two other groups in vocabulary comprehension. This could be justified on grounds that information presented at the same time through both the visual-pictorial and auditory-verbal channels in the VAC condition might have led to cognitive overload. Such a finding is in line with the split-attention theory, based on which information which is presented to learners through multiple channels will force learners to divide their attention. As a consequence, only a part of the learners' attention will be paid to each modality, hence reducing the effectiveness of learning. This study also showed that the mean score of the VC group was higher than that of the two other groups in vocabulary production. This indicates that captions can be useful in improving vocabulary production. But as it was mentioned before, differences among the three experimental groups were not statistically meaningful. Although this result was to some extent unexpected and in contradiction with other studies conducted in this domain, several factors could have led to these results. One of the factors may have been that the students needed more time to accommodate themselves to their related stimulus condition of multimedia instruction. Another cause may have been related to having opportunity to take the movie backward and forward, which might have helped them to have better performance on the post-test.

The result of this study is in line with Zarei and Oshnouei Mahmoudzadeh (2013), who concluded that the differences between the effects of various multimedia glosses including textual, pictorial, and textual pictorial were not statistically meaningful, although there were significant differences between these three experimental groups and the control group. In line with the result of this study, Jones (2004) concluded that the effects of different input modalities including pictorial, written, and pictorial-written conditions were not statistically significant. Similarly, the results of this study corroborate those of Zarei and Sadeghi (2011), who examined the effects of various types of captions in the area of multimedia instruction on vocabulary learning but found no significant differences among them.

On the other hand, unlike the present study, the result of several studies such as Yeh and Wang (2003) and Kim and Gilman (2008) indicated that the text-picture annotation had more positive effect than other combinations of multimedia on vocabulary learning. In addition, Rezaee and Sharbaf Shoar (2011) showed the positive effect of the combination of text and movies compared with text and picture, and text only, on vocabulary learning. Some parts of Zarei and Hasani's (2011) findings who worked on the effects of different input modalities in the domain of multimedia learning on language learning were to some extent similar to those of this study, and other parts were different. They found no significant differences among the four types of intralingual glosses including interlinear, marginal, pre-text, and post-text on vocabulary recognition and recall. But, with regard to interlingual glosses, the pre-text and marginal groups outperformed the two other groups in vocabulary recognition. Moreover, Zarei and Rashvand (2011) could found significant differences among different aspects of multimedia environments in vocabulary learning.

The present study failed to find any meaningful differences among the effects of different multimedia combinations on L2 vocabulary comprehension and production. On the other hand, based on the findings of several previous studies, to some of which reference was made in the 'literature review' and 'discussion' sections, various multimedia learning conditions have been shown to be significantly more effective on language learning than the control condition. One obvious conclusion to be drawn from this is that learning through multimedia has got fairly obvious, and to some degree undeniable, superiority over traditional learning.

At the same time, given the findings of this study and the mixed findings of previous studies, it may be concluded that it does not matter what combination of multimedia is used. In other words, regardless of the various forms and combinations of forms, multimedia is an asset that can positively influence L2 vocabulary learning.

The findings of the present study may have implications for teachers, learners and materials developers. Language teachers can improve L2 learners' vocabulary comprehension and production by using instructional movies and by making second language classes more interesting for L2 learners. The observations in all treatment sessions showed that using multimedia was very pleasurable for students and changed the boring instructional environment to a relaxing and interesting one for second language learning. In addition, the findings of this study can be useful for material developers and instructional book designers. By including exciting instructional movies and incorporating visual images in the instructional books, they can create a more pleasurable instructional environment for second language learners.

Nonetheless, the controversies among the findings of the previous studies as well as those between the results of the present study and those of the previous studies warrant further research in this little explored area.

## References

Anderson DR, Alwitt LF, Lorch EP, Levin SR, 1979. Watching children watch television. In G. A. Hale and M. Lewis (eds), Attention and cognitive Development (pp.331-361). New York: Plenum.

Anderson RC, Spiro RJ, Anderson MC, 1978. Schemata as scaffolding for the representation of information in discourse. American Educational Research Journal, 15 (3), 433- 440.

Baggett P, 1984. The role of temporal overlap of visual and auditory material in forming dual media associations. Journal of Educational Psychology, 76 (3), 408- 417.

Baggett P, Ehrenfeucht A, 1982. Information in content equivalent movie and text stories. Discourse Processes, 5 (1), 73-99.

Baggett P, Ehrenfeucht A, 1983. Encoding and retaining information in the visuals and verbals of an educational movie. Educational Communication and Technology Journal, 31 (1), 23- 32.

Baltova I, 1994. The impact of video on the comprehension skills of core French students. Canadian Modern Language Review, 50 (3), 507- 531.

Baltova I, 1999. The effect of subtitled and staged video input on the learning and retention of content and vocabulary in a second language. Unpublished doctoral dissertation, University of Toronto. Canada.

Bird S, Williams J N, 2002. The effect of bimodal input on impact and explicit memory: An investigation into the benefits of within-language subtitling. Applied Psycholinguistics, 23 (4), 509- 533.

Clark JM, Paivio A, 1991. Dual coding theory and education. Educational Psychology Review, 3 (3), 149- 170.

Danan M, 2004. Captioning and subtitling: undervalued language learning strategies. Meta, 49(1), 66-77.

Dickson, W. P., 1985. Thought-provoking software: Juxtaposing symbol systems. Educational Researcher, 14(5), 30- 38

Goodman N, 1976. Language of art . Indianapolis, IN: Bobbs-Merrill.

Haas C, 1989. Seeing it on the screen isn't really seeing it: Computer writers' reading problems. In G. Hawisher and C. Selfe (eds), Critical perspectives on computers and composition instruction. New York: Teachers College Press.

Jones L, 2004. Testing L2 vocabulary recognition and recall using pictorial and written test items. Language Learning and Technology, 8 (3), 122- 143.

Kellogg GS, Howe MJA, 1971. Using words and pictures in foreign language learning. Alberta Journal of Educational Research, 17 (2), 89-94.

Khazaie S, Ketabi S, 2011. Contribution to vocabulary learning via mobiles. English Language Teaching, 4 (1), 174-184.

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

Kintsch W, 1988. The role of knowledge in discourse comprehension: A construction integration model. Psychological Review, 95 (2), 163-182.

Kintsch W, Van Dijk TA, 1978. Toward a model of text comprehension and production. Psychological Review, 85 (5), 363-394.

Koolstra CM, Beentjes JWJ, 1999. Children's vocabulary acquisition in a foreign language through watching subtitled television programs at home. Educational Technology Research and Development, 47 (1), 51- 60.

Kozma RB, 1991. Learning with media. Review of Educational Research, 61(2), 179-212.

Levie W, Lentz R, 1982. Effect of text illustrations: A review of research. Educational Communication and Technology Journal, 30 (4), 195- 232.

Mayer RE, 2001. Cognitive theory of multimedia learning. In R. E. Mayer (ed), The Cambridge handbook of multimedia learning ( pp. 31-49). New York: Cambridge University press.

Mayer RE, 2002. Cognitive theory and the design of multimedia instruction: An example of the two-way street between cognition and instruction. New Directions for Teaching and Learning, 89, 55- 71.

Mayer RE, 2005. The Cambridge handbook of multimedia learning. Cambridge: Cambridge University Press.

Meringoff L, 1982. What pictures can and can't do for children's story understanding. New York: American Educational Research.

Minsky MA, 1975. A framework for representing knowledge. In P. H. Winston (ed), The psychology of computer vision. (pp. 211-280). New York: McGraw-Hill.

Moreno R. Mayer RE, 2007. Interactive multimodal learning environments, Educational Psychology Review, 19 (3), 309- 326

Neuman SB, Koskinen P, 1992. Captioned television as comprehensible input: Effects of incidental word learning from context for language minority students. Reading Research Quarterly, 27 (1), 94- 106.

Nugent GC, 1982. Pictures, audio, and print: Symbolic representation and effect on learning. Educational Communication and Technology Journal, 30 (3), 163- 174

Ogasawara S, 1994. Effectiveness of using English captioned video on listening comprehension proficiency. Bulletin of Faculty of Liberal arts, 35(1), 103-114.

Oxford R, Crookall D, 1990. Vocabulary learning: A critical analysis of techniques. TESL Canada Journal/ Revue TESL DU Canada, 7(2), 9-30.

Paivio A, 1986. Mental representation: A dual coding approach. New York: Oxford University Press.

Paivio A, 1991. Dual coding theory: Retrospect and current status. Canadian Journal of Psychology, 45 (3), 255- 287.

Paivio A, 2007. Mind and its evolution: A dual coding theoretical approach. Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Pavicic V, 2008. Vocabulary learning and foreign language acquisition. Multilingual Matters Ltd: Crowell Press Ltd.

Pezdek K, Hartman EF, 1983. Children's television viewing: Attention and comprehension of auditory versus visual information. Child Development, 54, 1015- 1023.

Pezdek K, Lehrer A, Simon S, 1984. The relationship between reading and cognitive processing of television and radio. Child Development, 55, 2072- 2082.

Pezdek K, Stevens E, 1984. Children's memory for auditory and visual information on television. Development Psychology, 20 (2), 212-218.

Plass J, Jones L, 2005. Multimedia learning in second language acquisition. In R. Mayer (ed), The Cambridge handbook of multimedia learning (pp. 467- 488). New York: Cambridge University Press.

Pourhosein Gilakjani A, 2012. The significant role of multimedia in motivating EFL learners' interest in English language learning. I. J. Modern Education and Computer Science, 4 (1), 57- 66.

Pressley M, 1977. Imagery and children's learning: Putting the pictures in developmental perspective. Review of Educational Research, 47(4), 585- 622.

Rezaee AA, Sharbaf Shoar N, 2011. Investigating the effect of using multiple sensory modes of glossing vocabulary items in a reading text with multimedia annotations. English Language Teaching, 4 (2), 25- 34

Rusted R, Coltheart V, 1979. The effect of pictures on the retention of novel words and prose passages. Journal of Experimental Child Psychology, 28 (3), 516- 524.

Salomon G, 1988. AI in reverse: Computer tools that turn cognitive. Journal of Education Computing Research, 4 (2), 123-139.

Salomon G, Clark R, 1977. Reexamining the methodology of research on media and technology in education. Review of Educational Research, 47 (1), 99-120.

Schank RC, Abelson RP, 1977. Scripts, plans, goals, and understanding. Hillsdale, NJ: Erlbaum.

Snow RE, 1989. Aptitude-treatment interaction as a framework for research on individual differences in learning. In P. Ackerman, R. Sternberg, and R. Glaser (eds), Learning and individual differences. (pp. 13- 59). New York: Freeman.

Sydorenko T, 2010. Modality of input and vocabulary acquisition. Language Learning and Teaching, 14 (2), 50-73.

Tabatabaei O, Shams N, 2011. The effect of multimedia glosses on online computerized L2 text comprehension and vocabulary learning of Iranian EFL learners. Journal of Language Teaching and Research, 2 (3), 714- 725.

Terrell TD, 1986. Acquisition in the natural approach: The binding/Access framework. The Modern Language Journal, 70 (3), 213- 227.

Underwood J, 1989. HyperCard and interactive video. CALICO Journal. 6 (3), 7-20.

Vanderplank R, 1993. A very verbal medium: Language learning through closed captions. TESOL journal, 3 (1), 13-17.

Vygotsky LS, 1978. Mind in society. Cambridge, MA: Harvard University Press.

Wang YC, 2012. Learning L2 vocabulary with American TV drama from the learner's perspective. English Language Teaching, 5 (8), 217-225.

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, 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 and Technology, 10 (3), 58- 101.

Zarei AA, Hasani S, 2011. The effect of glossing conventions on L2 vocabulary recognition and production. The Journal of Teaching Language Skills (JTLS), 3 (2), 209-233.

Zarei AA, Oshnouie Mahmoudzadeh P, 2013. Multimedia glosses in L2 vocabulary learning and reading comprehension. Deutschland: LAP LAMBERT Academic Publishing.

Zarei AA, Rashvand Z, 2011. The effect on interlingual and intralingual verbatim and nonverbatim subtitles on L2 vocabulary comprehension and production. Journal of Language Teaching and Research, 2 (3), 618- 625.

Zarei AA, Sadeghi M, 2011. The contribution of synchronous and asynchronous interlingual and intralingual transcript presentation to vocabulary comprehension and production. Journal of Teaching English Language and Literature Society of Iran. 5 (1), 101-123.