

A review of multimedia glosses and their effects on L2 vocabulary acquisition in CALL literature

MOHAMMED ALI MOHSEN

*Najran University, College of Arts & Sciences, P.O. Box. 1988, Najran.
Saudi Arabia
(email: mamohsen@nu.edu.sa)*

M. BALAKUMAR

*Central Institute of Indian Languages, Manasagangothri, Mysore 570006 India
(email: balakumar@ciil.stpmv.soft.net)*

Abstract

This article reviews the literature of multimedia glosses in computer assisted language learning (CALL) and their effects on L2 vocabulary acquisition during the past seventeen years. Several studies have touched on this area to examine the potential of multimedia in a CALL environment in aiding L2 vocabulary acquisition. In this review, the researchers investigate the studies that deal with different modes of multimedia annotations in a CALL environment. This article reviews the empirical studies of multimedia glosses in reading as well as in listening comprehension activities and discusses the factors under which these studies have been conducted and their impact on L2 vocabulary acquisition. The study aims to understand how multimedia glosses have been used in the past to support second language vocabulary acquisition and also to explore any evidence regarding how multimedia glosses in a CALL environment can enhance the acquisition of L2 vocabulary. Only empirical studies (18 studies published in international refereed journals and conference) have been reviewed exclusively. The article discusses the findings of the reviewed studies and recommends future research.

Keywords: Multimedia glosses, multimedia annotations, hypermedia, vocabulary acquisition, vocabulary learning

1 Introduction

The advent of multimedia related computer technology has created several opportunities for language teachers and researchers for adopting it in the field of language learning and teaching. The accessibility, interactivity, the ability for combining different media, the efficiency and other potential benefits of multimedia have encouraged designers to manipulate texts, sounds, images, and video clips. Multimedia glosses, either in reading comprehension activities or in listening comprehension activities, and their impact on L2 vocabulary acquisition have been

addressed in CALL literature. It is observed that multimedia glosses have been the subject of many considerable studies in vocabulary acquisition and reading comprehension (Chun, 2006).

Glosses are traditionally defined as “a short definition or note in order to facilitate reading and comprehension processes for L2 learners” (Lomicka, 1998: 41). Now multimedia glosses can provide learners with different modalities (textual, visual, and auditory) and modes (video, picture, and text). Verbal annotations in the texts for reading are usually indicated by hyperlinks. By clicking the hyperlinked words, learners can access different forms of annotations either at the end of the text, in the margin, at the bottom of the screen, or in a pop-up window. Annotations in multimedia are facilitative for language learning (Nation, 2001). They help readers understand words more accurately by preventing misleading guessing, and also to avoid interruption in reading comprehension when readers try to check dictionaries to find out unknown words (Ko, 2005). Chun and Plass (1996: 183) emphasized that when “words or phrases are presented with different types of media, retention is easier”. They pointed out that “foreign words associated with actual objects or imagery technique are learned more easily than words without”. Jacobs, Dufon, and Fong (1994) gave four advantages for multimedia glosses: enhancing comprehension, increasing vocabulary learning, catering to students’ preferences, and providing greater use of authentic texts.

In addition to the multiplicity of presenting information through multimedia glosses, there are other advantages of multimedia glosses when they are compared to traditional glosses. Gettys, Imhof, and Kautz (2001: 91) pointed out that “online glosses can increase general comprehension, improve vocabulary retention, and save students’ time and effort in reading L2 texts”. Online glosses are not only beneficial for students but for teachers as well since they can “significantly increase comprehensible input” which is found to be important for second language acquisition (Krashen, 1989). Multimedia glosses enable readers to approach the text both globally and linearly while traditional glosses only enable readers to approach the text linearly (Martinez-Lage, 1997).

Researchers in the field have tried to empirically investigate the potential of multimedia annotations on L2 incidental vocabulary acquisition. They have extensively examined different variables with multimedia annotations that might have an effect on aiding L2 vocabulary during the past two decades. The purpose of the current review is to explore:

1. How have the multimedia glosses been used to enhance the acquisition of foreign/ second language vocabulary?
2. Do multimedia glosses addressed in the selected studies yield positive impacts on foreign/second language vocabulary acquisition?

2 Reviewing the related literature

The reviews of the studies in CALL have been of interest to many researchers. Liu *et al.* (2003), for example, reviewed empirical studies published in refereed journals between 1990 and 2000. They concluded that students showed positive attitudes to

the use of technology. Also, using technology lowers their anxiety and this would lead to an effective interaction in learning a foreign language. Zhao (2003) investigated nine studies carried out between the years 1997 and 2001. He summed up the findings of the existing literature of using technology in language education as follows: (a) there was a lack of more empirical and well designed studies on the use of technology in language learning; (b) participants in all the studies were adult learners; (c) the target language in his reviewed studies was English; and (d) the experiments investigated one or two language skills such as grammar or vocabulary. Felix (2005) investigated 52 studies which were conducted between the years 2001 and 2005 in terms of CALL effectiveness using different criteria, i.e., “number of participants, research design used, educational settings, language taught, subjects/skill taught, and variable under investigation” (Felix, 2005: 7). She concluded that some studies gave a poor description of the research design, some failed to investigate previous research and some chose poor variables for investigation. Another review by Hubbard (2005) examined 78 articles between the period 2000 and 2003. He investigated “eight hypotheses regarding trends in CALL literature which were related to research subjects: small number of subjects, overuse of questionnaires, limited time, lack of experience with the application, task type and CALL in general, and lack of training before and during the study” (*op.cit.*, 2005: 351). Hubbard concluded that researchers failed to gather data relevant to these eight characteristics and participants in most of the studies were untrained users. Stockwell (2007) tried to ascertain how technology had been used to achieve the objectives of language learning. The result of the studies examined showed that there was a small proportion of them where the reason for using technology was not evident. Some features of the technology did not appear to be used.

In a recent review, Abraham (2008) did a meta-analysis on eleven studies of computer-mediated glosses in second language reading comprehension and incidental vocabulary learning that were published up to September 2007. He addressed the effect of glosses in relation to three factors; “level of instruction, text type, and task of assessment” (Abraham, 2008: 199). His meta-analysis revealed that the intermediate learners who had access to computer glosses performed significantly better than learners who did not have computer glosses on incidental vocabulary learning over time. It showed that computer-mediated glosses had a large effect on narrative text but the expository text gained a medium effect. In addition, learners’ passive vocabulary was larger than productive vocabulary since they performed better in vocabulary recognition tests than in productive vocabulary tests. He recommended replicating the individual differences in a multimedia learning environment with authentic texts. To show a clear and comparative picture of CALL literature reviews, Table 1 explains the studies including the number of articles, the main findings, and their objectives.

Given that multimedia vocabulary annotations have been used extensively by researchers, it is important to examine various issues relating to multimedia vocabulary annotations and how they might be tackled. In other words, why did some studies yield different results from others? How can the findings of these studies be interpreted in the light of multimedia learning theories? How did studies use the vocabulary assessment methods and what were the possible results gained by them?

Table 1 *Reviews of studies in the CALL literature*

Reviewers	Objectives of reviews	Period	No. of articles	Main Findings
Liu <i>et al.</i> (2003)	potential of computer in enhancing the acquisition of language skills	1990–2000	246	students showed positive attitude to the use of technology
Zhao (2003)	investigating the effectiveness of technology in language education	1997–2001	9	lack of empirical studies, skills focused were grammar & vocabulary, and TL was English
Felix (2005)	CALL effectiveness restricted to some criteria	2001–2005	52	poor description of the research design and poor choice of variables
Hubbard (2005)	investigating eight hypotheses in CALL trends	2000–2003	78	studies failed to gather data relevant to the eight hypotheses
Stockwell (2007)	use of technology on language skills	2001–2005	206	some features of technology were not used
Abraham (2008)	computer-mediated glosses and their effects on L2 reading	Up to 2007	11	computer glosses had more effect on passive vocabulary than productive vocabulary

In examining the potential of multimedia glosses for L2 vocabulary acquisition, the researchers find it important to examine how multimedia annotations in a CALL environment have been used to support the acquisition of L2 vocabulary in the past. Therefore, the task here is to review the use of multimedia glosses in CALL literature on aiding second or foreign language vocabulary acquisition and to discuss the different findings gained by these studies from 1993 to 2009.

It should be noted that, for the purpose of this article, the terms second language/foreign language and glosses/annotations have been used interchangeably. Multimedia glosses have been examined in this review and have also been examined in a CALL environment including hypermedia, online multimedia, and hyperlinks that are connected to different media. The studies examined vocabulary acquisition through non-multimedia annotations such as the one by Duquette, Renić, and Laurier (1998) or the study on multimedia instruction by Kim and Gilman (2008). The studies which examined hypermedia annotations but exclusively utilized textual annotations, like that of AbuSeileek (2008) are not consulted. Our review excludes pure computer textual glosses such as sentence definition and word translation. Moreover, multimedia glosses used in a non-CALL environment such as marginal glosses or pictorial glosses are also left out.

3 Methodology

A series of three steps were followed to identify relevant studies. First, a key word search using 'multimedia glosses' or 'multimedia annotations', 'annotation', 'gloss', 'vocabulary acquisition' and 'hypermedia annotations' were performed in the Education Resources Information Center (ERIC), SCOPUS, and EBSCO, databases. For the key words 'multimedia glosses and vocabulary' five articles were found in ERIC, the same in SCOPUS, and one article in EBSCO. For the key words 'multimedia annotations and vocabulary' fifteen studies were found in ERIC, one in EBSCO, and six in SCOPUS. The results of these database searches were carefully checked. All articles that did not have the term 'second language', 'vocabulary', 'multimedia' or 'hypermedia annotations' were excluded. Eighteen studies were selected for review since they matched the key words of this article and covered multimedia annotations in a CALL environment. References gathered from recent reviews of computer assisted language learning, second language reading comprehension, and vocabulary learning were consulted (Chun, 2006; Plass & Jones, 2005). Bibliographies accumulated from articles, chapters, and unpublished dissertations and theses were also consulted.

For the purpose of the study, the data selected here are the empirical studies that examined multimedia annotations and L2 vocabulary acquisition in refereed journals and conferences between the period 1993 and 2009. Other sources such as doctoral dissertations, master's thesis/ reports, books, unpublished technical reports, non-refereed articles, and abstracts are not selected for this review. However, a thorough search was made for unpublished doctoral dissertations and master's theses that tackled the area of multimedia glosses and their impact on vocabulary acquisition. Eighteen articles which were published in nine journals and one conference were selected for this review (please see Appendix 1).

4 Classification of studies

The studies of multimedia glosses under reading comprehension activities can be classified under the following categories and are summarized in Appendix 2.

4.1 Multimedia glosses are more significant than printed glosses

The first studies of the early “nineties tried to compare the efficacy of multimedia glosses via computer with conventional printed textual glosses in non-multimedia CALL environments. Lyman-Hager *et al.* (1993) examined the impact of a multimedia software program, “Une Vie de Boy”, on vocabulary acquisition. The program provided a French story with 660 words, together with textual glosses, i.e., English and French definitions, verbal glosses, and pictorial glosses. Participants in the study were split into two groups. One group read the story via the computer program with access to multimedia annotations; the other read the story from a conventional printed text with the same glosses provided to the computer group. Several tracking means were used like time, number of keystrokes, and number of requests for the definition of words in the computer-based annotations. After reading the story, participants wrote a recall protocol and took a vocabulary test one week later for twenty annotated words. The results demonstrated that students who used the computer program to complete the reading task scored significantly better on the vocabulary recall test than students who used the printed text. Also, these students were able to retain vocabulary for some time after the treatment. However, there are some limitations to the findings of this study which threaten its validity: first, no pre-test was administered to ascertain the students’ prior knowledge, second, time was not tracked for the control group as it was for the computer group, and third, the test used was not valid since it measured the retention of words in the memory.

4.2 Effective mode

The studies of multimedia glosses investigated the effectiveness of different modes in aiding L2 vocabulary, i.e., images or dynamic video. These studies argued a need for the best effective mode for second language vocabulary acquisition. The notable study in this category is the one undertaken by Chun and Plass (1996) who conducted three studies that looked at the “effectiveness of annotations with different media types for vocabulary acquisition” (op. cit.: 185). One hundred and sixty English speaking students enrolled in a second year German course used CyberBuch, a program designed by the researchers, which contained textual annotations in German reading texts (in English), and visual annotations (video and pictures). The researchers used a within subject design in which all participants were exposed to three types of annotations: pictorial, video, and zero annotation. Participants were free to look up any of the annotation types available or not to use the gloss at all. They were told to read for comprehension and surprise recognition and productive post tests were given. In the production test, students were asked to write a protocol in their native language. The result from all these studies revealed that students scored 25% in the production test and 77% in the vocabulary recognition test. Moreover, picture annotated with text was better than video plus text annotations, and both were better than text annotations alone.

Al-Seghayer (2001) compared the effects of two multimedia CALL strategies (video mode or static pictures mode) in aiding vocabulary acquisition. Thirty ESL participants were asked to read an English story that contained different annotations (text, graphics, video, and sound) for target words. The glosses were in English (the target language). Al-Seghayer used a within subjects design in which all subjects received video annotations, picture plus text annotation, and textual annotations. Participants had to fill in a questionnaire and to conduct an interview after the treatment. Results showed significant differences for video clip annotations in the recognition and production of L2 vocabulary post-test. Al-Seghayer's findings fully contradicted those of Chun and Plass (1996) in relation to the best effective mode (video or static picture) in aiding L2 vocabulary acquisition. The contradictory findings of the above two studies could be attributed to many issues including the type of participants who took part in the two studies, and their familiarity with the visual aids (pictures and video clips) that had been used in the two studies. The difference in results could also be attributed to the type of tests that had been used in the studies particularly those of Al-Seghayer, which might assess memory retention.

Unlike the last studies, Akbulut (2007) observed no significant differences between the two visual groups (video and picture groups) who were under the treatment receiving an online text through hypermedia annotations. The third (control) group received the same text but with definitions of words without accessing picture or video clip annotations. Sixty-nine advanced Turkish freshman TEFL students enrolled on an English course were assigned to three groups: the textual glosses group provided with word definitions and grammatical functions; the pictorial group provided with pictorial glosses in addition to the textual glosses; and the video glosses group provided with dynamic video glosses accompanied with text. Visual treatment groups (pictorial and video groups) were free to choose any type of annotations. All the groups were requested to indicate the words they remembered from the text (recognition test) and they wrote the L2 equivalent of every target word (production test), or were asked to take the multiple choice test which mixed the previous two tests (meaning recognition test). Results of immediate and delayed post tests revealed that both visual annotations had positive effects on incidental vocabulary learning but there was no superiority of one type over the other. Again, these findings contradicted the findings of the previous two studies, i.e., Chun and Plass (1996) and Al-Seghayer (2001). This could be attributed to the types of participants in this study since they were advanced learners who therefore did not gain much benefit from one visual annotation over the other. Moreover, they had the same background culture (Turkish students). The contradiction could also be ascribed to the type of tests administered and to the type of picture and video assigned.

4.3 Multiple glosses are better than a single gloss

Several studies investigated the effects of image-based annotations combined with text, or text only. Yoshii and Flaitz (2002) examined the effects of different types of annotations (text alone, picture alone, and the combination of the two) in aiding L2 incidental vocabulary retention. One hundred and fifty-one ESL students at beginner and intermediate levels of language proficiency were asked to read an online story

with fourteen annotated words for comprehension purposes. Participants were randomly assigned to one of the three groups: the first group read a story with text annotation; the second group read the same story but with picture annotations; and the third group received the same treatment but with a combination of the two annotations, i.e., picture and text annotations. Post-test and delayed tests were administered to all the groups. Results revealed that the group that received picture plus text annotations performed better than the other two groups in immediate and delayed post-tests. The result of the immediate and delayed tests showed that students who were treated with two glosses performed significantly better than those who were treated with one gloss or no gloss. The study is a replication of Kost, Foss, and Lenzini (1999) who carried out a study in a non-multimedia environment where subjects were instructed to read a story with pictorial, textual, and both pictorial and textual glosses. These findings support Paivio's (1971–1986) dual coding theory which assumes that information is coded dually in the human mind either verbally (text and sounds) or non-verbally (picture, or objects). These two systems are interconnected to each other, namely, when words are represented by one system (verbal) and they can be activated by the other system (non-verbal) or vice-versa. The theory states that when information is presented dually through two systems, learning will be more effective than when information is presented singularly.

In a notable study, Yeh and Wang (2003) replicated the favor of multiple annotations. They investigated the effect of three types of vocabulary annotations on vocabulary learning among EFL college students. Eighty-two students were instructed to read a text in which they were exposed to two types of language glosses (L1 and L2 glosses) under three conditions: (1) only text annotation; (2) text and image annotations; and (3) text, image and audio annotations.

The researchers examined whether students' learning styles had any impact on the effectiveness of a particular annotation type. Students' learning styles were classified into auditory, visual-verbal (with text), visual-nonverbal (with pictures), and mixed preferences. A questionnaire on learning styles was administered before the reading task to explore students' learning styles. Pre-tests, immediate and delayed tests were administered. Results indicated that students who had access to text plus picture annotation significantly outperformed those who had access to text plus picture plus audio annotation. The researchers concluded that text plus picture annotation was the most effective for vocabulary learning among the participants in the study. The researchers found that students would prefer visual annotations to auditory annotations. The findings did not show any clear influence of the learners' perceptual preference on the effectiveness of vocabulary annotation types in vocabulary learning.

4.4 L1 versus L2 glosses

Yoshii (2006) examined the effectiveness of adding pictorial cues in L1 and L2 glosses on L2 incidental vocabulary and whether L1 or L2 glosses had more effect on English vocabulary acquisition for EFL Japanese students. One hundred and fifty-five Japanese participants were split into four groups and assigned the following four tasks: (1) L1 text only; (2) L2 text only; (3) L1 text plus picture; and (4) L2 text plus picture. They were instructed to read an online story that included

fourteen annotated words. As learners clicked on the target word, a definition or a picture in the L1 or L2 was provided. After the treatment, an immediate test was administered. Two weeks later a delayed post-test was administered to check word retention. Results showed no significant differences in terms of gloss types (L1 or L2 glosses). They revealed that both L1 and L2 glosses were useful for incidental vocabulary learning but picture plus text seemed to be more effective than text only in retrieval of words over time.

Jacobs *et al.* (1994) investigated the effects of multimedia glosses on vocabulary learning as well as learners' preferences for types of glossing (L1 or L2). Eighty-five students of Spanish as a second language were asked to read a Spanish text with 32 annotated words under three conditions: (1) L1 (English) gloss; (2) L2 (Spanish) gloss; and (3) No gloss. Students were asked to write a recall of a passage, and were also asked to translate the glossed words. Unexpected immediate vocabulary post-tests were administered after the treatment, and the same test was administered four weeks later. In addition, participants were asked to fill in a questionnaire designed to elucidate their preferences on glosses. The results of the post-tests showed that students who had more glosses outperformed the others in translation. However, no significant differences were shown on the delayed post-test in relation to long-term retention for both types of glosses. Jacobs *et al.* (1994: 26) reported that "those learners with glosses outperformed their peers who did not have glosses on a vocabulary instrument administered shortly after reading the passage that difference disappeared when the vocabulary instrument was re-administered four weeks later." Participants showed their preferences in favor of L2 glosses (Spanish).

5 Learning differences and learning styles

5.1 Learners' preferences

The factor of learning differences in multimedia glosses has been extensively investigated in a study by Plass *et al.* (1998). It examined the effects of multimedia glosses on individual learning differences as well as learning styles (visualizers and verbalizers). One hundred and three fluent English speaking university students learning German were given a pre-view video that summarized the key events in the story. They were instructed to read an authentic German story delivered on the computer, which had previously been annotated by the program. Twenty-four annotated words had both verbal (verbal text accompanied with pronunciation) and visual annotations (a picture of the word or a short video). Participants were asked to take a post-test which required them to produce the L1 translation of each target word and mark whether the word reminded them of hearing it or seeing a picture or a video. The results of the study showed that participants acquired vocabulary better when they used both visual and verbal annotations. Moreover, the participants recalled the translation of German words better if they were given their preferred mode of annotations. They concluded that students should be given both visual and verbal modes of input in a multimedia setting so that they could choose the mode that best suited them. These findings demonstrated Mayer's (1997–2001) generative theory of multimedia learning which assumed that the learner had to select verbal and pictorial

annotations. As a result of selection, the human mind had to create verbal and visual representations of the information processed, then organize these representations into coherent mental representations, and also build connections between the two types of system and integrate this information with prior knowledge in working memory.

5.2 Cognitive load

In a subsequent study, Plass *et al.* (2003) examined the effect of multimedia annotations on acquisition of German vocabulary. The study focuses on how cognitive load affects the way learners with different cognitive abilities (high and low spatial and verbal abilities) process verbal and spatial information. The cognitive load can be operationalized as the load on working memory when learners are exposed to learning inputs. One hundred and fifty-two English-fluent students enrolled in a second-year German course read a 762-word German story presented by a multimedia computer program. Students either did not receive any annotations, or received verbal annotations, visual annotations, or both, for 35 key words in the story. The researchers used different cognitive tests to measure verbal and spatial abilities (whether high or low). Results showed that spatial learners were helped by graphic information, but learners with low spatial abilities were not aided by visual annotations of unknown words. They discovered that learners with different cognitive abilities were aided by their cognitive abilities. Learners with high spatial ability benefitted from visual annotations whereas learners with low spatial abilities were not helped by visual annotations. They concluded that when learners with high verbal ability processed both visual and verbal annotations, it would result in a high cognitive load and this would negatively affect their learning. A criticism of this study is that learners took the post-test on the day following the treatments. The interval between the treatments and the vocabulary post-test would threaten the validity of the test, as students would have had a chance to discuss their work with others after the treatment.

Another study is that of Acha (2009) which replicated the variable of the cognitive load in multimedia glosses and their effects on acquisition of vocabulary for children in third and fourth grades in a primary school. Acha argues that the presentation of picture and text for children may lead to the same result obtained by Plass *et al.* (2003) for low ability learners, who performed worse when picture and text were presented simultaneously rather than separately. One hundred and thirty-five participants were given a short story to read in a computer program under three conditions: (1) verbal annotations (written translation); (2) visual annotations (picture associated with words); and (3) both. The results of the post-tests and delayed tests revealed that the “word group” who had verbal annotations performed better than the other two groups in recalling words. The researchers concluded that the groups who were exposed to visual annotations performed worse because pictures contained “a higher cognitive load than the word and led to less effective learning” (*op. cit.*, 2009: 28).

The findings of Acha (2009) are consistent with those of Plass *et al.* (2003), that adults with low cognitive ability performed worse when pictures were added to multimedia programs. This study demonstrates Chandler and Sweller’s (1991) cognitive load theory that suggests that the working memory capacity is very limited.

Hence, presenting too many elements to be processed in visual and verbal working memory can lead to cognitive overload. Therefore, materials and instructions should minimize the chances of overloading.

5.3 Working memory capacity

In the context of individual differences, Chun and Payne (2004) examined the effect of working memory capacity on the acquisition of L2 vocabulary. They studied the effects of four variables that might have an effect on working memory capacity; i.e., look-up behavior, vocabulary test scores, comprehension test scores, and a phonological recall protocol test. Chun and Payne tried to examine how learners of different working memory capacities (WMC) look up words and whether there was any relationship between the multimedia annotations and students with low working memory capacity. Thirteen English native speaker students were instructed to read a multimedia-based German story on CD-ROM. The story contained 102 annotated words for which participants could access their L1 translation, i.e., English and their L2 (German) synonyms, or they could view images and video clips of those annotated words. After the treatment, a vocabulary recognition test was administered. In addition, students were asked to write a summary in English of everything they could remember about the story. In order to measure WMC, Chun and Payne used “non word repetition and reading span tests” (*op. cit.*, 2004: 489). They found that there was a relationship between look up behavior and the phonological working memory. They suggested that learners with a lower WMC look up more words in multimedia glosses to compensate for their memory limitations while reading.

6 Students' participation in authoring multimedia

Nikolova (2002) attempted to answer the question of whether students' participation in authoring multimedia materials could improve their vocabulary when the time factor was considered. Sixty-five second semester students of first year French as a second language were requested to read a French text, which was downloaded from the internet. Students were assigned to two groups: one group studied the text in a multimedia environment where the annotated target words were already developed by the teacher; the other group studied the text in the same multimedia environment but targeted words were not highlighted and students had to link words with their own annotations (text, sound, and picture) by using the authoring mode of the software. Pre-tests and post-tests were administered. One month later, the delayed test was administered to check for word retention. Results revealed that if time was not considered, students given the authoring multimedia treatment learned vocabulary better than the annotation group. However, when time was considered, no significant differences were found in favor of the creation multimedia group.

7 Noticing Hypothesis

Yanguas (2009) replicated Bowles's (2004) findings using Schmidt's (1990) Noticing Hypothesis. Bowles's (2004) interpretation of this hypothesis is that learners have to notice the form of the input and show awareness before it can be processed.

To measure the construct of “noticing” he used an online protocol to decide which type of glosses induced learners to notice more words. Yanguas’s (2009) replication of Bowles (2004) added one independent variable, which was multimedia glosses (not studied by Bowles, 2004). Yanguas studied the effects of different multimedia glosses on promoting noticing, and whether noticing could lead to better comprehension or vocabulary learning. Ninety-four Spanish students read an online passage under one of the following conditions: textual glosses, pictorial glosses, both textual and pictorial glosses, or no glosses (control group); students were also asked to think aloud as they read, and were given instructions and training prior to the treatment on how to think aloud while performing the tasks. Recognition and production tests (pre-test, immediate test, and delayed test) were administered. Results showed that target words were noticed and recognized more significantly by the multimedia glosses groups than the control group. In addition, no significant differences were found in the production of the target words among groups but the combination group showed higher performance in comprehension of the passage.

8 Pedagogical learning training

O’Bryan (2008) investigated how the pedagogical learning training discussed by Hubbard (2004) in a glossed online reading text would enhance the use of the glosses and improve the learning and retention of L2 vocabulary. The five principles set out by Hubbard (2004) were: (i) to encourage teachers to experience CALL from the learner’s perspectives; (ii) to provide learners with teacher training; (iii) to use a cyclic approach to training by reminding students of points that they might have forgotten over time; (iv) to encourage “collaborative debriefings” to create a balance between the task objective and the language learning objective; and (v) to help learners generalize strategies learned to other CALL activities. Twenty-two Asian ESL students were asked to read an online text written by the researchers, which contained 48 abstract words provided with text plus image annotations. Participants were divided into two groups: an experimental group received a pre-unit training session on Hubbard’s (2004) CALL principles for learner training (except the fifth principle) before the treatment; the control group did not receive any such training. Results of the test showed that the experimental group who received pre-training became aware of language learning in the annotated online text and retained the vocabulary even three weeks after the treatment.

9 Studies in listening comprehension activities

Researchers of listening comprehension tried to explore the effects of multimedia annotations for listening comprehension activities in aiding second language vocabulary acquisition. The objectives of listening tasks were restricted to comprehension rather than vocabulary acquisition even in the late ‘eighties and early ‘nineties, since the traditional audiotape presented audio without text. Some researchers used conventional media like video images, video subtitles, and audio accompanied with images but restricted their studies to examining audio comprehension (Mueller 1980; Baltova, 1999; Chung, 1994). Even some researchers such as Brett (1997) investigated

the effects of multimedia on comprehension of L2 listening in a CALL environment but did not check the effectiveness of choosing pictorial and textual annotations in a CALL environment in aiding L2 vocabulary acquisition. Several studies tried to address the effectiveness of presenting information aurally in multimedia CALL annotations and whether they could help L2 learners comprehend the aural text and acquire unknown words. They also tried to investigate certain variables that might have an effect on L2 vocabulary acquisition in listening comprehension activities with multimedia annotations. Studies are categorized as follows and are summarized in Appendix 3.

9.1 Multiple glosses, single gloss, and zero gloss

Jones and Plass (2002) investigated the effect of listening comprehension activities in a multimedia CALL program (choice of written and pictorial annotations) on the acquisition of French vocabulary. The study hypothesized that students who completed listening comprehension activities that contained pictorial and written annotations would acquire more vocabulary and remember more ideas from the text than those who received only one type of annotation (pictorial or written) or no annotations. One hundred and seventy-one fluent English students enrolled in a French course listened to a historical account in French presented by a computer program. The participants were randomly assigned to one of four listening treatments: the listening text (a) with no annotations available; (b) with only written annotations available; (c) with only pictorial annotations available; and (d) with both written and pictorial annotations available. Immediate and delayed recognition vocabulary tests were administered. Results showed that students remembered word translations better when they had multiple annotations than when they had one type of annotation. In addition, students who had pictorial annotations outperformed those who had written annotation in the delayed test. The study supported Mayer's (1997–2001) generative theory of multimedia learning.

9.2 Students' choices

In her subsequent study, Jones (2003) examined whether students' choices of information in multimedia annotations could enhance their ability in listening comprehension as well as vocabulary acquisition. She built her study on Jones Vogely's (1998) findings, which stated that "students struggled with listening comprehension because of the lack of choice of information" (Jones, 2003: 60). One hundred and seventy-one English students studying French as a second language listened to an aural text in different treatments: the aural text (a) with no annotations; (b) with only verbal annotations; (c) with only visual annotations; and (d) with both visual and verbal annotations. She administered recognition vocabulary tests to measure students' vocabulary gains. After the treatment, some students were asked to attend an interview. Results demonstrated that students performed best with the combination of visual and verbal annotations, moderately when single annotations were present, and poorest when no annotations were there. Also, choosing from different annotations helped them comprehend the aural text and acquire vocabulary

well. The students' views reflected the first component of Mayer's (1997, 2001) generative theory of multimedia learning, that students had to select from the input either verbal or visual annotations, then to organize the information and to integrate new information with their prior knowledge.

9.3 Consistency of tests with the type of annotation

In a subsequent study, Jones (2004) examined the effects of vocabulary tests (either recognition or production tests) that would consistently match the type of annotation accessed with gains in L2 vocabulary. She conducted two consecutive studies to investigate the effects of multimedia annotations (verbal and pictorial) on acquisition of French keywords. Participants were assigned to one of four aural multimedia groups: three groups received written, pictorial, or both annotations while listening and a control group received no annotations. In her first study, eighty-two English students were asked to listen to the French aural passage in which students could click on the annotated word which directly led him/her to written annotations, pictorial annotations or both written and pictorial annotations. Immediate and delayed recognition post-tests were administered to them. In the second study, seventy-seven English students received the same treatment as had been done in the first study but they were tested to recall vocabulary from their memory in the form of a production test. Results of the first study revealed that students with multiple annotations and single annotation performed equally well on the vocabulary recognition tests. The second study's findings showed that students in a written annotation mode performed better on recall vocabulary because the testing mode (production test) matched the treatment mode (written annotations).

9.4 Collaborative work

In a notable study, Jones (2006) investigated whether students who worked in groups could better recall L2 vocabulary and comprehend a passage when they listened to an aural French passage with multimedia annotations as compared to those who worked alone. Sixty-eight English-speaking college students enrolled in a French course listened to aural texts. They were randomly assigned to one of the four groups: the first group worked alone and received no annotations; the second one worked in pairs and also received no annotations; the third group worked alone but received pictorial and written annotations; and the fourth group received the same annotations but worked in pairs. The researcher investigated whether the collaborative work of students in multimedia annotations would affect L2 incidental vocabulary learning and their recall of the aural passage. Her results showed positive impacts on both annotation types, either alone or in pairs for recalling and identifying words, whereas collaborative work did not indicate a positive effect in aiding L2 vocabulary. However, students who worked collaboratively performed significantly better in terms of aural comprehension. She concluded that accessing written and pictorial annotations would help students better recall and identify vocabulary and would also improve their comprehension of an L2 aural passage. This study revealed a positive effect of annotation types on vocabulary acquisition regardless of the variable of collaborative work that indicated no effect.

9.5 Differences in students' ability

In a recent study, Jones (2009) investigated how different multimedia annotations would affect vocabulary acquisition while students of different spatial and verbal abilities listened to a French passage. She investigated whether the interaction of learners' cognitive abilities (spatial and verbal abilities) with multimedia annotations would improve L2 vocabulary acquisition in an aural multimedia environment. One hundred and seventy-one English-speaking students of French were randomly assigned to one of the four groups while listening to a passage in French: a control group received no annotations and three treatment groups received written annotations, pictorial annotations, or both written and pictorial annotations. A recognition vocabulary test of 25 words was administered first to rule out prior knowledge, and then learners completed a cognitive ability test to be grouped according to their performance on the test, which matched their degree of cognitive abilities. Groups were categorized as: (a) high spatial ability (HSA); (b) low spatial ability (LSA); (c) high verbal ability (HVA); and (d) low verbal ability (LVA). After the treatment, students were invited to complete a recognition vocabulary test and vocabulary recall test. Three weeks later, the same recognition and recall vocabulary tests were administered. Results of the tests revealed that when both types of annotation were present, there were few significant differences between HSA and LSA groups. The HVA learners performed significantly better than the LVA learners in recall and vocabulary tests when the pictorial annotations were presented alone. The zero effect of the combination of both annotations for HSA learners could be justified because when both annotations were processed, learners experienced extra cognitive load, which negatively affected their performance on the tests.

10 Discussion and suggestions for future research

The first question of this study addressed the ways in which multimedia annotations had been investigated in relation to second language vocabulary acquisition. Based on a review of the selected studies, multimedia glosses had used different variables to enhance second language vocabulary acquisition in listening and reading comprehension activities. The selected studies had used multiple (verbal, textual, and visual) annotations, different modalities (auditory, pictorial), and different modes (video, picture, and text). However, there are considerable issues related to the selected studies which are discussed below in the light of the research questions.

10.1 Participants and target languages

By examining the reviewed studies, the researchers found that almost all reviewed studies were carried out with college level participants (Acha, 2009 is an exception). This is in line with the findings of Zhao's (2003) review of CALL studies conducted between 1997 and 2001, that participants in his selected studies were adult learners. Thus, the findings of multimedia vocabulary annotation studies cannot be generalizable to children or school students. Studies are needed to examine the effects of multimedia vocabulary annotations on other subjects, such as school students, and to explore the modes which best suit their cognitive ability. Languages used for the

reviewed studies were limited to those with the same orthography (Latin form) as the participants' native languages were French, German, English, and Spanish (please see Appendices 2 and 3). Future studies need to explore the effects of multimedia annotations on learners whose native language has different orthography from their target language, for example, Arab or Japanese EFL learners.

10.2 Assessment tools

The main assessment tools used for the selected studies were vocabulary tests as dependent variables, although some studies had intensified tests with questionnaires and interviews to check learners' impressions and attitudes towards their preferred mode and glosses. Studies used two types of test, recognition test and production test, to assess the acquisition of L2 vocabulary. Vocabulary recognition tests took several forms such as word identification, image identification, multiple choice form, and ticking L1 equivalents. The vocabulary production tests were almost always limited to word translation into L1. Only Al-Seghayer (2001) used word definitions in L2 as a production test to assess the acquisition of English words. Some studies used tests of the synonyms of words in L1 (e.g., Yoshii & Flaitz, 2002; Yeh & Wang, 2003; Yoshii, 2006). The variation in the tests could be attributed to the type of language annotation in the courseware material or to the language annotation of the online multimedia texts that were used in studies. For example, the studies by Chun and Plass (1996), Plass *et al.* (1998, 2003) and Chun and Payne (2004) used the English annotations (L1) in their courseware, *CyberBuch*. Thus, they tested learners in word translation into L1 to match the language annotation type available in the program. It is observed that the production tests through word translation into L1 were not valid since they tested the receptive vocabulary knowledge of the learners.¹ In addition, word translation makes it difficult for learners to find out the exact term in L1 to correspond with the concept in L2. Word translation will help learners identify the word or recognize it in context but will not help them to use the word productively. Researchers need to administer productive tests for L2 learners and to focus on word definition in L2, word completion, and word synonyms in order to examine L2 vocabulary acquisition. In addition, a limitation of Plass *et al.* (2003) was the absence of a delayed post-test during the treatment to assess word retention after the treatment.

10.3 L1 versus L2 annotations

One considerable issue for multimedia annotation is the type of language annotations. As mentioned earlier, some researchers into L2 learning used L1 annotations in designing their courseware or material (Chun & Plass, 1996; Plass *et al.*, 1998, 2003; Chun & Payne, 2004; Nikolova, 2002; Yanguas, 2009), and some designed their programs in L2 (Al-Seghayer, 2001; Yoshii & Flaitz, 2002; Yeh & Wang, 2003; Yoshii, 2006). The selection of the text annotation language relied heavily on the homogeneity and heterogeneity of subjects participating in researchers' studies. While studies such

¹ Mondria and Wiersma (2004) argue that learning vocabulary through translation is considered as receptive vocabulary knowledge.

as those of Chun and Plass (1996), Plass *et al.* (1998, 2003) and Chun and Payne (2004) used learners who were native speakers of English and learned German as a second language, Al-Saghyer's participants were multi-cultural learners who spoke different languages and studied English as a second language. Therefore the choice of text annotation language depended on the language in which they could communicate with each other. However, in L2 listening activities only one type of language annotation, L1, was used (Jones & Plass, 2002; Jones, 2003, 2004, 2006, 2009). A few researchers, such as Jacobs *et al.* (1994) and Yoshii (2006) used both L1 and L2 to explore the type of annotation which best facilitated L2 vocabulary learning, but none of them gained significant differences in favor of one text type over the other.

Along with the issue of comparing the types of language annotation, one should consider the vocabulary assessment tools used for the different studies. While Jacobs *et al.* used a word translation test to assess the gain in Spanish vocabulary acquisition, Yoshii used different vocabulary assessment tests: definition test (either in L1 or in L2) and word recognition test. The assessment tool used by Jacobs *et al.* (1994) was limited to word translation into L1; this cannot be relied upon since it did not measure acquisition but rather retention, because it relied heavily on short-term memory.

One of the shortcomings of Yoshii's (2006) study is the limited number of annotated words used to teach subjects in L2 multimedia annotation treatments. He used only fourteen words in his treatment to measure the acquisition of L2 words. Another shortcoming of the same study is the short interval of two to three weeks between the immediate post-test and the delayed vocabulary test, which threatens the test's validity in assessing vocabulary retention. It is advisable to prolong the intervals between the post-test and the delayed post-test (Xu, 2010) and also to avoid tests that examine receptive vocabulary knowledge and focus on productive tests that assess word use in context. In an L2 listening context, the comparison between L1 and L2 annotation types has not yet been explored. All the studies reviewed used native language annotations (i.e., French) and this issue should be considered in future research studies.

10.4 Effective annotation mode

Studies in L2 reading activities are inconclusive about the most effective mode of annotation in aiding L2 vocabulary learning. Chun and Plass (1996) demonstrated a significant difference with the use of static pictures, Al-Seghayer, (2001) showed the superiority of dynamic video, and Akbulut (2007) demonstrated a zero effect for both modalities. The contradictory results could be attributed to the cultural background of participants, target language, level of word difficulty, and the type of assessment used (tests). However, all studies proved the positive impact of image-based annotations. This issue has not yet been investigated in relation to listening comprehension activities.

10.5 Cognitive ability

Another issue that has been extensively examined in both L2 reading and listening activities is the cognitive ability of learners in an L2 multimedia environment. Plass *et al.* (2003) found that "verbal ability had a positive effect on vocabulary learning but spatial ability did not" (*op. cit.*: 236). In other words, when low ability learners

processed both visual and verbal annotations, their performance in a vocabulary test was worse. However, no difference was found between low and high ability learners when they accessed only verbal annotations. This effect was due to the cognitive load imposed by visual annotations.

Jones's (2009) findings were consistent with Plass *et al.* (2003) in this respect. She found that HVA students outperformed LVA ones in a delayed recall protocol test when pictorial annotations, alone or combined with written annotations, were present. The two studies confirm the superiority of multiple annotations over a single annotation, which in turn substantially outperformed no annotations. Plass *et al.* (2003) concluded that learners who received both verbal annotations and visual annotations performed significantly better than those who received one type of annotation or received no annotations. However, studies are inconclusive in terms of the differences between the high and low spatial ability learners. Plass *et al.* found that HSA learners outperformed LSA learners when visual annotations were present but LSA learners scored significantly better than HSA learners when both annotations were processed and there was no significant difference between them when both types of learner accessed verbal annotations only. On the other hand, Jones did not find any significant difference between HSA and LSA learners who "performed similarly on immediate and delayed vocabulary recognition post-tests" (Jones, 2009: 284).

Nevertheless, whilst Plass *et al.* (2003) utilized a dynamic video and a static picture to indicate visual annotations, it was not clear which type of visual annotation most imposed the cognitive load that affected the performance of low ability learners, because both types were processed differently. Unlike Plass *et al.* (2003), Jones (2009) used only one type of visual annotation, i.e., pictorial annotations. Further study needs to investigate which mode of visual annotation (picture or video) has a stronger effect on working memory capacity that might hinder the learning of L2 vocabulary. Also, a limitation of Plass *et al.* (2003) was the absence of delayed post test during the treatment to assess word retention after the treatment.

10.6 Multiple glosses and zero gloss

All the studies of L2 reading proved the positive impact of dual glosses over a single gloss or no gloss. On the other hand, studies of L2 listening activities found contradictory results. While Jones and Plass (2002) and Jones (2003) proved that students learned L2 vocabulary best when pictorial and verbal annotations were processed, moderately when they processed one type of annotation, and worse when they processed no annotations, Jones (2004) concluded that significant differences were no longer found when annotations were absent while students were involved in different treatments with annotations or no annotations. Jones (2004) found that students performed equally well when annotations were present or absent. The contradictory results could be ascribed to the fact that learners were distracted while processing information relating to the multimedia annotations while they were listening.

10.7 The use of word concreteness

With one exception (O'Bryan, 2008), none of the relevant studies annotated abstract words in an L2 multimedia environment. Concrete words are easily imaged by

different visual modes as they are tangible to the senses. However, abstract words express ideas or feelings, so they cannot be perceived easily. Therefore, it is difficult to visualize them because they cannot fully correspond to the exact meaning of the terms. If a word indicates a high concreteness, the visualization will be more significant. Xu (2010) argues that words of high concreteness are more comprehensible than words of low concreteness. Furthermore, multimedia glosses did not investigate words in context and all the selected studies were restricted to dictionary words. Further studies should examine abstract words and carefully select images, which fully depict the meaning of the target words, and much care should be given to contextual words and idioms to be visualized in L2 multimedia annotations.

10.8 Ignorance of the use of animation

Another important issue that has been neglected in the use of L2 multimedia modes is the use of an animation mode. Multimedia modes in the above studies used video clips, static pictures and text to visualize the meaning of the annotated words but the efficacy of animation was not fully investigated. A criticism of the visual annotations is that they were used only as fragments to indicate the annotations. In the L2 listening context student exposure was restricted to aural multimedia passages and there was no investigation of audiovisual material. Therefore, the pictorial annotations in the selected studies represented different contexts. Future study should incorporate audiovisual input and the annotation should indicate the video, animation, or images that learners would process during involvement to audiovisual material.

11 Conclusion

Multimedia annotation studies have contributed to the body of literature with insight into facilitating L2 vocabulary acquisition. They have shown a high significant difference for multimedia annotations compared to traditional glosses in aiding L2 vocabulary acquisition. Multiple annotations were found to be more effective than a single annotation or no annotations in an L2 reading context. However, there is a need for a meta-analysis study to statistically address the positive effects of multimedia vocabulary annotations, since the researchers have not studied the mean differences between the control groups (no glosses) and the experimental groups (gloss access) in the relevant studies. The above-mentioned issues should be considered in order to make it more feasible to generalize from the findings. Vocabulary assessment methods should focus on the use of productive vocabulary in the target language. Consequently, it is preferable for the language annotation types to be in L2 to match the language testing. Further studies are needed to resolve the discrepancy that has been found in the issue of the superiority of dual glosses over a single gloss in an L2 listening context. Multimedia glosses need to use abstract words, contextual words, and idioms and to be carefully visualized. Furthermore, researchers into the use of multimedia glosses need to consider animation modes to aid L2 vocabulary acquisition rather than using the visual mode as fragments outside the context of the material that is being taught.

References

- Abraham, L. (2008) Computer-mediated glosses in second language reading comprehension and vocabulary learning: A meta-analysis. *Computer Assisted Language Learning*, **12**(3): 199–226.
- AbuSeileek, A. (2008) Hypermedia annotation presentation: learners' preferences and effect on EFL reading comprehension and vocabulary acquisition. *CALICO Journal*, **25**(2): 1–16.
- Acha, J. (2009) The effectiveness of multimedia programmes in children's vocabulary learning. *British Journal of Educational Technology*, **40**(1): 23–31.
- 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: An International Journal of the Learning Sciences*, **35**(6): 499–517.
- Al-Seghayer, K. (2001) The effect of multimedia annotations modes on L2 vocabulary acquisition: a comparative study. *Language Learning & Technology*, **5**(1): 202–232. <http://llt.msu.edu/vol5num1/alseghayer/default.html>
- 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.
- Bowles, M. A. (2004) L2 glossing: To CALL or not to CALL. *Hispania*, **87**(3): 541–552.
- Brett, P. (1997) A comparative study of the effects of the use of multimedia on listening comprehension. *System*, **25**: 39–53.
- Chandler, P. and Sweller, J. (1991) Cognitive load theory and the format of instruction. *Cognition and Instruction*, **8**: 293–332.
- Chun, D. M. and Plass, J. L. (1996) Effects of multimedia annotations on vocabulary acquisition. *The Modern Language Journal*, **80**: 183–198.
- Chun, D. and Payne, J. S. (2004) What makes learners click: Working memory and look-up behavior. *System*, **32**: 481–503.
- Chun, D. (2006) CALL technologies for L2 reading. In: Ducate, L. and Arnold, N. (eds.), *Calling on CALL: From theory and research to new directions in foreign language teaching*. San Marcos, TX: CALICO, 69–98.
- Chung, U. K. (1994) *The effect of audio, a single picture, multiple pictures, or video on second-language listening comprehension*. Unpublished doctoral dissertation, University of Illinois at Urbana–Champaign.
- Duquette, L., Renié, D. and Laurier, M. (1998) The evaluation of vocabulary acquisition when learning French as a second language in a multimedia environment. *Computer Assisted Language Learning*, **11**(1): 3–34.
- Felix, U. (2005) Analyzing recent CALL effectiveness research: Towards a common agenda. *Computer Assisted Language Learning*, **18**: 1–32.
- Gettys, S., Imhof, L. A. and Kautz, J. O. (2001) Computer-assisted reading: The effect of glossing format on comprehension and vocabulary retention. *Foreign Language Annals*, **34**: 91–106.
- Hubbard, P. (2004) Learner training for effective use of CALL. In: Fotos, S. and Browne, C. (eds.), *New perspectives in CALL for second language classrooms*. Mahwah, NJ: Lawrence Erlbaum, 45–67.
- Hubbard, P. (2005) A review of subject characteristics in CALL research. *Computer Assisted Language Learning*, **18**(5): 351–368.
- Jacobs, G. M., Dufon, P. and Fong, C. H. (1994) L1 and L2 vocabulary glosses in L2 reading passages: Their effectiveness for increasing comprehension and vocabulary knowledge. *Journal of Research in Reading*, **17**(1): 19–28.
- Jones, L. C. (2003) Supporting listening comprehension and vocabulary acquisition with multimedia annotations: The students' voice. *CALICO Journal*, **21**(1): 41–65.

- Jones, L. (2004) Testing L2 vocabulary recognition and recall using pictorial and written test items. *Language Learning and Technology*, **8**(3): 122–143. <http://llt.msu.edu/vol8num3/jones/default.html>
- Jones, L. (2006) Effects of collaboration and multimedia annotations on vocabulary learning and listening comprehension. *CALICO Journal*, **24**(1): 33–58.
- Jones, L. (2009) Supporting student differences in listening comprehension and vocabulary learning with multimedia annotations. *CALICO Journal*, **26**(2): 267–289.
- Jones, L. and Plass, J. (2002) Supporting listening comprehension and vocabulary acquisition in French with multimedia annotations. *The Modern Language Journal*, **86**(4): 546–561.
- Jones Vogely, A. (1998) Listening comprehension anxiety: Students' reported sources and solutions. *Foreign Language Annals*, **31**(1): 67–80.
- Kim, D. and Gilman, D. A. (2008) Effects of Text, Audio, and Graphic Aids in Multimedia Instruction for Vocabulary Learning. *Educational Technology & Society*, **11**(3): 114–126.
- Ko, M. (2005) Glosses, comprehension, and strategy use. *Reading in a Foreign Language*, **17**(2): 125–143.
- Kost, C., Foss, P. and Lenzi, J. (1999) Textual and pictorial gloss: Effectiveness on incidental vocabulary growth when reading in a foreign language. *Foreign Language Annals*, **32**: 89–113.
- Krashen, S. (1989) *Language acquisition and language education*. New York: Prentice Hall.
- Liu, M., Moore, Z., Graham, L. and Lee, S. (2003) A look at the research on computer-based technology use in second language learning: A review of the literature from 1990–2000. *Journal of Research on Technology in Education*, **34**(3): 250–273.
- Lomicka, L. (1998) To gloss or not to gloss: An investigation of reading comprehension online. *Language Learning & Technology*, **1**(2): 41–50. <http://llt.msu.edu/vol1num2/article2/default.html>
- Lyman-Hager, M. A., Davis, J. N., Burnett, J. and Chennault, R. (1993) Une vie de boy: Interactive reading in French. In: Borchardt, F. L. and Johnson, E. M. T. (eds.), *Proceedings of the CALICO 1993 annual symposium on "assessment"*. Durham, NC: Duke University, 93–97.
- Martinez-Lage, A. (1997) Hypermedia technology for teaching reading In: Bush, M. and Terry, T. (eds.), *Technology enhanced language learning*. Lincolnwood, IL: National Textbook Company, 121–163.
- Mayer, R. (1997) Multimedia learning: Are we asking the right questions? *Educational Psychologist*, **32**(1): 1–19.
- Mayer, R. E. (2001) *Multimedia learning*. New York: Cambridge University Press.
- Mondria, J. and Wiersma, B. (2004) Receptive, productive, and receptive+ productive L2 vocabulary learning: What difference does it make?. In: Laufer, B., and Paul, B., (eds.), *Vocabulary in a second language. Selection, acquisition, and testing*. Amsterdam/Philadelphia: John Benjamins Publishing Company, 79–99.
- Mueller, G. (1980) Visual contextual cues and listening comprehension: An experiment. *Modern Language Journal*, **64**: 335–340.
- Nation, I. S. P. (2001) *Learning vocabulary in another language*. New York: Cambridge University Press.
- Nikolova, O. (2002) Effects of learners' participation in authoring of multimedia materials on student acquisition of vocabulary. *Language Learning & Technology*, **6**(1): 100–122. <http://llt.msu.edu/vol6num1/nikolova/default.html>
- O'Bryan, A. (2008) Providing Pedagogical Learner Training in CALL: Impact on Student Use of Language-Learning Strategies and Glosses. *CALICO Journal*, **26**(1): 142–159.
- Paivio, A. (1971) *Imagery and verbal processes*. New York: Holt, Rinehart, & Winston. (Reprinted 1979, Hillsdale, NJ: Lawrence Erlbaum Associate).
- Paivio, A. (1986) *Mental representations: A dual coding approach*. Oxford, UK: Oxford University Press.

- Plass, J. L. and Jones, L. C. (2005) Multimedia learning in second language acquisition. In: Mayer, R. E. (ed.), *The Cambridge handbook of multimedia learning*. New York: Cambridge University Press, 467–488.
- Plass, J. L., Chun, D. M., Mayer, R. E. and Leutner, D. (1998) Supporting visual and verbal learning preferences in a second-language multimedia learning environment. *Journal of Educational Psychology*, **90**: 25–36.
- Plass, J. L., Chun, D. M., Mayer, R. E. and Leutner, D. (2003) Cognitive load in reading a foreign language text with multimedia aids and the influence of verbal and spatial abilities. *Computers in Human Behavior*, **19**: 221–243.
- Schmidt, R. (1990) The role of consciousness in second language learning. *Applied Linguistics*, **11**: 129–158.
- Stockwell, G. (2007) A review of technology choice for teaching language skills and areas in the CALL literature. *ReCALL*, **19**(2): 105–120.
- Xu, J. (2010) Using multimedia vocabulary annotations in L2 reading and listening activities. *CALICO Journal*, **27**(2): 311–327.
- Yanguas, I. (2009) Multimedia glosses and their effect on L2 text comprehension and vocabulary learning. *Language Learning & Technology*, **13**(2): 48–76. <http://llt.msu.edu/vol13num2/yanguas.pdf>
- Yeh, Y. and 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 & Technology*, **10**(3): 85–101. <http://llt.msu.edu/vol10num3/yoshii/default.html>
- Yoshii, M. and Flaitz, J. (2002) Second language incidental vocabulary retention: The effect of picture and annotation types. *CALICO Journal*, **20**(1): 33–58.110.
- Zhao, Y. (2003) Recent developments in technology and language learning: A literature review and meta-analysis. *CALICO Journal*, **21**: 7–27.

Appendix 1*Names of the journals and numbers of the relevant articles included*

Title of journal	Number of articles
Language Learning & Technology	5
CALICO Journal	6
Computers in Human Behavior	1
Journal of Educational Psychology	1
Journal of Research in Reading	1
System	1
The Modern Language Journal	1
An International Journal of the Learning Sciences	1
British Journal of Educational Technology	1

Name of conference and number of relevant articles included

Title of conference	Number of articles
Proceedings of the CALICO 1993 annual symposium on "Assessment"	1

Appendix 2: *Studies of reading comprehension activities and vocabulary acquisition*

Variables	Multimedia Glosses & traditional glosses	Effective Mode	Multiple glosses	L1& L2 Glosses	Learner preferences	Cognitive mode	Working memory capacity	Authoring multimedia	Noticing Hypothesis
Title of study	Lyman-Hager <i>et al.</i> (1993)	1-(Chun & Plass, 1996) 2- (Al-Seghayer, 2001) 3-Akbulut, 2007)	1-(Yoshii & Flaitz, 2002) 2- (Yeh & Wang, 2003)	a-(Jacobs <i>et al.</i> , 1994). b-Yoshii, (2006)	(Plass <i>et al.</i> , 1998)	1-Plass <i>et al.</i> , 1998. 2-(Acha, 2009)	(Chun & Payne, 2004)	(Nikolova, 2002)	(Yanguas, 2009)
Material	<i>Une Vie de Boy</i> software	1- CyberBuch 2- software designed by the author 3-online text	1-online story 2-reading text	1-reading text 2-multimedia reading program	CyberBuch	1-CyberBuch 2-short story multimedia learning	CyberBuch	Online multimedia text	Online multimedia text
Target Language	French	1-Spanish 2-English 3-English	1-English 2-English	1-Spanish 2- English	German	1-German 2-English	German	French	Spanish

Appendix 3: *Studies of listening comprehension activities and vocabulary acquisition*

Factors	Multiple glosses	Student Voice	Type of test	Students' collaborative work	Students' differences
Title of study	(Jones & Plass, 2002)	(Jones, 2003)	(Jones, 2004)	(Jones, 2006)	(Jones, 2009)
Material	French aural reading program	French aural reading program	French aural reading program	Aural multimedia program	Aural multimedia program
Target Language	French	French	French	French	French