THE COMPARATIVE EFFECT OF SINGLE AND MULTIPLE GLOSS CONDITIONS ON EFL LEARNERS' VOCABULARY RETENTION AND PRODUCTION

By

LALEHKHOJASTEH* RAHELEH AKBARPOUR**

** Assistant Professor, Applied Linguistics, Islamic Azad University, Central Tehran Branch.*
*** MA in TEFL, Islamic Azad University, Central Tehran Branch.*

ABSTRACT

Vocabulary glosses are considered effective learning tools since they decrease incorrect meaning inferences from context while keeping reading uninterrupted. The purpose of the present study was to examine the impact of different gloss conditions on intermediate EFL learners' vocabulary retention and production. The participants were 101 EFL learners selected based on their performance on PET and randomly assigned to one control and two experimental groups. In the single gloss experimental group vocabulary was taught using glossed texts with one definition for each gloss while the multiple gloss experimental group received the same texts with three definitions for each gloss. The participants in the control group received the texts without any glosses but were encouraged to consult their dictionaries instead. All the participants took part in vocabulary retention and production posttests after the treatment and those in the two experimental groups answered a questionnaire about their attitudes towards using glosses. The results of the one-way ANOVA confirmed that the two gloss conditions contributed to vocabulary retention and production. Moreover, the results indicated that single gloss group did significantly better on vocabulary retention test. However, no significant difference was found between the single and multiple gloss groups on the vocabulary production test.

Keywords: Gloss Conditions, Single Gloss, Multiple Glosses, Vocabulary Retention, Vocabulary Production.

INTRODUCTION

The size and knowledge of vocabulary have always been the sign of prestige for speakers of any native language. Moreover, vocabulary is of importance not only to the typical language learner (Zimmerman, 1997) but also to the EFL/ESL reader. As Anderson and Freebody (1981) have demonstrated, lexical development and reading comprehension are strongly related. The more words the learners are familiar with, the better and the more effectively they can comprehend texts. Likewise, reading contributes to vocabulary knowledge; it is actually a by-product of reading (Swanborn, 1999). Moreover, vocabulary is central to communication and often seen as the greatest source of problem by second language learners. The centrality of vocabulary to both acquisition and use is also expressed by Hatch (1983) who states that, "When our first goal is communication, when we have little of the new language at our command, it is the lexicon that is crucial … The words … will make basic communication possible" (p. 74).

Learners have always encountered situations in which their comprehension is impeded by a large number of difficult words. In order to help learners to become good and efficient readers, teachers encourage them to learn vocabulary and this is, in most cases, done by memorizing word meanings through word lists. Presenting words in list forms is an efficient study method in which learners can learn large numbers of words in a short time (Meara, 1995). There are drawbacks, however, to using word lists. As Stevick (1989) observed, "If you want to forget something, put it in a list" (as cited in Lewis, 1993, p. 118). This might be due to the fact that through word lists learners have to learn words in a decontextualized and thus inauthentic manner.

Sometimes the learners are told that the best and the easiest option is dictionary consultation, especially a monolingual one. Dictionary use is very helpful to know the definition of new words in a text, but because it is time-consuming, it can make the learners very slow in reading.

Moreover, the learners may be confused with several definitions provided by dictionaries, not knowing which one is appropriate for the text and thus, leading to an error-prone process which requires cognitive sophistication (Ellis, 1995). Some have even gone far to conclude that dictionary look-up may be detrimental to vocabulary acquisition/retention (for example, Laufer & Hill, 2000). There is also the possibility of distracting readers' attention from the text so that they lose track of comprehending the text they read.

Another option is to look for contextual clues and to make inferences. It is possible that the passage provides information about the meaning of a given word and that there are contextual clues to lead to the meaning of a word in the passage; yet sometimes the very clues may mislead the learners. Moreover, learners may not be aware of the types of clues – punctuation, restatement, or examples. They may also not know the meaning of the morphemes within the word to draw a conclusion as what the meaning of the new word is.

In order to tackle the aforementioned problems, vocabulary glosses have been introduced (Hulstijn, 1992; Hulstijn, Hollander, & Greidanus 1996; Watanabe, 1997). It is believed that vocabulary glosses keep reading uninterrupted, decrease incorrect meaning inferences from context, draw learners' attention to the unfamiliar words, and promote greater use of unsimplified texts that might contain too many difficult words for the readers (Nation, 2001). Vocabulary glosses are of two types: single and multiple-choice. In single gloss conditions, only one definition (L1 or L2) is given for each unknown word whereas in multiple-choice conditions, some distracters as well as the correct definition for the unfamiliar word are provided. Therefore, in the single gloss conditions, meaning is given, while in the multiple-choice conditions meaning is inferred and that requires mental effort.

Despite the advantages of glosses, some teachers are worried that learners are deprived of the opportunities to develop their inferential skills when the meanings are given to them directly through glosses. Such being the case, teachers can occasionally implement meaning-inferred glosses to elicit students' mental processing of the words in class, so that correct feedback can be presented immediately after learners finish the meaning-inferred glosses (Hulstijn et al., 1996; Mondria, 2003; Rott, 2005; Rott, William, & Cameron, 2002; Watanabe, 1997). In this way, the influence of wrong inference can be minimized; at the same time, learners can exert more mental effort on the new words, from which their vocabulary learning can be consolidated (Hulstijn, 1992).

Glosses

Glosses are notes that are written in L1 or in a simpler form in L2 to facilitate learners' reading. To attract learners' attention, glossed words or information can be boldface typed or underlined (Roby, 1999). With the provided information next to unknown words, learners know their meanings immediately and proceed with minimum interruption of reading process (Lomicka, 1998; Nagata, 1999). Glosses are viewed as a valuable tool that facilitates reading in a foreign language (Richgels & Hansen, 1984; Watanabe, 1997). They are largely used in textbook materials in which potential unknown words or words of low frequency to L2 learners are included (Davis, 1989).

Traditionally, glosses provided a short definition or note in order to facilitate reading and comprehension processes for L2 learners. Nation (1983) defined glosses as short definitions; Pak (1986) referred to them as explanations of the meanings of words. Typically located in the side or bottom margins, glosses are most often supplied for 'unfamiliar' words, which may help to limit continual dictionary consultation that may hinder and interrupt the L2 reading comprehension process.

Stewart and Cross (1991) noted that with annotated texts "three voices become involved in the reading: the inner voice of the reader, the voice of the author, and the voice of the teacher manifested in the gloss" (p. 5). Moreover, they maintained that "the purpose of glossing is to produce independent readers" (p. 11). Such comments are a reminder of what the ultimate goal of teaching should be. They also argued that glosses have a "focusing effect" (p. 10). This touches on the areas of arousal and selective attention which have been extensively researched within cognitive science.

Consequently, in addition to impacting reading comprehension, glosses can influence vocabulary learning as well. Boldfaced or underlined glosses can make unfamiliar words salient to the learners and lead them to pay more attention to the unknown words, which in turn can enhance their vocabulary learning (Jacobs, Dufon, & Hong, 1994; Kost, Foss, & Lenzini, 1999; Nagata, 1999). The presence of glosses enables learners to look back and forth between the text and target words, which creates multiple encounters of the words to facilitate word retention (Watanabe, 1997). Moreover, Hulstijn et al. (1996) found that incidental vocabulary learning was higher when L2 readers had access to the meanings of words through marginal glosses.

An important controversy in the literature on gloss condition is the number of choices in the multiple-choice glosses. To improve the deficiency of single marginal glosses which offered only one correct meaning and were believed to deprive the chance of inferring, Hulstijn (1992) proposed the use of multiple-choice glosses in which instead of giving the exact meaning or definition for the unfamiliar word, some distracters as well as one correct meaning were given for each newly introduced word. The justification of using such glosses, according to Hulstijn, was that the search and evaluation of the best choice that fits into the context increases the amount of processing. In fact, the design of multiple-choice glosses is based on Hulstijn's (1992, 2001) mental effort hypothesis that states inferring requires mental effort. The greater the mental effort, the better the learner's recall and retention of information will be. It is believed that the more learners try to process the new information they acquire, the better they can promote their long-term memory through the inference from the context.

But the only problem in the case of multiple-choice glosses was with the distracters. Hulstijn noticed that there was a high probability of incorrect inferring among readers when the offered choices were more than two, so he suggested just two choices one of which was the distracter. Likewise, Ke (2003) found out that two choices in the multiple-choice glosses facilitated vocabulary learning, while four options were more effective on correct inferring.

Vocabulary Retention and Vocabulary Production

One way to grasp the overall task of vocabulary learning is through making the distinction between knowing a word and using a word. In other words, the purpose of vocabulary learning should include both remembering words and the ability to use them automatically in a wide range of language contexts when the need arises (McCarthy, 1984). In fact, evidence suggests that the knowledge aspect (both breadth and depth) requires more conscious and explicit learning mechanisms whereas the skill aspect involves mostly implicit learning and memory (Ellis, 1994). Vocabulary learning strategies, therefore, should include strategies for 'using' as well as 'knowing' a word.

Laufer (2005, p. 29) specifies a number of aspects of a word learners need to bear in mind when stating that they have mastery of a word: form, word structure, grammatical features, verb patterns, different meaning types, and others. According to her, such a mastery not only involves learning a word but also remembering it in the future, which is called retention.

Moreover, as mentioned above, in order to claim learning, it seems essential that a learner be able to actively use vocabulary in his/her language production. Vocabulary production is often defined as producing acquired words (through reading) in writing or speaking tasks. Mahyer and Brause (1986) define vocabulary production as, "Being able to spontaneously recall words that are known not only by sight, but that are understood well enough to be used correctly" (p. 392).The impact of a learner's scope and depth of vocabulary on the descriptiveness, accuracy, and quality of his or her writing is noted by Ediger (1999). Ediger asserts, "Variety in selecting words to convey accurate meanings is necessary in speaking and writing, the outgoes of the language arts" (p. 1). Corona, Spangenberger, and Venet (1998) likewise concur that, "At any level, written communication is more effective when a depth of vocabulary and command of language is evident" (p.26).

Another way to view vocabulary learning is to see it as a process of related sub-tasks; thus, "acquiring a vocabulary requires not only labeling but also categorizing skills"

(Thornbury, 2002, p. 18). When learners first encounter a new word, they might guess its meaning and usage from available clues. Some learners might proceed to look it up in the dictionary. Others might take down notes along the margins, between the lines, or on separate vocabulary notebooks. Some learners will repeat the new word a number of times until they are comfortable with it. Others will go beyond simple rote repetition to commit the word to memory. Some would even try to use the word actively. Each of these activities demands metacognitive judgment, choice, and deployment of cognitive strategies for vocabulary learning. And each strategy a learner uses will determine to a large extent how and how well a new word is learned.

One may think that the best way to increase vocabulary knowledge is through copious exposure for example through extensive reading. Although no one can deny the importance of exposure, it does not seem to be the only remedy for a foreign language learner who needs to communicate through a new system. To justify this claim one can refer to Cobb's (2007) corpus analysis of extensive readers which confirmed that the contents of extensive reading materials did not have adequate lexical coverage to take learners beyond the bounds of the most frequent 2000 words. Thus, it seems that exposure needs to come along with intention, systematicity, and strategy.

Although vocabulary retention and production are two separate domains which are equally important, within the extensive literature on the effectiveness of different gloss types, no research has focused on the comparative effects of gloss conditions on EFL learners' vocabulary production and vocabulary retention. The issue is that learners may learn newly introduced words, but the acquired vocabulary items may not be active in their productions. In other words, teachers cannot make sure whether providing learners with reading texts will successfully result in helping them retain as well as produce the new vocabulary items. Moreover, research findings about the effect of meaning-giving (single) glosses and meaning-inferring (multiple-choice) glosses on vocabulary learning are controversial (for example, Mondria, 2003; Nagata, 1999; Rott, 2005; Rott et al., 2002; Watanabe, 1997).

Thus, the current study intended to investigate whether gloss condition had any significant impact on EFL learners' vocabulary retention and written vocabulary production and in case it did, which gloss condition was more effective for vocabulary retention (single gloss condition or multiple-choice gloss condition) and which for written vocabulary production. Moreover, the study also intended to qualitatively seek the attitude and opinion of the participants toward use of glosses. Therefore, the following null hypotheses were stated by the researchers:

$H_{(0)}1$: Using gloss conditions does not have any significant effect on EFL learners' vocabulary retention.

$H_{(0)}2$: Using gloss conditions does not have any significant effect on EFL learners' vocabulary production in written form.

$H_{(0)}3$: There is no significant difference between the effect of single and multiple gloss conditions on EFL learners' vocabulary retention.

$H_{(0)}4$: There is no significant difference between the effect of single and multiple gloss conditions on EFL learners' vocabulary production in written form.

Method

Participants

One hundred and one intermediate EFL students (low to high, male and female) from two branches of a language school in two cities of Iran, namely Chalous and Noshahr, participated in this study. These participants were selected from 149 students who comprised the entire intermediate students of the two language schools during two semesters. The students' English learning experience was at least three years.

Forty eight out of the 149 students were excluded from the main study after participating in the Preliminary English Test (PET). These students obtained scores which fell outside the range of one standard deviation above and below the mean of the sample. Therefore, the study entailed 101 intermediate participants who were homogeneous in terms of their overall proficiency. These students were then randomly assigned to three groups; one control and two experimental groups.

Moreover, another 43 intermediate students who had been

chosen one semester before the commencement of the main study formed the pilot group participants. They were all intermediate students from the Noshahr branch of the language school and took PET and the vocabulary retention posttest in the pilot phase and the researchers used the results to conduct item analysis and reliability estimate in order to modify and improve the two tests.

Instrumentation

The instruments that were utilized in this study can be divided into four main categories: tests, instructional materials, rating scales, and a questionnaire.

Tests

To fulfill the purpose of this study, the researchers used three tests. One of the tests (PET) was used for the purpose of homogenizing the participants; the other one was a vocabulary retention posttest and the last one was a vocabulary production posttest.

The Preliminary English Test (PET) including all four language skills was piloted with 43 intermediate students who bore similar characteristics with the participants of the target sample. The pilot study was done since the test was a mock test and the researchers needed to carry out item analysis to see whether there were any malfunctioning items to be omitted from the test as well as estimating the reliability of the test. The modified PET was then administered to all of the 149 intermediate students who attended the two language schools in two cities in North of Iran. The time allocated for the PET was two hours and eighteen minutes.

Five reading passages were selected and given to the participants in the pilot group in order to extract the vocabularies that the participants at this level did not know and could thus be included in the course content. The procedure for this selection is fully explained in the procedure section below.

Moreover, after the treatment period in order to test the null hypotheses of the research, participants took a vocabulary retention posttest which included 35 fill-in-the-blank items and a word bank including 40 words which were selected from the 62 words that were taught during the course. The purpose of administering this test was to check how many of the newly-learned vocabularies the participants could retain and to test the null hypotheses number 1 and 3 of the study. It has to be mentioned that since this test was developed by the researchers, it was piloted with the same pilot group for item analysis and reliability estimation. Moreover, since the test was meticulously designed in a way to contain a representative sample of the vocabularies taught during the course of instruction, it could be claimed that it had high content validity.

Finally, the participants also took a written vocabulary production posttest in which they were supposed to write 250 words on a topic (discussing life in ones' country to a group of young people while being abroad focusing on customs, tourism, discrimination and government) which was related to the reading topics covered in the treatment period. The purpose of administering this test was to investigate how many of the learned vocabularies the participants could use in their writings and thus, test the null hypotheses 2 and 4 of the study. The researchers set 20-25 newly-learned vocabularies as the expected number of the new words in the participants' writing which was equal to 40 percent of the vocabularies that were taught during the course.

Instructional Materials

Some reading materials and writing tasks were used during the course which are described below.

Reading materials: "Mosaic 1" was used in this study. The researchers chose this book because it is designed for intermediate level and also it was piloted in the researchers' language school in the previous year and the results showed that the texts were interesting for the students and they could use the topics for group discussions and also as a source of learning new information and new vocabulary.

Writing tasks: The participants wrote essays on five topics during the semester. Each writing was done after covering one of the reading passages. Although the participants were always encouraged to use the newly-introduced words in their writings, no exact number of the new words was required of them in order not to distract the participants from the main focus of writing. The topics of the writings were all related to the reading materials so that the participants were given the chance of employing new vocabulary items in their productions.

Rating Scales

For the assessment of parts two and three of the writing section of PET, the researchers used the PET 'general mark scheme'. The band scores were 0-5. The marks given to the speaking section of the PET were awarded by two interviewers including one of the researchers and her colleague on the basis of the following criteria: Grammar and Vocabulary, Discourse Management, Pronunciation, and Interactive Communication.

Questionnaire

The researchers designed a questionnaire and administered it only to the two experimental groups to find out about their attitudes and perspectives toward using glosses during reading. The questionnaire consisted of seven questions, four of them with a Likert scale and the other three followed by two options. The questionnaire basically asked questions such as whether the participants were familiar with glosses before the intervention, the number and type (L1 or L2 and marginal or in-text) of glosses they preferred, whether the glosses were effective in comprehending the text, learning new vocabularies, and retention of the vocabularies while writing.

Procedure

The procedure of the study is reported in several sections according to the nature of the steps taken in the study.

Piloting the instruments as the first step, the researchers piloted two of the instruments to make sure that they were appropriate for the purpose of this study. First, a sample mock PET was given to the 43 intermediate students in the pilot group and based on their performances item analysis was carried out and reliability of the closed-ended items was estimated by Cronbach alpha and inter-rater reliability for the writing and speaking sections of this test was computed by Pearson correlation. Moreover, the vocabulary retention posttest was also piloted for the same purpose with these pilot participants.

Participant selection and homogenization, as the next step, the piloted PET was administered to the 149 intermediate students who were the entire intermediate students studying at the two branches of the language school during the semester and the students whose scores fell between one standard deviation above and below the sample mean were selected as the homogenerous participants of the main study. Then, the selected 101 students were randomly assigned to three groups resulting in 30 participants in the control group, 32 in the single-gloss group, and 39 in the multiple-choice gloss group. The other 48 students were excluded from the main study, but they were not told so. All the students were present in their classes and all of them participated in reading the materials, writing, and the final vocabulary retention and production tests, but only the results of the main participants were analyzed for the present study.

One of the researchers and a colleague of hers, who also taught at the intermediate level, selected and bolded 83 words from the five reading passages of 'Mosaic 1' described earlier. They did that based on their experience of knowing which of the words might probably be unknown to the intermediate students. After piloting the PET with the 43 intermediate students, the researchers gave them the five reading texts. They were asked to read the texts and write the L2 synonyms or the L1 meaning for the bolded words. They had 60 minutes to do the task. Since these students were all intermediate and at the same level with the participants of the main study, the researchers decided that if 75% of the participants of the pilot group could write the meaning/synonym for each bolded word correctly, the word should be omitted from the list of bolded words for the main study and were thus not finally considered for the glosses. In fact, the researchers set the cut score at 75%. This way, the researchers selected the words they had to use during the intervention and thus provide glosses for.

Treatment

The treatment was carried out as part of the syllabus of an English course with focus on all language skills. Each session lasted 90 minutes from which around 30-45 minutes were dedicated to the reading skill. The participants were given the reading texts on 3rd, 7th, 11th, 15th, and 19th sessions under three different conditions. The control group received the reading texts without any glosses. The single-gloss group received the texts with one definition for each bolded word appearing on the right margin and the multiple-choice-gloss group received them with three definitions for each

word appearing on the right margin.

One of the researchers taught the classes. To ensure the instruction was implemented with an identical procedure in all groups except for the gloss condition, a list of steps was provided for the teacher to refer to as a part of a detailed and comprehensive lesson plan.

The participants of the main study were not told and thus, were not aware of the research project in order to let the classes go on as natural as the regular classes as well as avoiding the Hawthorne effect. The students were also not told that they were supposed to take part in a subsequent posttest to measure their vocabulary retention.

During the sessions named above in which reading and vocabulary were practiced for the purpose of the study, the teacher started with a warm-up to activate the students' schemata and prepare them for the reading. While-reading activities and post-reading activities were done in all groups.

In the multiple-choice gloss group, which was one of the two experimental groups, the reading materials included three definitions for each unfamiliar bolded word parallel to that word in the right margin. All three definitions were the correct definitions of the bolded words, but only one of them would fit the context. The students were asked to choose the best choice to fit the text using inferences. Then, they answered some comprehension questions in the form of discussions and received feedback from the teacher. The teacher let the students discuss their choices and as a class they talked about the vocabulary items. This was done to teach the students collocations, parts of speech, and the way to use the vocabulary items correctly in sentences.

In the second experimental group, which was the single-gloss group, the students read the same reading texts aided by a single definition for each bolded word parallel to that word in the right margin. They then answered the comprehension questions posed by the teacher in the form of discussions and received feedback from them. They would then discuss with the help of the teacher the information about the word such as its part of speech and important collocations.

In the control group, the students received no definitions. In fact, they had no glosses. Therefore, they read the texts for comprehension followed by a number of comprehension questions posed by the teacher and received feedback from her. They were encouraged to infer meaning from the text and consult their dictionaries for the unknown words. In fact, the dictionary was the substitute for glosses in this group. However, note has to be taken that for this group the teacher had briefly explained to the students how to use dictionaries to make sure that no one had any problem in this regard.

Moreover, in all the three groups, pre-, while-, and post-reading activities were carried out in same way. In the pre-reading phase, the students became motivated to read the passage and to participate more fully and with greater satisfaction. This was done mainly through brainstorming. In the while-reading phase, the teacher asked the students to read the texts and find out the main ideas and the supporting facts. The teacher helped those who needed assistance. And in the post-reading phase, the teacher clarified the meaning of any unclear parts and their relationships to the overall message. The teacher encouraged the students to ask any questions they had about the passages and then asked some comprehension questions and they would receive some time to discuss the topic.

Finally, in all the three groups, the students were asked to write about the topic presented to them after they were done with each reading text and received feedback from the teacher. All the topics were chosen on the basis of the reading texts. The participants were encouraged to make use of new learned words in their writings. The written productions of the participants in all the three groups were corrected by the teacher and feedback was given to them in the following two sessions. The participants were given no scores, but the percentage of the taught/learned words which were used in the students' writings was calculated and reported to them.

Questionnaire and Posttests

At the end of the treatment phase, a questionnaire was given to all the participants of the two experimental groups in order to find out about their attitudes toward having glosses in their reading texts.

Moreover, a 35-item vocabulary retention test was administered to the participants in all the three groups. The participants were asked to fill in each blank using the words provided in the word bank. There were five extra words in the word bank. The participants received one point for each correct fill-in; therefore, the total score added up to 35. This test was administered to compare the vocabulary retention of the participants in the three groups.

Finally, the participants were all given a topic to write on. Their writings were analyzed based on the percentage of the use of taught/learned words during the treatment to investigate how many of the learned vocabularies the participants would use in their writings. The participants were asked to write at least 250 words on the topic and the expected number of newly-introduced words in their writings was 20 to 25. Note has to be taken that use of each word was credited if the use was appropriate. That is, simply appearance of the new words in the writing was not a manifestation of the learners' active use of that word.

Results

As mentioned before, the first step of the analysis required analyzing the results of the pilot study of PET during which the test was administered to 43 intermediate EFL learners. The results of item analysis demonstrated no malfunctioning item in terms of item facility and discrimination and thus, none were discarded. The Cronbach Alpha as the estimate for the internal consistency of the closed-ended items came out to be 0.77. The inter-rater reliability as computed through Pearson correlation for the writing section of PET came out to be 0.81 and for the speaking 0.79, demonstrating a significant correlation between the ratings of the two raters. The results of the inter-rater reliabilities are presented in Table 1 and Table 2.

	Writing R1	Writing R2
Writing R1 Pearson Correlation	1	.810**
Sig. (2 -tailed)	43	.000
N		43
Writing R2 Pearson Correlation	.810**	1
Sig. (2 -tailed)	.000	43
N	43	

** Correlation is Significant at the 0.01 level (2-tailed)

Table 1. Inter-rater consistency between R1 and R2 for writing section of the PET – pilot study

	Speaking R1	Speaking R2
Speaking R1 Pearson Correlation	1	.790**
Sig. (2 -tailed)	43	.000
N		43
Speaking R2 Pearson Correlation	.790**	1
Sig. (2 -tailed)	.000	43
N	43	

** Correlation is Significant at the 0.01 level (2-tailed)

Table 2. Inter-rater consistency between R1 and R2 for the speaking section of the PET – pilot study

Moreover, since the vocabulary retention posttest was also piloted with the same group, items analysis was also carried out for this test. Few items fell outside the acceptable facility and discrimination indices and were thus modified. The reliability of this test as estimated by Cronbach alpha came out to be 0.83.

Homogenizing the participants

The piloted PET was used to homogenize the participants of the study. After administering the PET to 149 students, descriptive statistics and reliability estimates were obtained. Table 3 demonstrates the descriptive statistics and the Cronbach Alpha index of internal consistency for the language proficiency test which was used for the homogenization of the participants.

The Pearson correlation coefficient for inter-rater reliability between the raters for both the writing and speaking came out to be significant and 0.79 and 0.76, respectively.

The results of the vocabulary retention posttest

Once the treatment was over, the participants took part in two tests as fully described in the procedure section. The performance of the participants in all the three groups on the vocabulary retention posttest was measured. The reliability of this test as estimated by Cronbach alpha came out to be 0.86 in the main administration. Table 4

N	Mean	SD	Max.	Min.	Cronbach Alpha
149	54.85	7.12	66.00	5.00	0.70

Table 3. Descriptive statistics for the PET for homogenization

Group	N	Mean	SD	SEM	Skewness	Std.error of skewness	Skewness ratio
Single gloss	32	15.78	2.93	.51	.394	.414	.95
Multiplechoice gloss	39	13.41	3.51	.56	.520	.378	1.37
control	30	9.74	3.76	.68	1.020	.528	1.93

Table 4. Descriptive statistics for the vocabulary retention posttest

demonstrates the descriptive statistics on the vocabulary retention posttest.

To test the first null hypothesis which stated that, "Using gloss conditions do not have any significant effect on EFL learners' vocabulary retention", an ANOVA had to be run to compare the three groups. To check the assumptions of running an ANOVA, the normality and the homogeneity of variances were also checked. The skewness ratio of all the three groups fell within the acceptable range of ± 1.96 and thus all the scores were normally distributed (Table 4). Furthermore, the results of the Levene's test demonstrated that there was no significant difference among the variances of the three groups (F= 0.95, df=2,89, p = .39 >.05) and thus, homogeneity of variance was assumed.

According to Table 4, the highest mean score was obtained by the multiple choice gloss group and then the single gloss group. The lowest mean score was obtained by the control group. However, to check whether the differences were significant or not ANOVA was run, the results of which are demonstrated in Table 5.

As Table 5 indicates, the ANOVA results showed that the presupposed null hypothesis was rejected ($F_{(2,98)}$ = 24.38, p = .0005 <.05) meaning that the difference observed between sample means was large enough to be attributed to the differences in the treatment condition. Therefore, the mean obtained by the experimental groups on the posttest (15.78 and 13.41) were significantly higher than that obtained by the control group (9.74) (Table 4) meaning that using glosses significantly affected vocabulary retention of the participants.

However, to see the difference between the effect of single and multiple gloss conditions on EFL learners' vocabulary retention, post hoc comparison which compared the means of the three groups with each other was run. Table 6 demonstrates the results.

As demonstrated by Table 6, all mean differences came out to be significant at 0.05 level. That is the mean score of

	Sum of Squares	Df	Mean Square	F	Sig.
Between group	570.70	2	285.35	24.38	.000
Within Group	1146.63	98	11.70		
Total	1717.33	100			

Table 5. ANOVA results on vocabulary retention posttest

Gloss Conditions			Mean Difference	Std. Error	Sig.
1.	2		-6.0392*	.86928	.000
		3	-3.66708*	.83067	.000
2.	1		6.03192*	.86928	.000
		3	2.36484*	.81587	.013
3.	1		3.66708*	.83067	.000
		2	-2.36484*	.81567	.013

* The mean difference is significant at the .05 level

Table 6. Post hoc tests (Multiple Comparisons) vocabulary retention posttest

the single gloss group was significantly higher than that of the multiple-gloss group (mean difference of 2.36, p = .013 <.05), the single-gloss group significantly outperformed the control group (mean difference of 6.04, p = .0005 <.05) and the multiple-choice group significantly outperformed the control group (mean difference of 3.67, p = .0005 <.05). Thus, null hypotheses 1 and 3 were rejected.

The Results of the Written Vocabulary Production Posttest

After the treatment, the participants of the three groups took part in another posttest which was the written production posttest in order to see how much of the vocabularies the participants would actively use in their writing. Table 7 displays the descriptive statistics for the written vocabulary production posttest for all the three groups. The same procedure was used to analyze the obtained data, that is, checking the assumptions of and then running an ANOVA. As shown in Table 7, the scores on this posttest were normally distributed in the three groups (refer to skewness ratios in Table 7). Moreover, Levenes' test also revealed that the assumption of homogeneity of variance was also observed (F= 6.71, df= 2,89, p = .08 >.05).

As Table 7 demonstrates, the highest mean was obtained by the multiple-choice gloss group (17.62), then the single-gloss group (15.44) and finally by the control group (6.53). Table 8 shows the results of the ANOVA which tested

Group	N	Mean	SD	SEM	Skewness	Std.error of skewness	Skewness Ration
Single gloss	32	15.43	4.93	.87	.489	.414	1.18
Multiple-choice gloss	39	17.61	4.10	.65	.354	.378	.93
Control	30	6.53	3.13	.57	.989	.527	1.87

Table 7. Descriptive statistics for the written vocabulary production posttest

	Sum of Squares	Df	Mean Square	F	Sig.
Between group	2234.873	2	1117.437	65.239	.000
Within Group	1678.572	98	17.128		
Total	3913.446	100			

Table 8. ANOVA results on vocabulary retention posttest

whether these differences were significant or not.

The results of the ANOVA on Table 8 demonstrates that the differences came out to be significant ($F_{(2, 98)} = 65.23$, $p = .0005 < .05$), meaning that the second null hypothesis that stated, "Using gloss conditions do not have any significant effect on EFL learners' vocabulary production in written form" was rejected. This indicated that gloss conditions significantly affected the written vocabulary production of the participants of the study.

However, in order to check the difference between the effect of single and multiple gloss conditions on EFL learners' written vocabulary production post hoc comparison was run (Table 9).

As Table 9 indicates, the mean differences between the single-gloss group and the control group came out to be significant (mean difference of 8.9, $p = .0005 < .05$). Moreover the mean difference between the multiple-choice gloss group and the control group also came out to be significant (mean difference of 11.08, $p = .0005 < .05$). However, the mean difference between the two experimental groups did not become significant (mean difference of 2.18, $p = .075 > .05$). Therefore, there was no significant difference between the two gloss conditions in improving the vocabulary production of the participants in the written form and thus null hypothesis number 4 was not rejected.

Analysis of the questionnaire data

Findings related to the qualitative part of the study were based on the questionnaire given to the participants of the two experimental groups after finishing the treatment phase. There were 7 questions which the 71 participants of the two experimental groups were asked to answer. The results are presented as follows.

Gloss Conditions		Mean Difference	Std. Error	Sig.
1.	2	-8.90417*	1.05176	.000
	3	-11.08205*	1.00505	.000
2.	1	8.90417*	1.05176	.000
	3	-2.17788	.98714	.075
3.	1	11.08205*	1.00505	.000
	2	2.17788	.98714	.075

* The mean difference is significant at the .05 level

Table 9. Post Hoc Tests (Multiple Comparisons) vocabulary production posttest

Question 1: Were you familiar with glosses of any kinds?

Three out of seventy-one participants answered yes to this question. These three were university students and for the course of reading they had books in which definitions of difficult words had been presented in the right margin. The definitions, as they said, were in English, but they did not know anything about glosses or if those definitions were even called glosses.

Question 2: Which one do you think would be more effective in text comprehension?

Nineteen participants from the single gloss group believed that only one L2 definition for each difficult word would be effective, while 13 others said that one L1 synonym would help more in comprehending the texts. From among the participants of the multiple-choice gloss group, 16 believed in having only one L2 definition for each difficult word, since they claimed that they were not sure of the right choices they made while reading the texts. Seventeen participants specified that two choices would be less confusing and more retainable than three choices. The other six were satisfied with the three choices. Nobody chose the last option of the question.

Question 3: Do you think glossed words were effective in comprehending the texts?

All participants of the single gloss group but eight strongly agreed that glossed words were effective in comprehending the texts. Three of those eight agreed but not strongly while the other five did not have any opinions. Twenty participants of the multiple-choice gloss group claimed their strong agreement with the glossed words helping comprehending the texts, while 11 strongly disagreed and seven disagreed with glossed words helping them to comprehend the texts.

Question 4: Do you think glossed words were effective in learning the new difficult words?

Eighteen participants of the single gloss group strongly agreed that glossed words helped them learn difficult

words. Seven agreed that they were effective and seven gave no idea. In the multiple-choice gloss group, 22 strongly agreed with the effectiveness of glosses, 11 agreed and six did not have any ideas.

Question 5: Do you think glosses made you pay attention to the new words?

In the single gloss group 28, and in the multiple-choice gloss group all the participants strongly agreed that glosses made them pay attention to the new words. The other four participants of the single gloss group agreed with their effectiveness in making them notice the new words.

Question 6: Which one do you think would be more helpful while reading and comprehending the texts, marginal or in-text glosses?

Twenty five of the subjects in the single gloss group and 21 of the multiple-choice gloss group thought marginal glosses would be more helpful, while seven and 18 other participants from the single and multiple-choice gloss groups, respectably, chose the in-text glosses.

Question 7: Do you think glosses helped you remember the introduced words when writing?

Eight of the single gloss group participants strongly agreed, 13 agreed, and 11 did not have any opinion. In the multiple-choice gloss group, 16 strongly agreed, 12 agreed and 11 did not have any opinion.

Discussion

The findings of this study indicated that glosses, either multiple-choice meaning-inferred or single meaning-given, are beneficial to vocabulary learning. These results are in line with the findings of previous research findings (for example, Hulstijn et al., 1996; Jacobs et al., 1994; Mondria, 2003; Nagata, 1999; Paribakht & Wesche, 1997; Rott, 2005; Rott & William, 2003; Watanabe, 1997; Yoshii, 2006). Some factors can be accountable for this effectiveness. The use of gloss arouses learners' noticing to the target words, which is a crucial process in vocabulary-learning (Schmidt, 1992). Being bold-faced, gloss successfully draws learners' attention, creating an ideal vocabulary-learning condition of "consciousness-raising" and "input-enhancement" (Rutherford & Sharwood Smith, 1985; Sharwood Smith, 1993). Furthermore, the qualitative data gathered through the questionnaire which was given to the two experimental groups to elicit ideas about the use of glosses and answer the qualitative research question supported this finding as majority in both groups mentioned that glosses were effective in learning new words and retaining them while writing and almost all mentioned that glosses made them pay attention to the new words.

Besides its salience of textual input that enhances learners' attention to target words, gloss also helps learners to connect the word form to its meaning immediately, consolidating the form-meaning association, which is a vital component of knowing a word (Rott & William, 2003). Finally, with a view to comprehending the reading material, learners are more likely to read back and forth between the target words and the gloss, triggering more lexical processing. Such lexical processing of the target words is beneficial to vocabulary learning (Jacobs et al., 1994).

The present study also revealed greater effectiveness of single glosses than those of multiple-choice glosses in vocabulary retention. The superiority of single glosses to multiple-choice glosses in vocabulary learning does not correspond to the findings of previous studies (Hulstijn, 1992; Nagata, 1999; Rott, 2005; Rott & William, 2003). However, the reason behind the superiority of the single-gloss group in this study might have been the lack of the requirement for inferencing different meanings as only one synonym for each bolded word was presented in this group. However, the participants in the multiple-choice group had to infer the appropriate meaning for the bolded words from among the three alternatives; this could have made their pace of reading a bit slow or could have led them to choose the wrong definitions. Again the data from the questionnaire supports this interpretation as in the multiple gloss group majority maintained that they preferred either two choices or one for the glosses rather than three. This showed that with more choices the participants got confused.

Finally, the findings showed that the multiple-choice gloss participants performed better than their peers in the single-gloss group and the no-gloss ones in terms of actively using the words in their writings. However, the difference between the multiple-choice and the single-gloss groups was not

statistically significant. In fact, it can be argued on the whole that when learners confront a single-gloss they easily make a one-to-one correspondence between the word and the meaning and this makes later retention of the word easier for them compared to the situation in which they face alternatives for the meaning in the gloss. However, dealing with alternatives in a multiple-choice gloss requires inferring the meaning inside the text and when the students are required to write on a topic related to the previously encountered text, they better recall and thus use the word with the appropriately inferred meaning as the context is similar to what they have faced before. Thus, inference contributes more to active use or production of the word. However, note has to be taken that the contribution of meaning-inferring glosses was not significantly more than that of the meaning-giving gloss in this study.

Finally, based on the data gathered by the questionnaire, majority of the participants claimed they were not familiar with glosses prior to the intervention. Thus their opinion about the impact of glosses as well as the quantitative findings of this study could be attributed to the intervention in this particular study and not their previous exposure to gloss conditions. This was evidence for the internal validity of the research findings.

Conclusion

Since the findings of this study confirmed the benefits of using marginal glosses for both vocabulary retention and production, it can be concluded that learners can learn vocabulary from either single glosses or multiple-choice glosses when they are engaged in reading mainly for comprehension of the text. The bold-faced glosses can trigger learners' noticing of the new words and lead them to pay additional attention to the new words, which in turn facilitates their vocabulary learning. Of the two gloss types, single gloss (meaning-given) was more effective than multiple-choice glosses (meaning-inferred) in eliciting vocabulary gain and retention because of having only one correct or appropriate definition matching the text. At the same time, multiple-choice glosses showed better but not significant effect on learners' vocabulary production in written form. This can be related to more mental processing of words and more involvement load triggered by the decision-making process in inferring word meaning which led them have better use of the newly-acquired words in their writing performance.

No matter what kind of text enhancement teachers employ, one thing to keep in mind is that students should not abandon their inferring ability to learn words incidentally while they are engaged in reading. This study suggests that word processing via inferring the meaning may facilitate students' vocabulary production more than merely giving meaning directly and explicitly to them; however, further studies need to replicate this study to see whether significant differences can be found between single- and multiple-choice-glosses in vocabulary production or not. In other words, further research is still required to determine whether giving students the word definition directly deprives them of the opportunities to have deep processing of the words and thus lead them to inattentively browse the target words when producing them. Moreover, other studies can investigate the impact of different glosses on EFL learners' limited production (as this study considered extended production), reading comprehension, oral production, and delayed vocabulary retention (as immediate retention was the focus here).

Nevertheless, for the time being the implication of the findings of this study for the teachers and syllabus designers is the use of glosses as text enhancement for better and uninterrupted text comprehension and more effective vocabulary learning but along with constant encouragement of the students to infer the word meaning from the context before consulting the meaning of the new words with the glosses or a dictionary.

References

[1]. Anderson, J. C., & Freebody, P. (1981). Vocabulary knowledge. In J. T. Guthrie (Ed.), *Comprehension and Teaching* (pp. 77–117). International Reading Association, Newark, DE.

[2]. Cobb, T. (2007). Computing the vocabulary demands of L2 reading. *Language Learning Technology*, 11(3), 38-63.

[3]. Corona, C. , Spangenberger, S., & Venet, I. (1998). *Improving student writing through a language rich environment*. M.A. Action research Project, St. Xavier

University and IRI/Skylight.

[4]. Davis, J. N. (1989). Facilitating effects of marginal glosses on foreign language reading. *The Modern Language Journal,* 73(1), 41-48.

[5]. Ediger, M. (1999). Reading and Vocabulary Development. *Journal of Instructional psychology*, 26(1), 7-15.

[6]. Ellis, N. C. (1994). Vocabulary acquisition: The implicit ins and outs of explicit cognitive mediation. In. N. Ellis (Ed.), *Implicit and explicit learning of language* (pp. 211-282). London: Academic Press.

[7]. Ellis, N. C. (1995). Vocabulary acquisition: psychological perspectives and pedagogical implications. *The Language Teacher*, 19(2), 12-16.

[8]. Hatch, E. (1983). *Psycholinguistics*: *A Second Language Perspective*. Newbury House, Rowley, MA.

[9]. Hulstijn, J. H. (1992). Retention of inferred and given word meanings: Experiments in incidental vocabulary learning. In. P. J. L. Arnaud & H. Bejoint (Eds.), *Vocabulary and applied linguistics* (pp. 113-125).London: Macmillan.

[10]. Hulstijn, J. H. (2001). Intentional and incidental second language vocabulary learning: A reappraisal of elaboration, rehearsal, and automaticity. In P. Robinson (Ed.), *Cognition and Second Language Instruction* (pp. 258-286). Cambridge: Cambridge University Press.

[11]. 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(3), 327-339.

[12]. Jacobs, G. M., Dufon, P., & Hong, C. H. (1994). L1 and L2 glosses in reading passages: Their effectiveness for increasing comprehension and vocabulary knowledge. *Journal of Research in Reading*, 17, 19-28.

[13]. Ke, Y. (2003). The Effects of Input Modification on Incidental Vocabulary Learning. *Selected Papers from the Twelfth International Symposium on English Teaching and Learning*, 208-209. Taipei: Crane.

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

[15]. Laufer, B., & Hill, M. (2000). What lexical information do L2 learners select in a CALL dictionary and how does it affect retention? *Language Learning and Technology*, 3(2), 58-76.

[16]. Laufer, B. (2005). Lexical frequency profiles: From Monte Carlo to the real World. A response to Meara. *Applied Linguistics*, 26, 581-587.

[17]. Lewis, M. (1993). *The Lexical Approach. Language Teaching* Publications, Hove.

[18]. Lomicka, L. L. (1998). To gloss or not to gloss: An investigation of reading comprehension online. *Language Learning & Technology*, 1(2), 41-50.

[19]. Mayher, J. S., & Brause, R. S. (1986). Learning through teaching: Is testing crippling integrated language education? *Language Arts*, 63(4), 390-96.

[20]. McCarthy, M. J. (1984). A new look at vocabulary in EFL. *Applied Linguistics*, 5, 12-22.

[21]. Meara, P. (1995). The importance of an early emphasis on L2 vocabulary. *The Language Teacher,* 19(2), 8-10.

[22]. Mondria, J. (2003). The effects of inferring, verifying, and memorizing on the retention of L2 word meanings: An experimental comparison of the "meaning-inferred method" and the "meaning-given method". *Studies of Second Language Acquisition*, 25, 473-499.

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

[24]. Nation, I. S. P. (1983). *Teaching and learning vocabulary*. Wellington: English Language Institute, Victoria University.

[25]. Nation, I. S. P. (2001). *Learning Vocabulary in Another Language.* Cambridge: Cambridge University Press.

[26]. Pak, J. (1986). *The effect of vocabulary glossing on ESL reading comprehension*. Unpublished manuscript, University of Hawaii at Manoa.

[27]. Paribakht, T. S., & Wesche, M. (1997). Vocabulary enhancement activities and reading for meaning in second language acquisition. In. J. Coady, & T. Huckin

(Eds.), *Second Language Vocabulary Acquisition* (pp. 174-200). Cambridge: Cambridge University Press.

[28]. Richgels, D. J., & Hansen, R. (1984). Gloss: Helping students apply both skills and strategies in reading content texts. *Journal of Reading*, 27(4), 312-317.

[29]. Roby, W. B. (1999). What's in a gloss? *Language Learning & Technology*, 2(2), 94-101.

[30]. Rott, S. (2005). Processing glosses: A qualitative exploration of how form-meaning connections are established and strengthened. *Reading in a Foreign Language*, 17, 95-124.

[31]. Rott, S., & William, J. (2003). Making form-meaning connection while reading: A qualitative analysis of word processing. *Reading in a Foreign Language*, 15, 45-75.

[32]. Rott, S., Williams, J., & Cameron, R. (2002). The effects of multiple-choice L1 glosses and input-output cycles on lexical acquisition and retention. *Language Teaching Research*, 6, 183-222.

[33]. Rutherford, W., & Sharwood Smith, M. (1985). Consciousness-raising and universal grammar. *Applied Linguistics*, 6, 274-82.

[34]. Schmidt, R. W. (1992). The role of consciousness in second language learning. *Applied Linguistics*, 11,439-453.

[35]. Sharwood Smith, M. (1993). Input enhancement in instructed SLA: Theoretical bases. *Studies in Second Language Acquisition*, 15, 165-179.

[36]. Stevick, E. W. (1989). *Success with foreign languages: Seven who achieved it & what worked for them*. London: Prentice Hall.

[37]. Stewart, R. A., & Cross, T. L. (1991). The effect of marginal glosses on reading comprehension and retention. *Journal of Reading*, 35(1), 4-11.

[38]. Swanborn, M. S. L. (1999). Incidental word learning while reading: A meta-analysis. *Review of Educational Research*, 69(3), 261-286.

[39]. Thornbury, S. (2002). *How to teach vocabulary.* Essex, UK: Longman.

[40]. Watanabe, Y. (1997). Input, intake, and retention: Effects of increased processing on incidental learning of foreign language vocabulary. *Studies in Second Language Acquisition*, 19, 287-307.

[41]. Yoshii, M. (2006). L1 and L2 glosses: Their effects on incidental vocabulary learning. *Language Learning & Technology,* 10(3), 85-101.

[42]. Zimmerman, C. B. (1997). Does reading and interactive vocabulary instruction make a difference? An empirical study. *TESOL Quarterly*, 31, 121-140.

ABOUT THE AUTHORS

Mona Khabiri is working as Assistant Professor of Applied Linguistics at Islamic Azad University, Central Tehran Branch and Director of Journal of English Language Studies (JELS). She mainly teaches Language Testing, Research Methodology, seminar in TEFL issues, and teaching language skills at graduate level and her main areas of interest include Teacher Education, Critical Thinking, Learner's Autonomy, and Task-based Language Teaching and Learning. She has published papers in International and National Academic Journals and presented in several National and International seminars.

Raheleh Akbarpour holds an MA degree in TEFL from Islamic Azad University, Central Tehran Branch and has been teaching English in different cities in North of Iran at language schools and colleges for many years. Her special field of interest is teaching, reading comprehension and vocabulary.