The Effects of Glossing Conventions on L2 Vocabulary Recognition and Production

A. A. Zarei [*]
Assistant Professor
Imam Khomeini University, Qazvin
email: aazarei@ikiu.ac.ir

S. Hasani
M. A., TEFL
Islamic Azad University, Takestan
email: s_hasani2000@yahoo.com

Abstract

To investigate the effects of different glossing conventions on vocabulary recognition and recall, 158 participants were given a pre-test to make sure that they did not have any prior knowledge of the target words. Reading passages with four different glossing conventions (interlinear, marginal, pre-text, and post-text) were given to eight groups. Four groups received interlingual glosses and four groups were given intralingual glosses. Receptive and productive post-tests were administered to measure vocabulary recognition and recall. The collected data were analyzed using two one-way ANOVA procedures. The results showed that there were no significant differences among the effects of different types of intralingual glosses on vocabulary recognition and recall. As to the affect of the interlingual glosses on vocabulary recognition, the post-text group performed significantly worse than both the pre-text and the marginal groups. Moreover, the interlinear gloss was shown to be more effective than the post-text gloss in vocabulary recall.

Keywords: glossing conventions, interlingual gloss, intralingual gloss, vocabulary learning

1. Introduction

Vocabulary is without doubt one of the most important components of any language, without which the story of language learning is hardly worth telling. Alessi and Dwyer (2008) agree that the lack of vocabulary and frustration of looking up new words are the main obstacles of second language readers. One of the main ways of overcoming these obstacles is extensive reading for the purpose of the rapid expansion of vocabulary. Pitts, White, and Krashen (1989) agree that reading is an important source of vocabulary acquisition. Day, Omura, and Hiramatsu (1991) also point out the merits of vocabulary learning through reading. Huckin and Coady (1999) also emphasize vocabulary learning through extensive reading and provide three advantages of incidental vocabulary learning over direct instruction contending that it is contextualized, it is pedagogically efficient, and it is more individualized and learner-based.

Despite the obvious merits of extensive reading, it is believed that extensive reading alone is hardly enough, and that it needs to be complemented by some other activities or techniques to enhance the input and to raise the learners' consciousness. While acknowledging the effectiveness of extensive reading, especially at intermediate and advanced levels, Hulstijn, Hollander, and Greidanus (1996) emphasize that meaning-focused reading alone is not enough, and that learners have to pay attention to form-meaning relationship. They refer to some of the limitations of meaning-focused learning that impede learners' acquisition of unknown words in the text. One of these obstacles is lack of attention to unknown words. Another is that even when they notice an unfamiliar word, they ignore it or consider it a known word when actually it is not. Also, they may just pay attention to meaning but not form. The quality of the text is another factor; the meaning of unknown words may not always be inferable. In response to these shortcomings, Nation (2007) underscores the importance of glossing as a facilitating factor. Glossing, as an easy and fast access to vocabulary, is believed to provide word meanings more suitable to the context of what the learner is currently reading (Alessi & Dwyer, 2008).

'Gloss' is defined as a short definition of unfamiliar words (Lomicka, 1998) provided in different parts of a text. There are different categorizations of glossing. One of the determining variables is its position (Roby, 1999); marginal, interlinear, pretext and post text. Moreover, the language of the gloss is another dominant variable (Roby, *ibid.*).

There are arguments both for and against the use of glosses, but according to Yoshii (2006), the question is not whether to gloss or not to gloss. The central questions at present include when, where, and how to gloss. While most studies have investigated gloss types separately, this study will consider the most common glossing conventions simultaneously.

2. Review of Literature

Scholars have defined glossing in quite similar ways. For example, Otto and Hayes (1982) define it as 'adjunct aids' and hold that "The term gloss and glossing are being used to designate and describe the systematic use of marginal notes and other extra-text notations to direct readers' attention while they read" (p. 1). Stewart and cross (1991) offer a similar definition and contend that glosses are mostly used in content area reading texts. Chun and Plass (1996) distinguish between traditional glossing and multimedia glossing and state that "the traditional method for glossing or annotating words is to provide a definition or explanation of the word either in the L_2 or in the native language (L_1) of the readers" (p.183). Similar definitions are offered by Cheng and Good (2009) and Lin and Huang (2008). Lomicka (1998) defines glosses as "short definitions or notes in order to facilitate reading and comprehension processes for L_2 learners" (p. 41) According to Roby (1999), "glosses are many kinds of attempts to supply what is perceived to be deficient in a reader's procedural or declarative knowledge" (p. 96).

The advantages of using glosses in reading materials are relatively well-documented. Otto and Hayes (1982) emphasize applying glosses for two reasons: glosses enhance understanding and help less proficient

learners to optimally use their reading abilities. They also improve comprehension skills and strategy use.

Jacobs (1991) refers to glossing from three different points of view; bottom-up, top-down, and interactive views of reading. To him, from the bottom-up view, glossing is useful for reading comprehension. However, from the top-down view, glossing interferes in comprehension. The interactive view combines bottom-up and top-down processes and holds that both of them are essential. Moreover, Ariew and Ercetin (2004) confirm that less proficient learners exploit more top-down processes compared to more proficient learners. In view of that, according to Ariew and Ercetin, glosses are more useful for more proficient learners who use bottom-up processes.

Nagata (1999) points out some of the advantages of marginal glosses. Firstly, access of marginal glosses is easy compared to dictionary. Secondly, glossing supports consciousness-raising and input enhancement by attracting learners' attention to the target word and improves noticing. According to Ko (2005), one reason for using glosses is preventing incorrect guessing. The second reason is minimizing interruption throughout reading. Promoting comprehension and retention of a text by activating background knowledge and linking it to the new information on the text is the third one. In addition, glossing aids connecting background knowledge and the text as a result of recalling what the learners already know. Last but not least, Ko contends that glosses foster autonomy.

In line with Nagata (1999), Lin and Huang (2008) acknowledge the positive effect of glossing. They hold that glossing makes reading more enjoyable and decreases interruption in reading process and assists form-meaning connection. Moving backward and forward between the text and target words, it facilitates word retention through multiple encountering.

In spite of the fact that glossing has numerous merits for language teaching and learning, it must be noted that too much glossing may hinder students' understanding and too little glossing has no effect at all (Otto & Hayes, 1982). Despite this caveat, the question today is not

whether to gloss or not to gloss; the main issue is the determination of the type of glossing that has the most effect on vocabulary learning.

There are various categorizations of glossing based on different forms, different positions and different languages that are utilized. Categorization based on form includes textual, pictorial and multimedia glosses. Furthermore, there can be single versus multiple-choice glosses and meaning inferred glosses versus meaning given ones. The categorization based on position of glosses is composed of interlinear, marginal, pre-text and post-text glosses. With respect to language, there are two types of L_1 and L_2 glosses.

A number of studies have investigated the effects of various kinds of glossing on different aspects of language learning. Hulstijn et al. (1996) explored the effects of marginal glosses, dictionary look-up, and word frequency on incidental learning. Seventy-eight advanced level Dutch learners of French were provided with an adapted text containing 1306 words. They read the text under one of the three conditions of marginal glosses (L_1 translation), dictionary (bilingual dictionary), and control group (neither marginal glosses, nor dictionary). Three tests were administered. The first one was a recognition and recall test. The second one tested whether the learners knew the target words previously or not. In the third test, again they were supposed to write the meaning of 16 target words, this time not in a context but in isolation. The results suggested that marginal glosses are more efficient than dictionary use because learners rarely used dictionary in texts more than one page. Surprisingly, when learners used dictionary, the results changed and dictionary was proven to be as good as or even better than marginal glosses.

Ko (2005) compared L_1 and L_2 glosses in reading comprehension using qualitative and quantitative measures. The participants were 106 intermediate level Korean undergraduates. They read a text under three conditions of no gloss, L_1 gloss, and L_2 gloss. An advanced level article was used with 931 words of which twenty-two words were glossed. In the quantitative measure of reading comprehension, which was a

multiple-choice reading comprehension test, the L_2 glosses were proved to be more beneficial. But in the qualitative think aloud measure, both types of L_1 and L_2 glosses were shown to be efficient. Accordingly, he concluded that L_1 glosses should not be used for more advanced L_2 learners.

Yoshii (2006) compared the effects of L_1 and L_2 glosses on incidental vocabulary learning in a multimedia environment. The participants of this study were 195 EFL university students. The study was under four conditions of L_1 gloss, L_2 gloss, L_1 textual gloss and pictorial gloss, and L_2 textual gloss and pictorial gloss. The material of the study was a story with 390 words of which 20 were glossed, 14 target words and 6 familiar words. Both L_1 and L_2 glosses were proved to be effective for incidental learning and useful in improving vocabulary learning but regarding long term retention, it was suggested that the L_1 textual glosses led to better retention results compared to L_2 textual glosses or L_2 textual and pictorial glosses.

Lin and Huang (2008) compared the effect of meaning-inferred and meaning-given glosses on learners' incidental vocabulary learning. The participants of the study were 175 high and low proficiency level students in Taiwan. Results confirmed the facilitative effect of glossing on vocabulary learning in that bold faced glosses promoted learners' attention to the new words. They also concluded that although both meaning-inferred and meaning-given glosses can lead to incidental vocabulary learning in meaning-focused reading, meaning-inferred gloss was considered to be more effective in vocabulary gain and retention.

Alessi and Dwyer (2008) compared the effects of vocabulary assistance in different positions. The participants were 76 undergraduate university students in the United States studying Spanish. They were required to read a Spanish newspaper article with vocabulary assistance either before reading, while reading, both before reading and during reading, and without any assistance. They concluded that students who received vocabulary assistance during reading outperformed those who received it before reading. Alessi and Dwyer offer three reasons to

support during reading vocabulary assistance. One is the empirical support for during-reading activities in contrast to theoretical support for pre-reading activities. The other reason is better adjustment of individual differences. And the third important practical merit of during-reading assistance is being 'just-in-time'.

Referring to individual differences, they explain 'adjustment', they maintain that different readers have different background knowledge. As a result, pre-reading activities could only be useful for some readers and waste the time of others. As to being 'just-in-time', they maintain that before-reading vocabulary assistance may be forgotten over time and may not be accessible when it is actually needed. But during-reading vocabulary assistance is available any time during reading.

Al-Jabri (2009) investigated the effects of L_1 and L_2 glosses on reading comprehension and idea recall. Nineteen non-native learners were assigned into three groups of L_1 gloss, L_2 gloss, and no gloss groups. A 470-word English text with 19 glossed words was offered to all the three groups. Although the L_1 group performed better than the L_2 group on immediate multiple-choice reading comprehension test, there were no significant differences between the no-gloss group and the gloss groups. With regard to idea recall, the L_2 gloss group was the least successful group in the recall protocol.

Yanguas (2009) studied the effects of different types of textual, pictorial, and textual-pictorial glosses on text comprehension and vocabulary learning. The participants of this study were 94 university students assigned to four groups based on four conditions of glossing. The material used in the study was an internet-based passage from an online newspaper. The length of the text was 543 words. Both qualitative and quantitative measures were used. The results of both measures showed that all kinds of glosses are better than the no-gloss condition. Regarding the production of the target vocabulary items, the results did not show any significant differences among different conditions. But the combined glossing condition had more effect on comprehension compared to other conditions.

Cheng and Good (2009) explored the effect of glosses on reading comprehension and vocabulary acquisition. They compared three kinds of glosses – first language Chinese glosses together with second language English example sentences, first language in-text glosses, and first language marginal glosses – with no gloss condition. The participants of this study were 135 undergraduate business and engineering students at four proficiency levels. The study consisted of three phases. The first phase had four parts: a vocabulary pre-test, a reading session, a post-test composed of a reading comprehension test, an immediate vocabulary recall test, and a questionnaire about applying vocabulary gloss. The second phase was a delayed vocabulary recall test a week after reading. And the third phase was a second delayed vocabulary recall test two weeks after reading. Each phase was implemented in one session. Participants were randomly assigned to one of the four groups of L_1 gloss together with L_2 example sentence, L_1 in-text gloss, L_1 marginal gloss, and no-gloss condition. The results showed that although on all vocabulary recall tests, the participants in the three gloss provisions were more successful than the no-gloss group, there was no significant effect on reading comprehension. While the scores on the immediate vocabulary recall test was higher, learners' retention declined between the immediate and the first delayed recall tests. However, between the first and second delayed recall tests, a minute increase in retention was observed for all groups. They also confirmed that language proficiency is a factor in glossing. Based on the results of the questionnaire, the participants' attitudes toward glossing were positive.

Xu (2010) explored the effect of three different types of glossing (glossing in both Chinese and English, glossing in Chinese, and glossing in English) on vocabulary acquisition. An immediate retention test and a delayed retention test were administered. It was found that while glossing in Chinese was very effective in immediate retention compared with the two other kinds, it was not effective in delayed retention. English glossing was not very effective in immediate retention. And glossing

both in English and Chinese was equally effective in immediate and delayed retention.

Due to the contradicting results of the previous studies and the fact that they did not consider different types, positions, and languages of glossing together, this study aims to investigate the effects of the most common types of glosses on vocabulary recognition and recall. It intends to answer the following research questions:

1. Are there any significant differences among the effects of various interlingual glossing conventions on L_2 vocabulary recognition?

2. Are there any significant differences among the effects of various interlingual glossing conventions on L_2 vocabulary recall?

3. Are there any significant differences among the effects of various intralingual glossing conventions on L_2 vocabulary recognition?

4. Are there any significant differences among the effects of various intralingual glossing conventions on L_2 vocabulary recall?

3. Method

3.1 Participants

The participants of the present study were 158 male and female B.A freshmen taking a general English course at Imam Khomeini International University. The examination of their textbook (Aroma) showed that they were roughly at pre-intermediate level of proficiency. Out of the 237 participants who initially took part in the vocabulary subtest of a proficiency test, forty were excluded during the homogenization procedure. Another group of 39 participants were excluded due to their failure to complete their cooperation by being absent on the post-tests. The initial sample was, therefore, reduced to 158 participants.

3.2 Instruments

To conduct the present study the following instruments were employed:

1) A vocabulary subtest of a Michigan general proficiency test containing 30 multiple-choice items was used to homogenize the participants and to validate the posttests.

2) A pre-test comprising 145 items was given to all the participants. Each item contained one of the target words and required students to supply the Persian equivalent of the underlined English words in a sentence. It was given to make sure that the participants had no prior knowledge of the target words.

3) The reading comprehension passages for the purpose of this study were selected from 'Reading and Vocabulary Development' series. This series has four levels. Level three, entitled 'Cause and Effect', which roughly corresponded to the learners' level was chosen for the purpose of the present study. The passages were given to the participants under eight conditions. Four groups read the passages under the four conditions of interlinear gloss (IG), marginal gloss (MG), before the text (pre-text) gloss, and after the text (post-text) gloss. The other four groups were given the passages with the same glossing positions. This time, however, glosses were interlingual. In other words, the Persian equivalents of the target words were given as glosses.

Two post tests were also administered to measure the extent of vocabulary recognition and recall.

4) The recognition test included 30 items in multiple choice format.

5) The recall test consisted of 30 fill-in-the-blank items. In the fill-in-the-blank test, the first letter of each word was given along with its translation in Persian. This was done to ensure that the learners could produce the target words and to prevent the possibility of learners providing either partial synonyms or other words that fitted the context without necessarily being the intended words.

3.3 Procedures

Initially, a total number of 237 participants were selected. To homogenize the participants, a 30-item vocabulary subtest of a Michigan Test of English Language Proficiency (MTELP) was administered. As a

result, 40 participants, who had scored more than one standard deviation away from (above or below) the mean, were excluded from subsequent statistical analyses, and there remained 197 approximately homogeneous participants. Out of this sample, 39 participants failed to take part in the post-test of the study. Consequently, 158 participants were taken into account in data analysis.

A pre-test was developed to make sure that the participants had no knowledge of the target words prior to the treatment. The participants were required to supply the Persian equivalent of the underlined English words in 145 sentences. Each sentence contained one of the target words which had been extracted from the reading passages the learners were supposed to receive as treatment. 41 of the target words which turned out to be familiar to more than three participants were excluded from the subsequent lexical recognition and recall post-tests. The reason for the choice of three (and above) correct responses as a criterion for exclusion of items from the posttests was that a larger number of words (78) had been correctly translated by only one or two learners. For practical reasons, not all these words could be omitted; there would not be a sufficient number of words left. Nor could they be substituted because they were extracted from reading passages, and manipulation of texts could damage their authenticity. So, a practical compromise was made to exclude words that had been correctly responded to by at least three people (less than two percent of the sample).

Then, the treatment began, in which different versions of reading passages were given to the experimental groups. Each version consisted of one gloss type. For each word, only one of the related meanings was offered. These glosses were provided through six different reading passages. Each passage was less than two pages long. The participants were supposed to read for the purpose of text comprehension. After reading, three comprehension questions were posed to make sure that the participants read the text. The experimental period lasted for nine weeks, of which six weeks were allocated to the treatment (glossed texts), two weeks to the vocabulary subtest of the Michigan test and the pretest, and

one week to the posttests. It needs to be noted, however, that not all the class time was used for the treatment each session. Since the learners were taking their general English course, only a third of each class time every week (about 45 minutes) was allocated to the experiment.

As the participants were receiving their treatment, two post-tests were developed on the glosses supplied through the six passages: a recognition and a recall post-test. To validate the post-tests, and to avoid creating learner sensitivity toward the target words, the vocabulary subtest of a Michigan proficiency test was given to a group of 30 students with characteristics similar to the target groups (another class at the same university, taking the same course) concurrently with the post-tests. To check the validity of the post-tests, a correlation procedure was used. The validity index turned out to be .79 and .83 in the recall and recognition tests, respectively. The reliability of the post-tests was also estimated using the KR-21formula. The reliability indices of the recall and recognition post-tests were estimated to be .74 and .73, respectively.

The validated post-tests were then administered to the 158 approximately homogeneous participants. The obtained data were then summarized and submitted to statistical analyses. To answer the research questions, four one way ANOVA procedures were used; two of them to measure the recognition and the other ones to measure the recall of vocabulary.

4. Results and Discussions

The first research question sought to investigate the effect of the various interlingual glossing conventions on L_2 vocabulary recognition. To this end, a one-way ANOVA was used. Descriptive statistics are summarized in the following table:

Table 1. Descriptive statistics for the ANOVA on vocabulary recognition

	N	Mean	Std. Deviation	Std. Error	Minimum	Maximum
Marginal	13	6.53	3.40	.94	.00	11.00
Pretext	23	6.65	2.60	.54	1.00	13.00
post-text	15	4.06	2.49	.64	.00	9.00
interlinear	20	5.00	3.72	.83	.00	13.00
Total	71	5.61	3.20	.38	.00	13.00

As it can be seen in Table 1, the pretext group has the highest mean, followed closely by the marginal group. The mean score of the post-text group is noticeably lower than the other groups. Graphic representation of the means of the groups shows the differences among the means more conspicuously.

Chart 1. The means of the groups on vocabulary recognition test
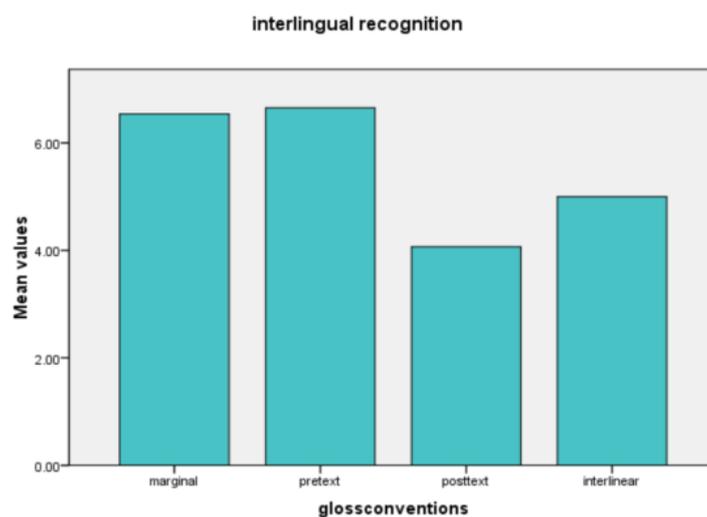
To see whether or not the differences among the groups are statistically significant, the one-way ANOVA procedure was utilized. The results of the ANOVA procedure are given in Table 2.

Table 2. ANOVA on the learners ' vocabulary recognition

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	79.35	3	26.45	2.77	.048
Within Groups	639.38	67	9.54		
Total	718.73	70			

As it can be seen in Table 2, the observed F value and the significance level are indicative of significant differences among the groups. To locate the significant differences, a post hoc Scheffe test was used, the results of which are summarized in Table 3.

Table 3. Multiple comparisons of means for the recognition ANOVA

(I) position	(J) position	Mean Difference (I-J)	Std. Error	Sig.
marginal	Pretext	-.11	1.07	1.00
	post-text	2.47	1.17	.22
	interlinear	1.53	1.10	.58
pretext	marginal	.11	1.07	1.00
	post-text	2.58	1.02	.10
	interlinear	1.65	.944	.38
post-text	interlinear	-.93	1.05	.85

Although the result of the one-way ANOVA showed a significant difference among means, the post hoc Scheffe test was unable to locate the significant differences. This was probably because the differences, though statistically significant, were not big enough to be identifiable.

So, the Duncan test was utilized to locate the differences. The results are shown in Table 4.

Table 4. Multiple comparisons of means for the recognition ANOVA

position	N	Subset for alpha = 0.05	
		1	2
post-text	15	4.06	
interlinear	20	5.00	5.00
marginal	13		6.53
pretext	23		6.65
Sig.		.38	.14

A look at Table 4 makes it clear that the only significant difference is between the post-text group and the marginal and pre-text groups. In other words, the participants of both pre-text and marginal groups have outperformed their counterparts in the post-text group.

The aim of the second question was to investigate the effect of various interlingual glossing conventions on L_2 vocabulary recall. To this end, another one-way ANOVA was used. Descriptive statistics are given in Table 5.

Table 5. Descriptive statistics for the ANOVA on vocabulary recall

	N	Mean	Std. Deviation	Std. Error	Minimum	Maximum
Marginal	13	4.15	2.26	.62	.00	9.00
Pretext	23	9.39	9.09	1.89	.00	28.00
Post-text	15	.86	1.64	.42	.00	6.00
Interlinear	20	13.35	9.41	2.10	2.00	29.00
Total	71	7.74	8.59	1.02	.00	29.00

Table 5 indicates that the highest mean obviously belongs to the interlinear group, followed by the pre-text group. The third highest mean belongs to the marginal group. The post-text group has the lowest mean. Graphic representation of the means of the groups shows the differences among the means more conspicuously.

Chart 2. The means of the groups on vocabulary recall test
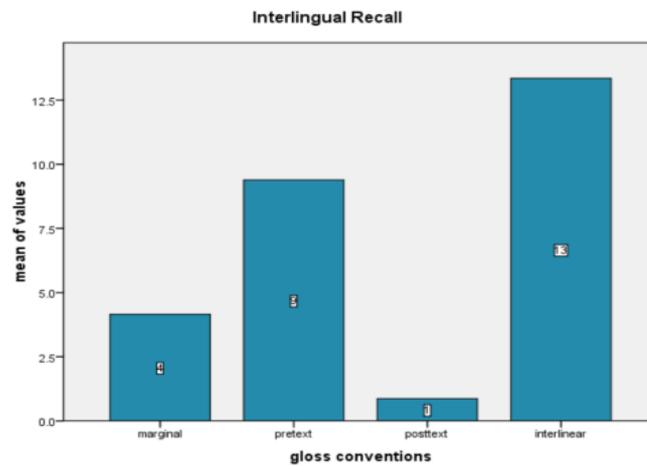
In order to see whether or not the differences between the means are statistically significant, the one-way ANOVA procedure was used. The results are presented in Table 6.

Table 6. ANOVA results on learners' vocabulary recall

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	1567.98	3	522.66	9.71	.000
Within Groups	3603.45	67	53.78		
Total	5171.43	70			

As it can be seen in Table 6, since the F-value of 9.71 is statistically significant (sig. = .000), we can safely claim that there are significant differences among the groups. The post-hoc comparisons of pairs of means (the Scheffe' test) helped to locate the differences as shown in the Table 7.

Based on Table 7, the difference between marginal and interlinear glossing is statistically significant (sig<0.5) and interlinear glossing is more effective in vocabulary recall.

Table 7. Multiple comparisons of means for the recall ANOVA

(I) position	(J) position	Mean Difference (I-J)	Std. Error	Sig.
marginal	Pretext	-5.23	2.54	.247
	Post-text	3.28	2.77	.707
	interlinear	-9.19*	2.61	.010
pretext	Post-text	8.52*	2.43	.010
	interlinear	-3.95	2.24	.381
post-text	interlinear	-12.48*	2.50	.000
*. The mean difference is significant at the 0.05 level.				

The difference between the means of the pre-text and post-text groups is also significant (sig<0.05) with the pre-text group having higher mean. In addition, the difference between the post-text and the interlinear groups is statistically significant (sig<0.05). This means that interlinear glosses have a noticeably better effect on vocabulary recall compared to post-text glosses.

The third research question considered the effects of various intralingual glossing conventions on L_2 vocabulary recognition. The descriptive statistics was as follows:

Table 8. Descriptive statistics for the ANOVA on vocabulary recognition

	N	Mean	Std. Deviation	Std. Error	Minimum	Maximum
Marginal	16	5.81	1.86	.46	3.00	10.00
Pretext	24	6.16	3.45	.70	2.00	16.00
Post-text	19	6.63	3.54	.81	1.00	13.00
Interlinear	28	6.10	4.84	.91	.00	18.00
Total	87	6.18	3.72	.39	.00	18.00

As it can be seen in Table 8, the means of the pre-text, post-text, and interlinear groups are close to each other. The highest mean belongs to the pre-text group, and the lowest belongs to the marginal group. Graphic representation of the means of the groups shows the differences among the means more conspicuously.

Chart 3. The means on the vocabulary recognition test
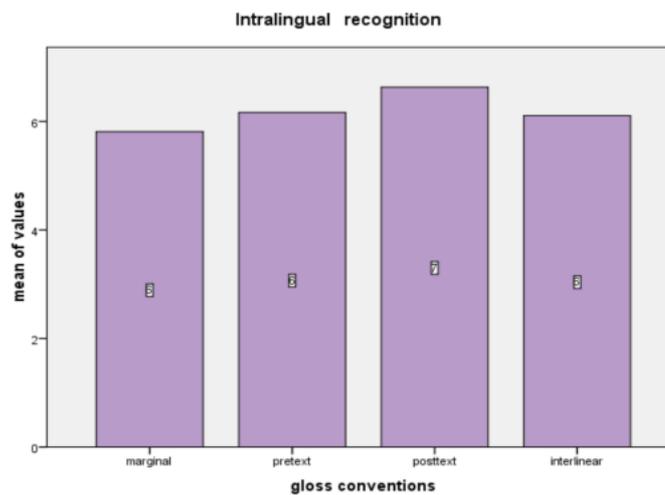
To see whether this difference is statically significant, another one-way ANOVA procedure was used. The results are shown in Table 9.

Table 9. ANOVA results on learners' vocabulary recognition

	Sum of Squares	Df	Mean Square	F	Sig.
Between Groups	6.18	3	2.06	.14	.93
Within Groups	1186.87	83	14.30		
Total	1193.05	86			

It can be observed in Table 9 that there are no significant differences among the groups.

The last research question explored the effects of various intralingual glossing conventions on L_2 vocabulary recall. The descriptive statistics is summarized in Table 10.

Table 10. Descriptive statistics for the ANOVA on vocabulary recall

	N	Mean	Std. Deviation	Std. Error	Minimum	Maximum
Marginal	16	6.56	2.96	.74	3.00	12.00
Pretext	24	5.12	3.45	.70	.00	14.00
Post-text	19	6.73	4.06	.93	.00	18.00
Interlinear	28	4.60	5.22	.98	.00	15.00
Total	87	5.57	4.19	.44	.00	18.00

Graphic representation of the means of the groups makes the comparisons of the means easier.

Chart 4. The means of the groups on vocabulary recall test
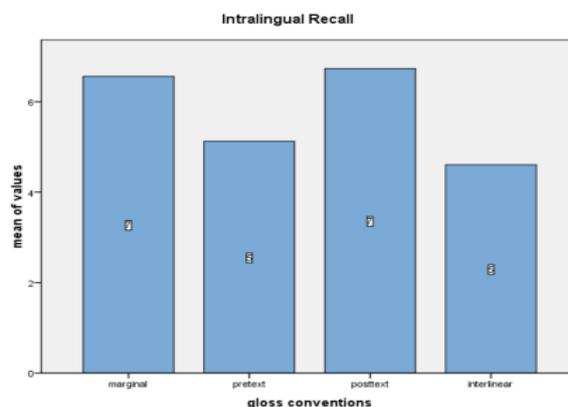
In order to see whether or not the observed differences among the means are statistically significant, the one-way ANOVA procedure was used, the results of which are presented in Table 11.

Table 11. ANOVA results on learners' vocabulary recall

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	72.33	3	24.11	1.38	.25
Within Groups	1440.92	83	17.36		
Total	1513.26	86			

Based on table 11, it can be concluded there are no significant differences among the groups.

The findings of the present study can be discussed in relation to a number of previous studies. For one thing, much like Stewart and Cross (1991), the present study showed that L_1 marginal gloss was more effective than L_1 post-text gloss in vocabulary recognition.

In addition, Alessi and Dwyer (2008) showed that interlinear glosses led to better results than pre-text glosses. In much the same vein, in the

present study, interlinear glosses had more effect on vocabulary recall compared with the other conventions. As was also suggested by Alessi and Dwyer (*ibid*.), the superiority of interlinear glosses may be partially accounted for on grounds of 'being just in time'. The participants' better performance in the interlinear gloss group might also be attributed to the fact that such glosses impede the reading process much less than the other types. The reader does not have to interrupt the reading process to check the meaning of the unknown word elsewhere. In other words, they are not only 'just in time' but only 'in the right place'.

On the other hand, such an outcome might look a bit unexpected since based on schema theory, one expects pre-text glosses to act as advance organizers and create some sort of background which can then be activated during exposure to the words in context to facilitate learning.

Furthermore, Yanguas (2009) showed that there were no significant differences among different types of glossing in vocabulary production. In the same way, the present study did not reveal any statistically significant differences among different intralingual glossing conventions in vocabulary recall. One possible reason why no significant difference was found among the effects of various intralingual glossing conventions on vocabulary recall is that, by their very nature, glosses are made use of when reading, hence improving one's comprehension not production. In other words, while reading, attention is drawn to glosses because they are needed to help resolve certain reading problems or uncertainties. Therefore, the reader might focus only on one aspect of gloss which facilitates comprehension in a particular context and fail to even notice other aspects (presumably, those facilitating production such as spelling, collocations, syntactic properties, etc.). And it is well-documented that for vocabulary learning to take place, the first stage is noticing. In simple terms, since the attention of readers is not drawn to the productive features of words when reading, vocabulary production is not affected by such glosses.

Moreover, unlike Cheng and Good (2009), who report no significant differences among the effects of different glosses on vocabulary

production, this study found significant differences among the effects of the interlingual glossing conventions on vocabulary recall. Although this might seem surprising in the light of the above discussion, it might be justifiable on several grounds. For one thing, it may be argued that in intralingual glosses, readers may grasp the overall meaning of the glossed word to overcome the task of text comprehension, but they might as well fail to grasp the exact meaning of the word. This holds true, especially when there are several partial synonyms in the target language which share the same general area of meaning, but differ in their semantic features, collocations, and syntactic behaviour. Such glosses may cause learners to confuse words with their partially synonymous counterparts in productive use. In interlingual glosses, on the other hand, the exact equivalents of target words given in the native language eradicate such problems. Alternatively, it might stand to reason that when learners get exposed to interlingual glosses, they may be able to draw upon their native language intuitions as to how to make productive use of the target words. Of course, this depends, to a large extent, on the degree of similarities between the two languages with respect to contextual features determining or affecting the productive use of lexical items.

5. Conclusion and Implication

Despite the apparent areas of discrepancy as to which gloss type is more beneficial than other types, there seems to be almost a consensus that, overall, glosses are effective and conducive to language learning in general and vocabulary learning in particular. Yet, the discrepancies among the findings of various studies as to the effectiveness of different glossing conventions on vocabulary learning, coupled with areas of gap between the findings of this study and those of other similar studies warrants more studies in the future.

The findings of the present study can have implications not only for teachers and learners, but also for materials developers. These findings can be used by teachers in reading and even listening activities in the classroom. The results may give teachers hints as to when and where to

offer the meaning of new words; before, during, or after the main text. Although the focus of this study was on reading, it may be useful in both reading and listening contexts. Furthermore, the results may be helpful for improving learner's autonomy. Glosses function as an aid which can compensate for the lack of vocabulary knowledge of learners, who are often dependent on their instructors or dictionary. In addition to teachers and learners, these results can be useful for material developers. The knowledge of the effects of the position and the language of glossing on vocabulary learning can help materials developers make more informed decisions as to whether or not to include glosses in the materials, and if yes, where and how.

References

Alessi, S., & Dwyer, A. (2008). Vocabulary assistance before and during reading. *Reading in a Foreign Language, 20,* 246-263.

Al-Jabri, S. (2009). The effects of L1 and L2 glosses on reading comprehension and recalling ideas by Saudi students. *Umm Al-Qura University Journal of Social Sciences, 1*, 11-27.

Ariew, R., & Ercetin, G. (2004). Exploring the potential of hypermedia annotations for second language reading. *Arizona Working Papers in SLAT, 10,* 1-16.

Cheng, Y. H., & Good, R. L. (2009). L1 glosses: Effects on EFL learners' reading comprehension and vocabulary retention. *Reading in a Foreign Language, 21,* 119-142.

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

Day, R. R., Omura, C., & Hiramatsu, M. (1991). Incidental EFL vocabulary learning and reading. *Reading in a Foreign Language, 7,* 541-551.

Huckin, T., & Coady, J. (1999). Incidental vocabulary acquisition in a second language. *SSLA, 21,* 181-193.

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. (1991). *Second language reading recall as a function of vocabulary glossing and selected other variables.* Unpublished doctoral dissertation, University of Hawaii- US.

Ko, M. H. (2005). Glosses, comprehension, and strategy use. *Reading in a Foreign Language, 17,* 125-143.

Lin, C., & Huang, H. (2008). Meaning-inferred gloss and meaning-given gloss on incidental vocabulary learning. *Journal of national Taiwan normal university, 53,* 87-116.

Lomicka, L. L. (1998). To gloss or not to gloss: An investigation of reading comprehension online. *Language Learning and technology, 1,* 41-50.

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

Nation, P. (2007). Teaching vocabulary. *Asian EFL Journal, 7,* 1-9.

Otto, W., & Hayes, B. (1982). Glossing for improved comprehension: Progress and prospect. *Journal of Reading, 26* (2), 176 – 179.

Pitts, M., White, H., & Krashen, S. (1989). Acquiring second language vocabulary through reading: A replication of the Clockwork Orange study using second language acquirers. *Reading in a Foreign Language, 5,* 571-575.

Roby, W. B. (1999). What's in a gloss?. *Language Learning and Technology, 1,* 94-101.

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

Thornbury, S. (2002). *How to teach vocabulary.* England: Pearson Education.

Xu, X. (2010). The effects of glosses on incidental vocabulary acquisition in reading. *Journal of language teaching and research, 1,* 117-120.

Yanguas, I. (2009). Multimedia glosses and their effect on L2 text comprehension and Vocabulary learning. *Language learning and Technology, 13,* 48-67.

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