Effects of Multimedia Annotations on Thai EFL Readers' Words and Text Recall

Jenjit Gasigijtamrong[1]

[1] Department of Teaching English to Speaker of Other Languages, Payap University, Chiang Mai, Thailand

Correspondence: Jenjit Gasigijtamrong, Institute of Southeast Asian Studies, Payap University, Chiang Mai, 50000, Thailand. E-mail: jenjit.iseas@gmail.com

Abstract

This study aimed to investigate the effects of using multimedia annotations on EFL readers' word recall and text recall and to explore which type of multimedia annotations – L1 meaning, L2 meaning, sound, and image – would have a better effect on their recall of new words and text comprehension. The participants were 78 students who enrolled in an Intermediate EFL course at a university in the north of Thailand. Five instruments – a questionnaire, a vocabulary pre-test, a computerized log-file, and two posttests: a vocabulary test and a text recall test – were used to collect data. First, the participants responded to the questionnaire and the pre-test in the classroom. One week later, they read a hypertext equipped with multimedia annotations in the computer laboratory while the log files recorded their selections of words and annotations. Immediately after that, they answered the two post-tests. Three weeks later, a vocabulary test was given to the participants again without informing in advanced. The results revealed that (1) accessing multimedia annotations led to significantly greater vocabulary recall (42%) among Thai EFL learners; (2) their retention lasted at least three weeks after being exposed to the multimedia-annotated text; (3) they recalled words with images significantly better than those with no images; (4) they recalled about 31% of the text with multimedia annotations; and (5) no significant difference was found regarding the effects of using different types of multimedia annotations on learners' word recall and text recall when compared between learners.

Keywords: word recall, text recall, multimedia annotations, Thai EFL readers

1. Introduction

In such a country as Thailand, where English is merely a foreign language, vocabulary is always claimed to be the most difficult language aspect to learn. With the availability of computer technology, the possibility for enhancing students' vocabulary knowledge is broadened. The use of such multimedia technology as audio, image, video, and animation in the design of annotations expands the provision of vocabulary meanings that match individual learning preferences.

Vocabulary annotations have been used to assist learners in reading second/foreign language (L2) texts. Unlike dictionaries, annotations provide accurate targeted meanings of unknown words that may be difficult or impossible to guess from the context, so authentic texts can be used. Annotations allow L2 readers to follow the text with very little interruption (Nation, 2001); they increase L2 learners' independence from relying on teachers for unknown words (Stewart & Cross, 1991).

An investigation on the effects of using annotations on learning new words in L2 reveals that L2 learners learn better when they had access to vocabulary annotations. However, most available empirical studies were conducted in the West, in which learners' first and second languages are similar, for example English and French, English and Spanish, or English and German. The current study, however, investigated learners whose first language is Thai and their second language is English. Thai does not use Roman alphabets, and it is a tonal language. Having a tonal language as a mother tongue may lead the learners to utilize the sound annotation more frequently than those who do not.

1.1 Review of Literature

The current study adopted the Cognitive Theory of Multimedia Learning (Mayer, 2009) as its framework. This theory comprises three assumptions: dual channels, limited capacity, and active processing. The dual-channel assumption states that human process visually and auditorily represented materials in two separate information-processing channels: the auditory/verbal channel and the visual/pictorial channel. The second assumption states that the amount of information that can be processed in each channel at one time is limited. Readers can hold only a few images or a few sounds in their working memory at any one time. The last assumption asserts that humans actively engage in cognitive processing to construct a coherent mental representation. These active processes include paying attention, organizing incoming information, and integrating new knowledge with prior knowledge and/or with new knowledge from the other channel.

Studies that compare the effects of reading L2 texts with and without accessing annotations lead to the following conclusions. With respect to L2 vocabulary learning, annotations led to better performances (Hulstijn, Hollander, & Greidanus, 1996; Jacobs et al., 1994; Rashkovsky, 1999). Moreover, L2 readers' word retention lasted at least one to two weeks (Ko, 1995; Kost et al., 1999; Watanabe, 1997). Regarding L2 text comprehension, research suggested that annotations had a more positive effect on text recall (Davis, 1989; Jacobs, 1994; Leffa, 1992; Rashkovsky, 1999).

In addition, empirical studies had further investigated which types of annotations or combinations was the most effective in facilitating L2 learners' incidental vocabulary learning and text recall (see Mohsen & Balakumar, 2011; Xu 2011). The first aspect investigates which language (L1 or L2) of definition annotations better helps L2 readers learn new words and comprehend the text. Jacobs, Dufon, and Hong (1994) conducted a study with American students learning Spanish as an L2; while Ko (1995) conducted similar study with Korean students learning English. Jacobs and his colleagues did not find any differences in the vocabulary recall between the two different language groups, but Ko found that the L1 group performed better in both the immediate and the one-week delayed vocabulary test. It is noticeable that another variable, the test types, comes into play here. Jacobs and colleagues used a production test whereas Ko used a recognition test. Hence, the different outcome may stem from test difficulty. A decade later, Yoshii (2006) conducted a similar study with Japanese students learning English using both recognition and production vocabulary tests. Yoshii found no significant difference in the test scores between the two groups in both types of vocabulary tests. Therefore, it is still not conclusive if the language of the definition annotations has any significant effect on incidental learning of L2 vocabulary although it is highly likely that there may not be any difference effects. In terms of text comprehension, both Jacobs et al. (1994) and Ko (1995) did not find any significant difference between the two languages. Their finding corresponds to what was found in a study by Bell and LeBlanc (2000) and another study by Ko (2005) with another group of Korean students learning English. It, therefore, can be concluded that the language of annotations did not affect L2 readers' text comprehension.

Another aspect of studies regarding multimedia annotations is the comparison between the effects of textual and pictorial annotations on incidental vocabulary learning. These studies divided their participants into three groups reading L2 texts in three different types of annotations: textual, pictorial, and both textual and pictorial. Yoshii and Flaitz (2002) found that the combination group outperformed the other groups in both an immediate test and a two-week delayed test. In 2006, Yoshii conducted another study, which revealed that the combination performed significantly better than the textual only group in the production test. These results are in contrast to the result of a study by Yanguas (2009), who found no significant differences among the three groups' vocabulary test scores.

Regarding the effect of textual and pictorial annotations on L2 text comprehension, only Yanguas (2009) investigated this issue and found that the combination group performed significantly better than the other groups in the text comprehension test. In conclusion, the combination of textual and pictorial annotations tend to have better effects on incidental L2 vocabulary learning than either textual or pictorial only annotations (Yoshii, 2006; Yoshii & Flaitz, 2002). This may also be true for L2 text comprehension (Yanguas, 2009). Xu (2011) explained that it may be because most of the participants in these studies were intermediate or advanced L2 readers, therefore, they relied little on vocabulary annotations for help.

Another dimension of an investigation of the effects of visual annotations on L2 vocabulary learning and text comprehension is a comparison of still pictures and video clips. In these studies, words are annotated with three conditions: textual definitions; textual definitions and still pictures; and textual definitions and video clips. Chun and Plass (1996) conducted a study with American students learning German. They found that learners received significantly higher scores for words with textual and picture annotations than for words with other types of

annotations. This is contrast to the result from Al-Seghayer's study (2001), which revealed that ESL learners remember words with textual definitions and video clips the most. Another study by Akbulut (2007) investigated the same types of annotations but with a different design. While the first two studies employed a within-subject design, Akbulut used a between-subject design. The participants in his study were randomly assigned to three types of annotations: definitions of words, definitions and associated pictures, and definitions and associated short videos. The results revealed that the groups that had access to definitions with both types of visuals had significantly higher vocabulary scores on both immediate and delayed post-tests than the definition only group. But no significant difference was found in the text comprehension test. It is, therefore, still inconclusive regarding the effects of still pictures and video clips on L2 learner's incidental word learning and reading comprehension.

Up to now, studies investigating the effects of multimedia annotations on L2 learners' vocabulary learning and text comprehension have focused on the comparison of textual and pictorial annotations; the other aspect of multimedia annotations – the sound or the pronunciation of the word – has been neglected. Hegelheimer (1998) compared the effects of textual and sentence-level audio annotations on L2 vocabulary learning and text comprehension. He discovered that the subjects who used annotations on two or more occasions performed significantly better on the vocabulary test, but no significant difference was found in the text comprehension. Another study that includes sound annotation in the investigation is a study by Yeh and Wang (2003). EFL students in Taiwan were divided into three groups with three types of annotations: textual definition; textual definitions and pictures; and textual definitions, pictures, and sounds. The results reveled that the textual definitions and pictures performed significantly better than other groups.

The last element of studies about multimedia annotations is an attempt to find out for whom annotated texts work the best. Plass, Chun, Mayer, and Leutner (1998) found that visualizers were more effective in using visual cues for remembering words, whereas verbalizers were more effective in using verbal cues. Furthermore, visualizers performed better on propositions with both verbal and visual annotations than on verbal annotations alone, but verbalizers performed well on both types. Moreover, visualizers recalled texts better when both visual and verbal annotations were available than when only verbal annotations were available, whereas verbalizers performed well on a text recall task in both conditions. The findings support the first principle of cognitive theory of multimedia learning (Mayer, 2001), which states that people learn better when information is processed in both verbal and visual channels.

In sum, empirical studies that examined effects of annotations on incidental L2 vocabulary learning through reading activities indicated that L2 learners learn vocabulary better when they have access to annotations. There is also a tendency that theyL2 learners would learn vocabulary better when they have access to both textual and picture annotations; however, the differences between the effects of still picture annotations and video clip annotations is still inconclusive. Furthermore, no significant difference has been found between L1 and L2 annotations. Next, studies that investigated the effects of sound annotations on incidental vocabulary learning and text comprehension are rare. Finally, the effect of multimedia annotations on L2 learners' text comprehension is also inconclusive.

The purposes of this study were twofold. First, it investigated how Thai EFL readers' utilization of multimedia annotations while performing a reading task affects their incidental learning of vocabulary and their abilities to comprehend the text read. Second, it examined which type of multimedia annotations - L1 meaning, L2 meaning, pronunciation, and pictures - is the most effective for readers' incidental vocabulary learning and text comprehension.

1.2 Research Questions

1) What is the effect of multimedia annotations on Thai EFL learners' vocabulary recall?

2) Which type of multimedia annotations best affect Thai EFL learners' vocabulary recall?

3) What is the effect of multimedia annotations on Thai EFL learners' text recall?

2. Method

2.1 Participants

The population of the current study was Thai students who enrolled in an English II course, which was one of the four required English language courses at a university in the north of Thailand. The students' English language ability is at intermediate-low level. The participants were selected by using a cluster sampling method. Since the students enrolled in this course were grouped according to their major fields of study, the researcher randomly selected one class from the following fields: pharmacy, law, business administration, hotel and tourism

management, and psychology. Initially, the total number of participants was 84. However, due to the participants' absences during the data collection procedures, the final number of the participants was 78. Twenty-two were male, and 56 were female. Their average age was 19.

2.2 Material

Three instructors who had been teaching English II for at least 10 years unanimously selected the same text from five proposed texts of the same reading difficulty level. The length of the selected text was 411 words with the Flech Kincade Readability level[1] at 9.8.

The text was then distributed to 30 students who did not participate in the later study. They were told to read the text and to underline the words that they did not know the meanings. Three other English instructors were also asked to read and to select the words that they believed their students did not know the meanings. Words that were selected 75% of the times or more by both the students and the teachers were used as target words in this study. The total number of target words was 42. All words were annotated with three types of annotations: meaning in Thai (L1), meaning in English (L2), and the sound of a native speaker pronouncing the word. In addition, 24 words were also annotated with pictures.

The hypertext consisted of five pages, with the text appeared in the left column (see Figure 1). The annotated words were highlighted in blue color. When a word was selected, a box would appear in the right column with choices of annotations. Additionally, the sound of a native English speaker reading the text on each page was also available as an optional choice for learners. The participants could select to go backward or forward to any page at anytime. The finished hypertext was submitted to a technological expert and a language expert for reviewing. Suggestions by both experts were corrected and changed accordingly. Then, the hypertext was field tested with 30 students who did not participate in the later study.

Figure 1. Screenshot of the text when a meaning in Thai was selected

2.3 Instruments

Five instruments were used to collect data: a questionnaire, a pre-test, a computerized log file[2], a vocabulary recall test, and a reading recall test. The questionnaire was used to collect demographic data. The pre-test was used to check whether the participants had any previous knowledge about the target words, and the computerized log file was used to track the participants' selection of multimedia annotations. Two production post-tests were also used to collect data. The vocabulary recall test asked the participants to write the meanings of the target words, and the reading recall asked the participants to retell what they had just read in their own words. In both tests, they could write their answers in either Thai or English.

2.4 Procedure

The data collection procedure was as follows.

Week 1: The questionnaire and the pretest were distributed to the participants in class. They were allowed as much time as they needed to finish both forms.

Week 2: The participants read the text in the computer lab during which the computerized log file recorded their selections of multimedia annotations. They were informed that the purpose of the activity was to read for comprehension, and they were allowed to work as long as they wanted. After they exited the program, a

vocabulary test and a reading test were distributed.

Week 5: The participants were given a vocabulary test again without being informed.

3. Results

3.1 Multimedia Annotations and Word Recall

To examine the effects of utilizing vocabulary annotations on the participations' ability to recall words, means and standard deviations of their three vocabulary test scores – pretest (PT), immediate post-test (IT), and delayed post-test (DT) -- were calculated. The mean of the participants' pretest score was 1.86 (S.D. = 2.362). They recalled about 42% of the words in the immediate post-test (Mean = 17.66, S.D. = 9.973); and they recalled about 30% of the words in the delayed post-test (Mean = 12.97, S.D. = 9.899). Next, a repeated measured ANOVA (see Table 1) was then used to compare these mean scores to find out if the participants could recall the words significantly better after being exposed to multimedia annotations and if there was any retention after three weeks.

Table 1 shows that the means of the participants' immediate post-test scores was significantly higher than that of their pretest scores and that of their 3-week delayed test scores ($p > .05$). Furthermore, the mean of their 3-week delayed test scores was significantly higher than that of their pretest scores ($p > .05$).

Table 1. Repeated measure ANOVA of the participants' test scores

	Source	Tests	SS	df	Mean Square	F	Sig.*
Tests of Within-subject Contrasts	Tests	Linear	4383.951	1	4383.951	109.577	.000
		Quadratic	4969.542	1	4969.542	143.315	.000
	Error (Tests)	Linear	2800.549	77	40.008		
		Quadratic	2427.291	77	34.676		
Test of Between-Subjects Effects	Intercept		24987.085	1	24987.085	194.677	.000
	Error		8984.582	77	128.351		

*$p > .05$

3.2 Types of Multimedia Annotations and Word Recall

To investigate which type of multimedia annotations helped Thai EFL learners to recall words the most, two analyses, a within-subject analysis and a between-subject analysis, were performed.

3.2.1 Within-Subject Analysis

A within-subject analysis was performed to investigate the participants' ability to recall words with and without picture annotations. The 42 annotated words were divided into two groups, those with pictured annotations (P) and those without pictured annotations (NP). That is, 18 annotated words (NP) have meaning in Thai, meaning in English, and sound annotations; while 24 words (P) have all of the aforementioned annotations and picture annotations.
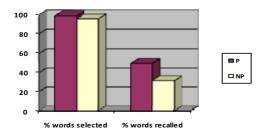

Figure 2. Percentage of selected and recalled words

A comparison of the number of words that the participants selected and the number they recalled is shown in Figure 2. The percentage of selections of P words was 98.88% (Mean = 23.73, SD. = 0.715), whereas the percentage of selections of NP words was 95.80% (Mean = 17.24, SD. = 1.219), which were only 3% less than the selections of P words. Figure 2 also showed that the participants recalled 49.84% of the P words (Mean = 11.96, SD. = 5.92), but they recalled only 31.70% of the NP words (Mean = 5.71, SD. = 4.664), or about 18% less than their recall of P words.

Next, a paired *t*-test was used to compare the participants' ability to recall words with and without picture annotations.

Table 2. Learners' ability to recall words with and without picture annotations

	N	Mean	SD.	Std. Error Mean	*t*	df	Sig. (2-tailed)
P recalled	78	11.96	5.916	.670			
NP recalled	78	5.71	4.646	.526			
P recalled – NP recalled	78	6.256	3.819	.432	14.469	77	.000*

*p < .01

Table 2 shows a significant difference between the numbers of P and NP words recalled ($t = 14.469$, $p < .01$). In other words, the participants recalled words with pictured annotations significantly more than words with no picture annotations although they looked up both types of words at a similar frequency.

3.2.2 Between-Subject Analysis

A between-subject analysis was used to further investigate the effects of different types of multimedia annotations on the participants' vocabulary recall and their text recall. The frequencies of their responses from the log files were counted; then means, standard deviations, and percentage were calculated to identify the participants' preference types of multimedia annotations.

The participants were grouped by the percentage of their annotation-type selections, a method adapted from a study by Laufer and Hill (2000). In the current study, two patterns of annotations selections emerged from the analysis. First, the participants looked up 75% or more of Thai meaning only (T). The second group looked up 75% or more of Thai meaning and 75% or more of other annotation types. Based on these patterns, the participants were divided into three subgroups: those who looked up only Thai meaning (N=47), those who looked up Thai meaning and picture (TP, N=15), and those who looked up Thai meaning, sound, and picture (TSP; N=16).

Table 3. A one-way ANOVA of the three test scores of the three groups

Tests		Sum of Squares	df	Mean Square	F	Sig.
Pretest	Between Groups	12.824	2	6.412	1.216	.302
	Within Groups	395.547	75	5.274		
	Total	408.372	77			
Immediate Post-test	Between Groups	83.983	2	41.992	.420	.659
	Within Groups	7507.350	75	100.098		
	Total	7591.333	77			
Delayed Post-test	Between Groups	19.842	2	9.921	.099	.906
	Within Groups	7490.838	75	99.878		
	Total	7510.679	77			

*p<.01

Next, a one-way ANOVA was used to analyze the three groups' word recall scores. Table 3 shows the results of an analysis of the pre-test, the immediate post-test, and the delayed post-test of the three groups.

Regarding the pretest scores, it is found that there was no significant difference between the scores of the three groups. That is, the participants in each group had no significant differences in their vocabulary knowledge prior to their access of multimedia annotations.

In the immediate post-test and the three-week delayed post-test, however, no significant differences were found in the test scores of the three groups. The Bonferroni analysis also revealed no significant differences among each pair of the three groups in all three tests. In brief, the types of multimedia annotations that the Thai EFL learners selected had no significant effects on their word recall when comparing between learners.

3.3 Multimedia Annotations and Text Recall

A within-subject analysis revealed that after reading the text equipped with multimedia vocabulary annotations, the participants recalled about 30% of the text (Mean= 4.55, SD = 3.645).

Table 4. A one-way ANOVA of the text recall

		Sum of Squares	df	Mean Square	F	Sig.*
Text Recall	Between Groups	70.128	2	35.064	2.759	.070
	Within Groups	953.167	75	12.709		
	Total	1023.295	77			

*$p<.01$

A between-subject analysis was used to compare the text recall scores among the three groups described earlier. Table 4 shows the result of a one-way ANOVA analysis of the text recall scores of the participants in the three groups. No significant difference was found among the three groups' text recall scores.

In sum, Thai EFL learners recalled 30% after they read a multimedia text. But when comparing between learners that used different patterns of annotation selections, there was no significant difference between their text recall score.

4. Discussion

The results of the current study revealed that Thai EFL learners recalled about 42% of words after utilizing multimedia vocabulary annotations. This recall was significantly higher than their pretest scores. This result is in congruent with the results from other studies conducted with learners of other foreign languages. For instance, Chun and Plass (1996) found that L2 learners of German recalled 24-27% of words, while Laufer and Hill (2000) found that Israel EFL learners recalled 33% of words and the Hong Kong students recalled 60% of words. Utilizing multimedia annotations do enhance foreign language learners' ability to recall words.

Furthermore, the current study found that learners were able to recall 30% of words in a production test after three weeks, and this retention is significantly higher than their pre-test scores. This is an addition to the line of research. In the past, there was an argument that L2 learners recalled words better and longer in a recognition test than in a production test. The present study confirms that even in a production test, L2 learners could still recall words with multimedia annotations.

Regarding the types of multimedia annotations that better help learners to recall words, the current study found that EFL learners recalled words with pictured annotations significantly more than words with no picture annotations although they looked up those two groups of words at similar frequency. This finding supports Mayer's Cognitive Theory of Multimedia Learning (2001, 2005, 2009) that learners learn better when they access more than one types of multimedia. The finding agrees with the results from other research studies that compared the groups that had access to textual annotations to those that had access to textual and pictorial annotations (Kost, Foss, & Lenzini, 1999; Yoshii, 2006; Yoshii & Flaitz, 2002).

A between-subject analysis, however, revealed no significant difference between the post-test scores of the three different groups of participants who used different patterns of selection. This finding is incongruent with other studies that compared the test scores between groups that used different types of annotations, for instance, a study by Yanguas (2009) and a study by Jacobs et al. (1994). The possible explanation is the cognitive theory of multimedia learning explains how each individual learner processes information, whereas a between-subject

analysis is a comparison of the effects of multimedia utilization between learners. This variable must be taken into consideration when comparing the types of annotation used by different learner. It is highly likely that multimedia annotations would work better when considering each individual learning rather than comparing between persons.

Regarding text comprehension, the current study found that learners could recall about 31% of the text with multimedia annotations; however, there was no significant difference among the three groups that used different patterns of annotation selections. This finding corresponds with the results from earlier studies, for example a study by Davis and Lyman-Hager's (1997), and a study by Jacobs et al. (1994). However, it contrasts with the result from Yanguas's study (2009), which found that learners who read a text with textual and picture annotations understand the text better than those who read a text with textual annotations only and those who read a text with picture annotations only. The difference in the findings from Yanguas's study and the current study may stem from different types of the reading tests. Yanguas used a recognition test in her study which was less demanding than a production test used in the current study. Furthermore, other variables, for example learners' L2 competency or their background knowledge, may lead to different outcomes in the two studies.

5. Conclusion

The results of the current study support both the theory of multimedia learning and previous empirical studies, therefore instructors should utilize multimedia annotations in their EFL classrooms. It will help enhance learners' ability to learn new words better and faster, and will help them comprehend what they read without stopping to consult teachers, classmates, or dictionary. Ko (1995) and Miall (2000) noted that the utilization of annotations in most textbooks and software were limited to meanings in L1 or L2. Adding multimedia annotations into the texts would help enhance learners' comprehension especially those in the beginning and intermediate levels, whose foreign language proficiency have not reached the critical threshold yet.

In addition, using multimedia in an electronic text makes it possible to bring authentic texts into the classroom. As stated by Krashen (1992), using authentic texts will help increase learners' word knowledge and text comprehension. However, authentic texts are not written to be used as language instructional materials; therefore, L2 readers have to rely on such sources as teachers, friends, or dictionary to find the meaning of unknown words. Multimedia annotations, therefore, will be the best solution for this problem.

Furthermore, the provision of multimedia annotations in the electronic text as used in the current study also encourages learner autonomy. As discussed above, learners will be less likely to rely on teachers, friends, or dictionary. Long-term utilization of multimedia texts will motivate learners to ultimate become an independent learner.

5.1 Limitations and Recommendation for Future Studies

First, the current study investigated five types of multimedia annotations. However, the between-subject analysis relied on the number of students' selections of annotations. Therefore, the numbers of each group are not equal. Different designs may yield different results since Thai EFL students' first language is a tonal language. There is an evidence that learners whose L1 or L2 is tonal tends to rely on sound annotations for meaning as well as other types, especially lower level learners (Gasigijtamrong, 2003). Future study investigating learners with tonal L1 or L2 should add to this line of research.

Second, the current study investigated the effects of using multimedia annotations on intermediate-low level learners of English. Although there was a tendency that students relied on annotations for meaning, the effect of different types of annotations on reading is still inconclusive. Goldman (1996) postulated that there might be greater facilitative effects of multimedia on younger learners than on older learners. Besides, Thai students need to be prepared for their English ability at lower ages. Studies conducted with Thai EFL students at the beginner level, therefore, may yield different results regarding the effects of different types of annotations on reading recall.

References

Akbulut, Y. (2007). Effects of multimedia annotations on incidental vocabulary learning and reading comprehension of advanced learners of English as a foreign language. *Instructional Science, 35*(6), 499-517. http://dx.doi.org/10.1007/s11251-007-9016-7

Al-Seghayer, K. (2001). The effect of multimedia annotation modes on L2 vocabulary acquisition: A comparative study. *Language Learning & Technology, 5*(1), 202-232. Retrieved from http://llt.msu.edu/vol5num1/alseghayer/default.html

Bell, F. L., & LeBlanc, L. B. (2000). The language of glosses in L2 reading on computer: Learners' preferences. Learners' preferences. *Hispania, 83*, 274-285.

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

Davis, J. N. (1989). Facilitating effects of marginal glosses on foreign language reading. *The Modern Language Journal, 73,* 41-48.

Davis, J. N., & Lyman-Hager, M. A. (1997). Computers and L2 reading: Student performances, student attitudes. *Foreign Language Annals, 30,* 58-71.

Goldman, S. R. (1996). Reading, writing, and learning in hypermedia environments. In H. Van Oostendorp, & S. De Mul (Eds.), *Cognitive aspects of electronic text processing* (pp. 7-42). Norwood, NJ: Ablex.

Hegelheimer, V. H. (1998). Effects of textual glosses and sentence-level audio glosses on on-line reading comprehension and vocabulary recall. (Doctoral dissertation, University of Illinois at Urbana-Champaign). *Dissertation Abstracts International, 59*(11), 4053A.

Hulstijn, J. H., Hollander, M., & Greidanus, T. (1996). Incidental vocabulary learning by advanced foreign language students: The influence of marginal glosses, dictionary use, and reoccurrence of unknown words. *The Modern Language Journal, 80*, 327-339.

Jacobs, G. M. (1994). What lurks in the margin: Use of vocabulary glosses as a strategy in second language reading. *Issues in Applied Linguistics, 5*(1), 115-137.

Jacobs, G. M., Dufon, P., & Hong, 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.

Ko, M. H. (1995). Glossing in incidental and intentional learning of foreign language vocabulary and reading. *University of Hawai'i Working Papers in ESL, 13*(2), 49-94.

Ko, M. H. (2005). Glosses, comprehension, and strategy use. *Reading in a Foreign Language, 17*(2). Retrieved from http://nflrc.hawaii.edu/rfl/October2005/ko.html

Kost, C. R., Foss, P., & Lenzini, J. J. (1999). Textual and pictorial glosses: Effectiveness on incidental vocabulary growth when reading in a foreign language. *Foreign Language Annals, 32*, 89-113.

Krashen, S. D. (1992). *Fundamentals of language education.* Torrance, CA: Laredo.

Laufer, B., & Hill, M. (2000). What lexical information do L2 learners select in a CALL dictionary and how does it affect word retention? *Language Learning & Technology, 3*(2), 58-76. Retrieved from http://llt.msu.edu/vol3num2/laufer-hill/index.html

Leffa, V. J. (1992). Making foreign language texts comprehensible for beginners: An experiment with an electronic glossary. *System, 20*, 63-73.

Lomicka, L. L. (1998). "To gloss or not to gloss": An investigation of reading comprehension online. *Language Learning & Technology, 1*(2), 41-50. Retrieved from http://llt.msu.edu/vol1num2/article2/default.html

Mayer, R. E. (2001). *Multimedia learning* (1st ed.). Cambridge, UK: Cambridge University Press.

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

Mayer, R. E. (2009). *Multimedia learning* (2nd ed.). New York: Cambridge University Press.

Mohsen, M. A., & Balakumar, M. (2011). A review of multimedia glosses and their effects on L2 vocabulary Acquisition in CALL literature. *ReCALL, 23,* 35-159. http://dx.doi.org/10.1017/S095834401100005X

Miall, D. S. (2000). *Reading hypertext*. Retrieved from http://www.Ualberta.ca/~dmiall/Brazil/Brazil_hypertext.htm

Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.

Plass, J. L., Chun, D. M., Mayer, R. E., & Leutner, D. (1998). Supporting visual and verbal learning preferences in a second-language multimedia learning environment. *Journal of Educational Psychology, 90*, 25-36.

Rashkovsky, A. L. (1999). The foreign language glossed reader: Authentic literacy text as comprehensible input. (Doctoral dissertation, University of Tennessee, Knoxville). *Dissertation Abstracts International, 60*(9), 3296A.

Stewart, R. A., & Cross, T. L. (1991). The effect of marginal glosses on reading comprehension and retention. *Journal of Reading, 35*, 4-12.

Watanabe, Y. (1997). Input, intake, and retention: Effects of increased processing on incidental learning of foreign language vocabulary. *SSLA, 19*, 287-307.

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), 48-67. Retrieved from http://llt.msu.edu/vol13num2/ yanguas.pdf

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 L2 incidental vocabulary learning. *Language Learning & Technology, 10*(3), 85-101. Retrieved from http://llt.msu.edu/ vol10num3/yoshii.pdf

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

Notes

Note 1. The Flesch-Kincaid readability test is designed to indicate comprehension difficulty when reading a passage of contemporary academic English. The Flesch-Kincaid Readability Level Formula translates the 0-100 score to a U.S. school grade level, making it easier for teachers, parents, librarians, and others to judge the readability level of various books and texts.

Note 2. Log file is a file that lists actions that have occurred. For example, Web servers maintain log files listing every request made to the server. With log file analysis tools, it's possible to get a good idea of where visitors are coming from, how often they return, and how they navigate through a site. In this study, it recorded what words and which type of annotations learners accessed, how often they did so, etc.

Copyrights