

# A Comparative Study of the Effects of Different Glossing Conditions on EFL Learners' Vocabulary Recall

SAGE Open  
July-September 2016: 1–11  
© The Author(s) 2016  
DOI: 10.1177/2158244016669548  
sgo.sagepub.com  


Tayebeh Danesh<sup>1,2</sup> and Mohammad Taghi Farvardin<sup>2</sup>

## Abstract

This quasi-experimental study examined the effects of different glossing conditions on English as a foreign language (EFL) learners' vocabulary recall. To this end, five glossing conditions were adopted (i.e., inference-gloss-gloss, gloss-retrieval-gloss, inference-gloss-retrieval-gloss, gloss-retrieval-gloss-retrieval, and full glossing). The participants were 140 MA students of Teaching English as a Foreign Language (TEFL). They were randomly assigned to one glossing condition to read an English reading passage. Five target words were glossed in different glossing conditions within a reading passage. To ensure the participants' attention focused on the reading material, the participants were told that a multiple-choice reading comprehension test would be administered. Afterward, two vocabulary tests (i.e., form recall and meaning recall) were conducted. The results of one-way MANOVAs and the post hoc Scheffé tests revealed that the full glossing condition group did significantly better than other glossing groups in vocabulary form recall, whereas the gloss-retrieval-gloss-retrieval condition group outperformed other four groups in vocabulary meaning recall.

## Keywords

EFL learners, glossing conditions, vocabulary recall

## Introduction

Glossing has been found as a means of facilitating vocabulary learning and reading comprehension (e.g., Davis, 1989; Hulstijn, 1992; Jacobs, Dufon, & Hong, 1994; Ko, 2005; Watanabe, 1997). Glossing in second language (L2) learning refers to providing a summary of words' meaning in a text usually found as notes in the margin or between the lines of the text (Richards & Schmidt, 2010). Glosses can be definitions, interpretations, explanations, or translations. They can have positive effects on student's vocabulary learning due to consciousness raising attained from their bolded form, their availability, shortening the time needed for searching for the word in the dictionary, and providing the appropriate meaning for glossed words or vocabulary (Hashemian & Fadaei, 2012). Taking into account the positive findings of the effectiveness of marginal glosses, researchers have recently attempted to determine what types of glosses or glossing conditions generate more positive effects on L2 learners' reading comprehension and vocabulary learning (e.g., Farvardin & Biria, 2011; Huang & Lin, 2014; Ko, 2012).

Laufer and Hulstijn (2001) proposed Involvement Load Hypothesis based on which tasks activating need, search, and evaluation of the meaning of unfamiliar words will lead to higher vocabulary learning than those which do not trigger such processes. Moreover, Laufer and Hulstijn argued that

compared with other tasks, glossing does not trigger search or evaluation processes and encourages weak involvement load. Recently, Huang and Lin (2014) said that "glossing can increase the possibility of learning correct word meanings in context, but leaving words unglossed for learners to infer or retrieve their word meanings may increase the involvement load and mental effort, thus contributing to better retention" (p. 128). Therefore, preceding glossing with inference (i.e., guessing a new word's meaning from context) or following glossing with retrieval (i.e., the act of remembering words' meaning) may complement vocabulary learning (Rott, 2007). Retrieval is defined as the ability to recall words or items after an interval of time (Richards & Schmidt, 2010). It is also suggested that repeated retrieval would decrease the possibility of forgetting an item (Nowzan & Baryaji, 2013).

Among studies done on glossing as a means of vocabulary learning, some found that marginal glossing has a positive effect on vocabulary learning (e.g., Hashemian & Fadaei,

<sup>1</sup>Department of ELT, Khouzestan Science and Research Branch, Islamic Azad University, Ahvaz, Iran

<sup>2</sup>Department of ELT, Ahvaz Branch, Islamic Azad University, Ahvaz, Iran

## Corresponding Author:

Mohammad Taghi Farvardin, Department of English Language Teaching, Ahvaz Branch, Islamic Azad University, Ahvaz, Postal Code: 61349-37333, Iran.  
Email: farvardin@iauhvaz.ac.ir



**Table 1.** Studies on Glossing Conditions and Vocabulary Recall.

Researcher(s)	Year	Participants	Type of instruction/manipulation
Rott, Williams, and Cameron	2002	76 fourth-semester learners of German whose first language was English	L1 multiple-choice marginal gloss and L2 text reconstruction
Rott	2007	54 fourth-semester learners of German whose first language was English	L1 marginal gloss and increasing the frequency of target words
Peters	2014	35 first-year EFL business students whose first language was Dutch	L2 gloss repetition; type of target items (i.e., single words and collocations)
Nowzan and Baryaji	2013	40 intermediate-level EFL learners whose first language was Persian	L2 gloss and retrieval tasks
Goossens, Camp, Verkoeijen, and Tabbers	2014	62 primary school students whose first language was Dutch	L2 gloss and retrieval tasks
Huang and Lin	2014	124 freshmen at the high-intermediate level whose first language was Taiwanese	L1 marginal glossing conditions (i.e., combining inference, retrieval, and glossing)

Note. EFL = English as a foreign language.

2012; Ko, 2012; Rott, 2007; Rott, Williams, & Cameron, 2002), whereas some argued that it does not necessarily lead to vocabulary learning as it does not encourage inferencing or word meaning retrieval (Huang & Lin, 2014).

The effects of combining inferencing with glossing, retrieval with glossing, or the effect of combining these three factors together on L2 vocabulary learning constitute an important research gap. To help overcome and close this gap, the present study has been designed to compare the effects of manipulating three variables (i.e., glossing, inferencing, and word meaning retrieval) on vocabulary recall of English as a foreign language (EFL) learners.

## Literature Review

There have been a number of studies examining the effect of glossing on L2 vocabulary recall. In some studies (e.g., Peters, 2014; Rott, 2007; Rott et al., 2002), texts were manipulated by increasing the number of word occurrences or providing additional semantic clues for target words (e.g., first language [L1] or L2 marginal glosses, appositives, and dictionary access). Table 1 shows a brief description of some previous studies investigating the effect of glossing conditions and vocabulary recall.

Rott et al. (2002) conducted a study to examine the effect of L1 multiple-choice marginal glosses, L2 text reconstruction with opportunities to recheck input, and combination of these two treatments on vocabulary acquisition and retention. Seventy-six learners of German were randomly assigned to one of the four conditions (three experimental and one control). They were selected from fourth-semester language classes. Productive and receptive word gains were tested immediately after the treatment and again 5 weeks later. Findings revealed that the multiple-choice marginal gloss treatment resulted in significantly better productive and receptive word gains immediately after the treatment. However, 5 weeks later, a significant receptive word gain

was found for the combined treatment condition. Later, Rott (2007) examined the effect of increasing the frequency of target words as well as L1 (English) marginal glosses in a reading text on vocabulary learning. The participants were 54 learners of German who were native speakers of English. They were selected from four intact fourth-semester language classes. Three reading conditions were compared: In the first condition, the target words were glossed 4 times in the margin of the text in all word occurrences; in the second condition, the target words were first glossed, then they were not glossed (for meaning retention based on the previously seen gloss), and bolded twice in the third and fourth word occurrences; and in the third condition, the target words were glossed for the first occurrence, and then they were bold-faced without glossing. Rott found that the first and second reading conditions resulted in more productive word gain than the third condition where readers were presented with the target words only once. She finally concluded that the comprehension of main ideas was the highest when the target words were glossed 4 times followed by the gloss-bolding reading condition and the gloss-retrieval task.

In another study, Peters (2014) investigated how form recall is influenced by repetition (1, 3, or 5 number of occurrences), type of target item (single words vs. collocations), and time of posttest administration (immediately or 1 week after the learning session) for EFL learners. The participants were 35 first-year EFL business students from two parallel classes at a Flemish university. The participants' first language was Dutch. A word list including 12 single words and 12 collocations accompanied by their English (L2) definition was the learning material. The participants were asked to do eight written vocabulary exercises, in which they had to supply the correct form (i.e., to provide synonyms, antonyms, hyponyms, or superordinates; to match collocations; to fill in the gap in sentences; to supply the correct item when given a definition) of the target items. Except for the matching exercise, the participants always had to provide the target item's form. The

participants' learning was measured via two form recall tests, which were administered immediately and 2 weeks after the experimental session. The results showed a large effect of repetition and also the durability of repetition effect in form recall of target words. Moreover, it was found that collocations were more difficult to learn than single words.

Nowzan and Baryaji (2013) further examined the effectiveness of a retrieval task on Iranian EFL learners' vocabulary learning. They selected 40 EFL learners at the intermediate level based on their performance on a preliminary English placement test. Then, a teacher-made vocabulary pretest was administrated to extract participants' unknown words. Afterward, the participants were randomly assigned to two groups: experimental and control groups. The experimental group was taught target words through a retrieval activity that consisted of four steps: learning phase in which a list of target words were presented to the participants with definitions and synonyms; retrieval activity phase, in which participants were required to retrieve the target words according to clues consisting of the first letters of target vocabulary; retrieval interval test phase in which participants were asked to provide the synonym of the target words according to the given clues consisting of first two letters of target words; and posttest phase that was provided for both controlled and experimental groups to assess their vocabulary learning. The participants in the control group received the same content without any retrieval activity processes above. Finally, a posttest was given in both groups at the end of the instruction. The results revealed that retrieval activity had a significant effect on Iranian EFL learners' vocabulary learning. It was also concluded that repeated retrieval can help vocabulary learning, and it can be helpful when long-term retention of information is the goal. They also suggested that the act of retrieval can fix items in the memory, improve memory traces, and increase the accessibility of information in the memory.

In the same vein, Goossens, Camp, Verkoeijen, and Tabbers (2014) conducted a study on the role of retrieval practice in Dutch primary school vocabulary learning. To this end, 62 participants were selected from three primary schools; they were required to learn the meaning of 20 target English words. The participants learned the meaning of 10 words, on one hand, by retrieval practice in which words and their synonyms were shown for 8 s, and the participants were asked to read them aloud and then try to retrieve their synonyms. The other 10 words, on the other hand, were taught by restudy in which the participants read the story or list of target words again and did a puzzle book after a 2-min break. After 1 week, learners' recall of the target words was measured. The results revealed that words learned by retrieval practice were statistically recalled better than words learned by restudy.

Recently, Huang and Lin (2014) introduced a new approach to using L1 (Taiwanese) glosses. They agreed that glossing facilitates the creation of correct form-meaning connections but does not encourage inference or retrieval of words' meaning. Thus, they investigated the effects of combining inferring

and word meaning retrieval on vocabulary learning. One hundred twenty-four freshmen at the high-intermediate proficiency level were selected from a university in Taiwan. A reading text was created by combining multiple sources, including three online articles and two articles from *CNN English magazine*, and eight target words judged to be difficult for upper-intermediate students were selected; each of them appeared 3 times in the reading text. Then, the text was revised to reflect three different glossing conditions (i.e., inference-gloss-gloss, gloss-retrieval-gloss, and full glossing). Target words were either glossed in Chinese or unglossed depending on the different conditions. In the inference-gloss-gloss condition, the participants inferred the word meaning in the first word encounter and were provided with glosses for the subsequent two word encounters; in the gloss-retrieval-gloss condition, the participants were provided with glosses for the first and last word encounters but were expected to retrieve the word meaning in the second word encounter; and in the full glossing condition, the participants were provided with glosses for all the word encounters. Participants were assigned randomly in different glossing conditions. Immediately after reading the text, they received a reading comprehension test and an immediate vocabulary test, and 2 weeks later, they received delayed vocabulary posttest. Results showed that the gloss-retrieval-gloss condition was statistically the most effective condition for improving vocabulary learning.

Previous studies have examined the effects of textual glossing on L2 vocabulary learning and reading comprehension. These studies have supported the effectiveness of textual glosses in facilitating L2 vocabulary learning. Moreover, most studies have investigated the role of different factors such as repeating target words in the text (Peters, 2014; Rott, 2007) and promoting vocabulary recall (Goossens et al., 2014; Nowzan & Baryaji, 2013; Rott, 2007) on L2 vocabulary learning; however, the role of combining such interventions has been little examined so far (Huang & Lin, 2014). The present study, therefore, aimed to fill this gap in the literature by combining inference, retrieval, and glosses together in vocabulary learning and examining the effects of such glossing conditions on vocabulary recall. Thus, the following research question was addressed:

**Research Question:** Do different glossing conditions (such as those investigated here, inference-gloss-gloss, gloss-retrieval-gloss, inference-gloss-retrieval-gloss, gloss-retrieval-gloss-retrieval, and full glossing) have different effects on EFL learners' vocabulary recall?

## Methods

### Participants

One hundred forty (98 female and 42 male) MA students of Teaching English as a Foreign Language (TEFL) in five intact classes were selected to participate in this study. The reason

behind selecting MA students of TEFL was their accessibility. The setting of the present study was Islamic Azad University of Ahvaz, Khuzestan, Iran. The participants' ages ranged from 22 to 40 ( $M = 30.57$ ,  $SD = 4.03$ ). The participants were randomly assigned to one of the five glossing conditions (i.e., inference-gloss-gloss, gloss-retrieval-gloss, inference-gloss-retrieval-gloss, gloss-retrieval-gloss-retrieval, and full glossing) with the same number of participants ( $n = 28$ ) in each group.

## Instruments

**Vocabulary pretest.** To ensure that the target words to be glossed were unknown to the participants, a vocabulary pretest was given to them. The pretest included 20 words selected from the reading section of Barron's TOEFL iBT (Sharpe, 2010). The participants were required to provide the Persian translation of the words. Finally, five words unknown to all participants were chosen as the target words.

**Reading passage and target words.** The reading passage, 'Thermoregulation', was adopted from Barron's TOEFL iBT (Sharpe, 2010; see Appendix A). Despite the reading passage title, it did not have much technical vocabulary in it. This passage was selected based on different criteria: difficulty level, text length, syntactic complexity, and content. The reading passage had 784 words (including the title), six paragraphs, at a Flesch Reading Ease of 42.6 and Flesch-Kincaid Grade Level of 13 indicating that the passage was suitable for the university students (Zamanian & Heydari, 2012). The passage was also given to two experienced EFL professors who taught MA courses, and they judged its difficulty appropriate for the participants. Thus, it was assumed that the participants were comfortable with the difficulty level of the passage. Five target words, included three nouns and two verbs, were selected based on the pretest's results. The five target words were terrestrial, endotherms, blubber, insulation, and vasodilation. The reading passage was modified in a way that each target word appeared 3 or 4 times in the reading passage based on the intended glossing condition. Thus, the reading passage differed in terms of text length, target word occurrences, and number of paragraphs in different glossing conditions. A target word did not appear twice in the same paragraph. Each paragraph appeared on one page to ensure that participants could not see the target gloss during inferring or retrieval (recall). Moreover, before moving to the next paragraph on the next page, the participants were told to put the previous paper on the ground and do not return to the previous page.

**Glossing conditions.** Target words were boldfaced in the passage, and five different L1 (Persian) marginal gloss conditions were included to investigate whether they had different effects on L2 vocabulary recall.

In the inference-gloss-gloss condition, each target word appeared 3 times in the passage; the participants were expected to infer the meanings of target words (based on their surrounding sentences) in the first word encounter in the passage; then, target words were glossed in the second and the third word occurrence.

In the gloss-retrieval-gloss condition, each target word appeared 3 times in the passage; they were glossed in the first and the third occurrence, but in the second word occurrence, the participants were expected to recall the word's meaning (based on the previously seen glosses in the first word encounter).

In the inference-gloss-retrieval-gloss condition, each target word appeared 4 times in the passage; the participants were expected to infer the word's meaning in the first occurrence (based on the context); then the word's meaning was presented as a gloss in the second occurrence; in the third word's occurrence, the participants were expected to recall the meaning again (based on the previously seen gloss in the second word encounter). Finally, for the last occurrence of the words, the meaning of target words was presented as a gloss (see Appendix B).

In the gloss-retrieval-gloss-retrieval condition, each target word appeared 4 times in the passage; they were glossed in the first occurrence; then participants were expected to recall the word's meaning in the second occurrence; again, glosses were presented in the third occurrence, and the participants were expected to recall the word's meaning in the last occurrence.

In the full glossing condition, each target word appeared 3 times in the passage; they were glossed in all occurrences, and so no inferencing or recalling had to be done by the readers.

**Multiple-choice reading comprehension test.** Immediately following the reading, the participants were asked to answer a reading comprehension test including four multiple-choice questions. For each question, the participants were asked to choose the best answer from the four given choices. The questions were selected from the original reading comprehension test in Barron's TOEFL iBT (Sharpe, 2010).

**Vocabulary posttests.** The vocabulary posttests consisted of a form recall test and a meaning recall test. In cognitive psychology, recall test is a test of memory in which participants are presented with stimuli and then, after a delay, are asked to remember as many of the stimuli as possible (Goldstein, 2011). Following Huang and Lin's (2014) study, the form recall test, on one hand, required the participants to translate L1 (Persian) words into their corresponding L2 (English) target words in the passage (see Appendix C). On the other hand, the meaning recall test required the participants to translate the L2 (English) target words into L1 (Persian) (see Appendix D). Each test consisted of 10 items, the five target words and five distractors in a mixed order. There were five



target items in form recall and five target items in meaning recall posttest. Distractors were neglected in scoring procedure. A correct answer received 1 point, and incorrect one received no point. Answers containing minor spelling mistakes such as “imediatly” (instead of “immediately”) were scored as correct.

### Procedures

The study was conducted during the participants' regular class times, and it required two sessions. In the first session, a vocabulary pretest was administered to the participants. In the second session, the participants read a passage and then answered a reading comprehension test and a vocabulary posttest. Before the beginning of the study, the reading passage was modified to ensure that the participants could not see the glosses during inferring or retrieval. A target word did not appear twice in the same paragraph, and once participants moved on to the next paragraph on the next page, they could not return to the previous one.

The different glossing conditions were predefined, and the reading passage was modified for each of the five different glossing conditions. The participants were randomly assigned to one of the five glossing conditions. Each group was asked to read one passage containing the target words that had been adopted for the glossing conditions they were assigned to. The participants were instructed to read for comprehension. After reading the passage, the participants completed the reading comprehension test and then the vocabulary posttest. The whole intervention session required approximately 30 min. The reading comprehension test was used to ensure that the participants would not skim through the passage.

### Data Analysis

All scores were recorded in SPSS, Version 21. To answer the research question, one-way MANOVA was conducted.

### Results

In this section, the statistical analyses are described. The discussion of the results can be found in a following section.

#### Results of Descriptive Statistics

Table 2 shows the descriptive statistics of the five groups in their vocabulary form and meaning recall posttests.

As shown in Table 2, the study had five groups ( $n = 28$  each) with different glossing conditions (i.e., inference-gloss-gloss, gloss-retrieval-gloss, inference-gloss-retrieval-gloss, gloss-retrieval-gloss-retrieval, and full glossing). In the form recall test, the lowest mean score belonged to gloss-retrieval-gloss group ( $M = 2.57$ ,  $SD = 1.133$ ), whereas the full glossing group obtained the highest mean score

**Table 2.** Descriptive Statistics of the Scores on Form and Meaning Recall Tests.

Gloss conditions	<i>M</i>	<i>SD</i>	<i>n</i>
<b>Form recall test</b>			
Inference-gloss-gloss	3.54	1.105	28
Gloss-retrieval-gloss	2.57	1.317	28
Inference-gloss-retrieval-gloss	3.61	1.133	28
Gloss-retrieval-gloss-retrieval	4.04	0.922	28
Full glossing	4.21	0.995	28
Total	3.59	1.229	140
<b>Meaning recall test</b>			
Inference-gloss-gloss	4.21	0.686	28
Gloss-retrieval-gloss	3.57	1.136	28
Inference-gloss-retrieval-gloss	4.29	0.535	28
Gloss-retrieval-gloss-retrieval	4.32	0.819	28
Full glossing	4.25	0.518	28
Total	4.13	0.812	140

**Table 3.** Levene's Test of Equality of Error Variances.

Test	<i>F</i>	<i>df1</i>	<i>df2</i>	Significance
Form recall test	1.216	4	135	.31
Meaning recall test	2.45	4	135	.08

( $M = 4.21$ ,  $SD = 0.995$ ). The total mean for all conditions for the form recall test was 3.59, and the total standard deviation was 1.229. In the meaning recall test, the lowest mean score was related to gloss-retrieval-gloss group ( $M = 3.57$ ,  $SD = 1.136$ ). However, the gloss-retrieval-gloss-retrieval group obtained the highest mean score ( $M = 4.32$ ,  $SD = 0.819$ ). Total mean for the meaning recall test was 4.13, and the total standard deviation was 0.812. The participants' overall mean score was higher in the meaning recall test ( $M = 4.13$ ,  $SD = 0.812$ ) than in the form recall test ( $M = 3.59$ ,  $SD = 1.229$ ).

#### Results of Inferential Statistics

A one-way between-groups MANOVA was conducted to investigate the effects of different glossing conditions on vocabulary recall. The dependent variables were participants' scores on form and meaning recall tests. The independent variable was glossing conditions. Table 3 shows the results of Levene's test of equality of variances across all groups.

As shown in Table 3, the assumption for between-groups homogeneity of variance across different glossing conditions in vocabulary form and meaning recall tests is met as each  $p$  value is greater than .05. Table 4 depicts the multivariate test of significance for the immediate posttest. This test indicates whether there are statistically significant differences among the gloss groups on a linear combination of vocabulary form and meaning recall tests.

**Table 4.** Multivariate Test of Immediate Vocabulary Posttests.

Effect		Value	<i>F</i>	Hypothesis <i>df</i>	Error <i>df</i>	Significance	Partial $\eta^2$
Intercept	Wilks's lambda	.031	2,094.969	2.000	134.000	.000	<b>0.969</b>
Gloss conditions	Wilks's lambda	.750	5.195	8.000	268.000	.000	<b>0.134</b>

**Table 5.** MANOVA Tests of Between-Subjects Effects.

Source	Dependent variable	Type III sum of squares	<i>df</i>	Mean square	<i>F</i>	Significance	Partial $\eta^2$
Intercept	Form recall test	1,807.207	1	1,807.207	1,486.02	.000	0.917
	Meaning recall test	2,386.314	1	2,386.314	3,994.80	.000	0.967
Gloss conditions	Form recall test	45.614	4	11.404	9.377	.000	<b>0.217</b>
	Meaning recall test	11.043	4	2.761	4.622	.002	<b>0.120</b>
Error	Form recall test	164.179	135	1.216			
	Meaning recall test	80.643	135	0.597			
Total	Form recall test	2,017.000	140				
	Meaning recall test	2,478.000	140				
Corrected total	Form recall test	209.793	139				
	Meaning recall test	91.686	139				

According to Table 4, there was a statistically significant difference between glossing conditions on vocabulary recall,  $F(8, 268) = 5.195$ ,  $p = .000$ ; Wilks's lambda = 0.750; partial  $\eta^2 = 0.134$ . As probability value is less than .05, it can be concluded that the participants' performance on vocabulary form and meaning recall tests was significantly dependent on glossing conditions. Table 5 displays the results of MANOVA on different groups' performance on the form and meaning recall tests.

Table 5 shows that for the form recall test,  $F(4, 135) = 9.377$ ,  $p = .000$ , and partial  $\eta^2 = 0.217$ . For the meaning recall test,  $F(4, 135) = 4.622$ ,  $p = .002$ , and partial  $\eta^2 = 0.120$ . As the probability value is less than .05, it can be concluded that results for both vocabulary form and meaning recall immediate posttests are significant; thus, gloss conditions had significant effect on vocabulary recall. Scheffé post hoc tests were used to find where the specific differences were located. Table 6 presents the results of Scheffé post hoc tests on vocabulary tests across gloss conditions.

As Table 6 shows, for the form recall test, the inference-gloss-gloss group outperformed the gloss-retrieval-gloss group ( $MD = 0.96$ ,  $p = .035$ ). Moreover, the inference-gloss-retrieval-gloss group did better than the gloss-retrieval-gloss group ( $MD = -1.04$ ,  $p = .018$ ). The gloss-retrieval-gloss-retrieval group surpassed the gloss-retrieval-gloss group ( $MD = -1.46$ ,  $p = .000$ ). Also, the full glossing group outperformed the gloss-retrieval-gloss group ( $MD = -1.64$ ,  $p = .000$ ). As depicted in Table 5, for the meaning recall test, the inference-gloss-retrieval-gloss group did better than the gloss-retrieval-gloss group ( $MD = -0.71$ ,  $p = .021$ ). Moreover, the gloss-retrieval-gloss-retrieval group surpassed the gloss-retrieval-gloss group ( $MD = -0.75$ ,  $p = .013$ ). Also, the full

glossing group outperformed the gloss-retrieval-gloss group ( $MD = -0.68$ ,  $p = .033$ ).

## Discussion

### Form Recall Test

The results of the study revealed that all gloss conditions yielded different vocabulary recall scores. According to the results of the form recall test, of the five interventions, the full glossing condition most benefited vocabulary learning, followed by gloss-retrieval-gloss-retrieval, inference-gloss-retrieval-gloss, inference-gloss-gloss, and gloss-retrieval-gloss conditions. In addition, the difference between full glossing and gloss-retrieval-gloss conditions was significant (Table 6).

The results obtained for the full glossing condition contradict Huang and Lin's (2014) contention that repeated encounters with glosses may not necessarily lead to learners' active processing of glosses in the margin; their study revealed that the full glossing condition generated the least vocabulary learning in comparison with the other glossing conditions (gloss-retrieval-gloss and inference-gloss-gloss conditions). In this study, the full glossing condition outperformed the other four glossing conditions on the vocabulary form recall test. This confirms Rott's (2007) claim that repeated encounters with glosses could offer learners repeated opportunities to isolate the glossed word from its surrounding context to rehearse the word meanings in their working memory.

The results also showed that the gloss-retrieval-gloss-retrieval condition was the second best glossing arrangement for a word that occurred in the passage according to the vocabulary form recall test. These findings support the claim

**Table 6.** The Results of Scheffé Post Hoc Tests on Immediate Vocabulary Tests Across Gloss Conditions.

(I) Gloss conditions	(J) Gloss conditions	Mean difference (I – J)	Significance
<b>Form recall test</b>			
Inference-gloss-gloss	Gloss-retrieval-gloss	0.96	<b>.035<sup>a</sup></b>
	Inference-gloss-retrieval-gloss	–0.07	1.000
	Gloss-retrieval-gloss-retrieval	–0.50	.580
	Full glossing	–0.68	.264
Gloss-retrieval-gloss	Inference-gloss-retrieval-gloss	–1.04	<b>.018<sup>a</sup></b>
	Gloss-retrieval-gloss-retrieval	–1.46	<b>.000<sup>b</sup></b>
	Full glossing	–1.64	<b>.000<sup>b</sup></b>
Inference-gloss-retrieval-gloss	Gloss-retrieval-gloss-retrieval	–0.43	.715
	Full glossing	–0.61	.378
Gloss-retrieval-gloss-retrieval	Full glossing	–0.18	.985
<b>Meaning recall test</b>			
Inference-gloss-gloss	Gloss-retrieval-gloss	0.64	.051
	Inference-gloss-retrieval-gloss	–0.07	.998
	Gloss-retrieval-gloss-retrieval	–0.11	.992
	Full glossing	–0.04	1.000
Gloss-retrieval-gloss	Inference-gloss-retrieval-gloss	–0.71	<b>.021<sup>a</sup></b>
	Gloss-retrieval-gloss-retrieval	–0.75	<b>.013<sup>a</sup></b>
	Full glossing	–0.68	<b>.033<sup>a</sup></b>
Inference-gloss-retrieval-gloss	Gloss-retrieval-gloss-retrieval	–0.04	1.000
	Full glossing	0.04	1.000
Gloss-retrieval-gloss-retrieval	Full glossing	0.07	.998

<sup>a</sup> $p < .05$ .<sup>b</sup> $p < .001$ .

that retrieval can play a positive role in enhancing vocabulary learning (Goossens et al., 2014). Moreover, the gloss-retrieval-gloss-retrieval condition's significant mean difference with the gloss-retrieval-gloss condition in the form recall test would be related to additional reviewing of the target items (Rott, 2007).

The inference-gloss-retrieval-gloss condition served as the third most beneficial glossing condition on the vocabulary form recall test. The results of this study confirm Rott's (2007) claim that repeated visual enhancement itself seems to have no effect on extending word processing and vocabulary learning. Participants might ignore the first contextual clues provided in the first condition in the inference-gloss-retrieval-gloss condition, as they did not have any background knowledge about the topic of the reading text or even about the test's goal which was vocabulary learning (Zaid, 2009). On the other hand, the inference-gloss-retrieval-gloss condition significantly outperformed the gloss-retrieval-gloss condition. This could be attributed to the participants' increased attention because of increased target word occurrence and providing a bolded form for the target words.

As for the form recall test, the inference-gloss-gloss condition was the fourth best intervention condition in this study. It generated less word learning than the full glossing condition. The lower vocabulary recall in the inference-gloss-gloss condition than in the full glossing condition could be linked to Zaid's (2009) assumption that presenting target words with contextual clues should help with inferencing meanings of

new items, but it depends heavily on the background knowledge. Thus, it may not lead to successful meaning guessing, whereas providing the correct word meaning in accordance with context would result in more vocabulary gain.

Vocabulary learning through the inference-gloss-gloss and gloss-retrieval-gloss conditions significantly differed from each other, and the inference-gloss-gloss condition had a higher mean than the gloss-retrieval-gloss condition. This may relate to the additional contextual clues accompanying inference situation in the inference-gloss-gloss condition, which gave the readers the opportunity to check their assumed meaning with the contextual clues.

Of the five interventions, the gloss-retrieval-gloss condition generated the least vocabulary learning in both form recall and meaning recall tests. It was significantly inferior to full glossing condition on the form recall test. Contrary to Huang and Lin's (2014) study where the participants had a better performance on the gloss-retrieval-gloss condition than on the full glossing condition, the participants in this study focused more on glossed words than unglossed ones and relied on the accessible input for vocabulary learning than activating the act of retrieval in the memory.

### Meaning Recall Test

Concerning the meaning recall test, of the five interventions, the gloss-retrieval-gloss-retrieval condition most benefited vocabulary learning, followed by inference-gloss-retrieval-

gloss, full glossing, inference-gloss-gloss, and gloss-retrieval-gloss conditions (Table 6).

The gloss-retrieval-gloss-retrieval condition was the most influential intervention for vocabulary meaning recall. This finding supports the claim that retrieval can fix items in the memory, improve the memory traces, and increase the accessibility of information in the memory (Nowzan & Baryaji, 2013). Moreover, it implies that retrieval can strengthen the link between form and meaning and hence enhance the recall of the words' meaning (Rott, 2007). The involvement load of retrieval is assumed to be three because it triggers the need to understand a word, search for its meaning in the mental lexicon, and evaluation of whether the searched meaning is contextually correct (Rott, 2007). In addition, the gloss-retrieval-gloss-retrieval condition's significant mean difference with the gloss-retrieval-gloss condition in the meaning recall test would be related to the additional word encounter's role in activating word processing in the mind.

The inference-gloss-retrieval-gloss condition performed as the second most beneficial glossing condition on the meaning recall test. It was not significantly different from the gloss-retrieval-gloss-retrieval condition (the best intervention condition on this test), but the existing difference would be related to the presentation of the target word's meaning in the first encounter in the gloss-retrieval-gloss condition, which would increase the reader's attention to the target word in following word encounters and thus result to more vocabulary gain (Rott, 2007). The inference-gloss-retrieval-gloss condition's statistically significant superiority over the gloss-retrieval-gloss condition could arise from more repeated opportunities for target word processing as suggested in Peters (2014).

The full glossing condition served as the third most beneficial gloss type based on meaning recall test scores; it outperformed inference-gloss-gloss and gloss-retrieval-gloss conditions in the vocabulary meaning recall test. As for the meaning recall test, the inference-gloss-gloss condition was also the fourth best intervention condition found in the study. The best glossing condition on meaning recall test was the gloss-retrieval-gloss-retrieval condition, but it was not statistically better than the inference-gloss-gloss condition. One possible reason may be due to the increased frequency of the target words in the gloss-retrieval-gloss-retrieval condition compared with the inference-gloss-gloss condition (Peters, 2014). On this meaning recall test, the gloss-retrieval-gloss-retrieval condition showed the highest vocabulary learning results, and it significantly outperformed the gloss-retrieval-gloss condition. According to Barcroft (2015), providing opportunities for target word retrieval can positively affect incidental L2 vocabulary learning, one additional recall chance in the gloss-retrieval-gloss-retrieval condition would be the reason of gloss-retrieval-gloss-retrieval group's outperformance.

## Conclusion

Few studies on L2 learning have examined textual glosses in combination with meaning inferencing and retrieval as glossing condition models. This study focused on this issue by comparing five glossing conditions in their effects on L2 vocabulary recall. The results of the study suggested that full glossing was the best intervention condition for vocabulary form recall, and it was more effective than the other four conditions in promoting L2 vocabulary learning. The gloss-retrieval-gloss-retrieval condition was the best intervention condition for vocabulary meaning recall, and it was more effective than inference-gloss-retrieval-gloss, full glossing, inference-gloss-gloss, and gloss-retrieval-gloss conditions in promoting recalling the target words' meaning. This study provides much-needed information on the effects of different glossing conditions on L2 vocabulary learning. By comparing five glossing conditions together, it sheds light on the use of different glossing conditions for L2 vocabulary learning.

The findings of the present study are subject to a number of limitations. The participants were MA students of TEFL, selected from a university where all instruction is done in English; accordingly, it is uncertain whether the findings can be readily applied to other students at a different educational level with different levels of language knowledge/proficiency. The difficulty of inferring each target word in the study might be different, so presenting these words in different contexts or conducting the research on different target words might generate different results. As only five target words were included in the passage, choosing a text with different lexical coverage may produce different results.

Future studies could be carried out to explore the effects of these glossing conditions on the non-English major students at the tertiary level. Moreover, this study only investigated words that occurred 3 or 4 times in the text; other studies could investigate words that occur less or more than 3 or 4 times to determine whether other glossing conditions can lead to more vocabulary learning.

## Appendix A

### *The Original Reading Passage*

**Thermoregulation.** Mammals and birds generally maintain body temperature within a narrow range (36 °C-38 °C for most birds) that is usually considerably warmer than the environment. Because heat always flows from a warm object to cooler surroundings, birds and mammals must counteract the constant heat loss. This maintenance of warm body temperature depends on several key adaptations. The most basic mechanism is the high metabolic rate of endothermy itself. Endotherms can produce large amounts of metabolic heat that replace the flow of heat to



the environment, and they can vary heat production to much changing rates of heat loss. Heat production is increased by such muscles activity as moving or shivering. In some mammals, certain hormones can cause mitochondria to increase their metabolic activity and produce heat instead of ATP (energy that drives certain reactions in cells). This nonshivering thermogenesis (NST) takes place throughout the body, but some mammals also have a tissue called brown fat in the neck and between the shoulders that is specialized for rapid heat production. Through shivering and NST, mammals and birds in cold environments can increase their metabolic heat production by as much as 5 to 10 times above the minimal levels those occur in warm conditions.

Another major thermoregulatory adaptation that evolved in mammals and birds is insulation (hair, feathers, and fat layers that prevent transfer of heat), which reduces the flow of heat and lowers the energy cost of keeping warm. Most land mammals and birds react to cold by raising their fur or feathers, thereby trapping a thicker layer of air. Humans rely more on a layer of fat just beneath the skin; goose bumps are a vestige of hair-raising left over from our furry ancestors. Vasodilation and vasoconstriction also regulate heat exchange and may contribute to regional temperature differences within the animal. For example, heat loss from a human is reduced when arms and legs cool to several degrees below the temperature of body core, where most vital organs are located.

Hair loses most of its insulation power when wet. Marine mammals such as whales and seals have a very thick layer of fat called blubber just under the skin. These mammals swim in water colder than their body core temperature, and many species spend at least part of the year in nearly freezing polar seas. The loss of heat to water occurs 50 to 100 times more rapidly than heat loss to air, and the skin temperature of a marine mammal is close to water temperature. Even so, the blubber insulation is so effective that marine mammals maintain body core temperatures of about 36 °C to 38 °C with metabolic rates about the same as those of land mammals of similar size. The flippers or tail of a whale or seal lack blubber, but countercurrent heat exchangers greatly reduce heat loss in these extremities, as they do in the legs of many birds.

Through metabolic heat production, insulation, and vascular adjustments, birds and mammals are capable of astonishing feats of thermoregulation. For example, small birds called chickadees, which weigh only 20 g, can remain active and hold body temperature nearly constant at 40 °C in environmental temperatures as low as -40 °C—as long as they have enough food to supply the large amount of energy necessary for heat production.

Many mammals and birds live in places where thermoregulation requires cooling off as well as warming. For

example, when a marine mammal moves into warm seas, as many whales do when they reproduce, excess metabolic heat is removed by vasodilation of numerous blood vessels in the outer layer of the skin. In hot climates or when vigorous exercise adds large amounts of metabolic heat in the body, many terrestrial mammals and birds may allow body temperature to rise by several degrees, which enhances heat loss by increasing the temperature gradient between the body and a warm environment.

Evaporative cooling often plays a key role in dissipating the body heat. If environmental temperature is above body temperature, animals gain heat from the environment as well as from metabolism, and evaporation is the only way to keep body temperature from rising rapidly. Panting is important in birds and many mammals. Some birds have a pouch richly supplied with blood vessels in the floor of the mouth; fluttering the pouch increases evaporation. Pigeons can use evaporative cooling to keep body temperature close to 40 °C in air temperature as high as 60 °C, as long as they have sufficient water. Many terrestrial mammals have sweat glands controlled by the nervous system. Other mechanisms that promote evaporative cooling include spreading saliva on body surfaces, an adaptation of some kangaroos and rodents for combating severe heat stress. Some bats use both saliva and urine to enhance evaporative cooling.

## Appendix B

### Inference-Gloss-Retrieval-Gloss Condition

#### Thermoregulation

Mammals and birds generally maintain body temperature within a narrow range (36-38°C for most birds) that is usually considerably warmer than the environment. Because heat always flows from a warm object to cooler surroundings, birds and mammals must counteract the constant heat loss. This maintenance of warm body temperature depends on several key adaptations. The most basic mechanism is the high metabolic rate of endothermy itself. **Endotherms** can produce large amounts of metabolic heat that replace the flow of heat to the environment, such warm-blooded animals can vary heat production to much changing rates of heat loss. Heat production is increased by such muscles activity as moving or shivering. In some mammals, certain hormones can cause mitochondria to increase their metabolic activity and produce heat instead of ATP (energy that drives certain reactions in cells). This nonshivering thermogenesis (NST) takes place throughout the body, but some mammals also have a tissue called brown fat in the neck and between the shoulders that is specialized for rapid heat production, this layer is called **blubber** in whales and seals. Through shivering and NST, mammals and birds in cold environments can increase their metabolic heat production by as much as 5 to 10 times above the minimal levels those occur in warm conditions.

page 1

Another major thermoregulatory adaptation that evolved in **endotherms** including mammals and birds is **insulation** (hair, feathers, and fat layers), which reduces the flow of heat and lowers the energy cost of keeping warm. Most **terrestrial** or land mammals and birds react to cold by raising their fur or feathers, thereby trapping a thicker layer of air. Humans rely more on a layer of fat just beneath the skin ; goose bumps are a vestige of hair-raising left over from our furry ancestors. **Vasodilation** also regulate heat exchange by widening the blood vessels and may contribute to regional temperature differences within the animal. For example, heat loss from a human is reduced when arms and legs cool to several degrees below the temperature of body core, where most vital organs are located.

page 2

Hair loses most of its **insulation** power when wet. Marine mammals have a very thick layer of fat called **blubber** just under the skin. These mammals swim in water colder than their body core temperature, and many species spend at least part of the year in nearly freezing polar seas.

page 3

The loss of heat to water occurs 50 to 100 times more rapidly than heat loss to air, and the skin temperature of a marine mammal is close to water temperature. Even so, the **blubber insulation** is so effective that they maintain body core temperatures of about 36-38°C with metabolic rates about the same as those of **terrestrial mammals** of similar size when they do **vasodilation**.

زمین، خشکی

گشاد شدن رگ ها

page 4

Through metabolic heat production, **insulation**, and vascular adjustments, birds and mammals are capable of astonishing feats of thermoregulation. For example, small birds called chickadees, which weight only 20 grams, can remain active and hold body temperature nearly constant at 40°C in environmental temperatures as low as -40°C, as long as they have enough food to supply the large amount of energy necessary for heat production.

حلق بندگی

Many **endotherms** including mammals and birds live in places where thermoregulation requires cooling off as well as warming. For example, when a marine mammal moves into warm seas, as many whales do when they reproduce, excess metabolic heat is removed by **vasodilation** of numerous blood vessels in the outer layer of the skin. In hot climates or when vigorous exercise adds large amounts of metabolic heat in the body, many **terrestrial mammals** and birds may allow body temperature to rise by several degrees, which enhances heat loss by increasing the temperature gradient between the body and a warm environment.

page 5

Like **vasodilation**, evaporative cooling often plays a key role in dissipating the body heat. If environmental temperature is above body temperature, animals gain heat from the environment as well as from metabolism, and evaporation is the only way to keep body temperature from rising rapidly. Panting is important in birds and many mammals. Some birds have a pouch richly supplied with blood vessels in the floor of the mouth; fluttering the pouch increases evaporation. Pigeons can use evaporative cooling to keep body temperature close to 40°C in air temperature as high as 60°C, as long as they have sufficient water. Many **terrestrial mammals** have sweat glands controlled by the nervous system. Other mechanisms that promote evaporative cooling include spreading saliva on body surfaces, an adaptation of some kangaroos and rodents for combating severe heat stress. Some bats use both saliva and urine to enhance evaporative cooling.

گشاد شدن رگ ها

زمین، خشکی

The flippers or tail of **endotherms** such as whales or seals that are marine mammals lack **blubber**, but countercurrent heat exchangers greatly reduce heat loss in these extremities, as they do in the legs of many birds.

جانوران خونگرم

چربی پستانداران دریایی

page 6

## Appendix C

### Form Recall Test

Please recall the English words that have the following Persian meanings in the reading and write them down:

1. عایق بندی
2. دریایی
3. جانوران خونگرم
4. لرزیدن
5. زمینی، خاکی
6. سازگاری
7. پستانداران
8. سوخت و ساز
9. چربی پستانداران دریایی
10. گشاد شدن رگ ها

## Appendix D

### Meaning Recall Test

Please translate the following words into Persian based on their meaning in the reading:

1. Terrestrial
2. Feather
3. Endotherms
4. Vestige
5. Blubber
6. Evaporation
7. Insulation
8. Flipper
9. Vasodilatation
10. Chickadee

### Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

### Funding

The author(s) received no financial support for the research and/or authorship of this article.

### References

- Barcroft, J. (2015). Can retrieval opportunities increase vocabulary learning during reading? *Foreign Language Annals*, 48, 236-249.
- Davis, J. (1989). Facilitating effects of marginal glosses on foreign language reading. *The Modern Language Journal*, 73, 41-48.
- Farvardin, M. T., & Biria, R. (2011). Textual glosses, text types, and reading comprehension. *Theory and Practice in Language Studies*, 1, 1404-1415.
- Goldstein, B. (2011). *Cognitive psychology: Connecting mind, research, and everyday experience* (3rd ed.). Belmont, CA: Wadsworth Cengage Learning.
- Goossens, N., Camp, G., Verkoeijen, P., & Tabbers, H. (2014). The effect of retrieval practice in primary school vocabulary learning. *Applied Cognitive Psychology*, 28, 135-142.
- Hashemian, M., & Fadaei, B. (2012). The effect of lexical glossing types on Persian L2 learners' reading and listening. *Procedia: Social and Behavioral Sciences*, 70, 490-500.
- Huang, L. L., & Lin, C. C. (2014). Three approaches to glossing and their effects on vocabulary learning. *System*, 44, 127-136.
- Hulstijn, J. H. (1992). Retention of inferred and given word meanings: Experiments in incidental vocabulary learning. In P. Arnaud & H. Bejoint (Eds.), *Vocabulary and applied linguistics* (pp. 113-125). London, England: Macmillan.
- Jacobs, G. M., Dufon, P., & Hong, F. C. (1994). L1 and L2 vocabulary glosses in L2 reading passages: Their effectiveness for increasing comprehension and vocabulary knowledge. *Journal of Research in Reading*, 17, 19-28.
- Ko, M. H. (2005). Glosses, comprehension, and strategy use. *Reading in a Foreign Language*, 17, 125-143.
- Ko, M. H. (2012). Glossing and second language vocabulary learning. *TESOL Quarterly*, 46, 56-79.
- Laufer, B., & Hulstijn, J. H. (2001). Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. *Applied Linguistics*, 22, 1-26.

- Nowzan, S. Z., & Baryaji, T. (2013). The impact of retrieval task on learning vocabulary of Iranian EFL learners. *Journal of Advances in English Language Teaching*, 1, 24-30.
- Peters, E. (2014). The effects of repetition and time of post-test administration on EFL learners' form recall of single words and collocations. *Language Teaching Research*, 18, 75-94.
- Richards, J. C., & Schmidt, R. (2010). *Longman dictionary of language teaching and applied linguistics* (4th ed.). London, England: Pearson.
- Rott, S. (2007). The effect of frequency of input-enhancements on word learning and text comprehension. *Language Learning*, 57, 165-199.
- Rott, S., Williams, J., & Cameron, R. (2002). The effect of multiple-choice L1 glosses and input-output cycles on lexical acquisition and retention. *Language Teaching Research*, 6, 183-222.
- Sharpe, P. J. (2010). *Barron's TOEFL iBT: Internet-based test*. New York, NY: Barron's Educational Series.
- 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.
- Zaid, M. A. (2009). A comparison of inferencing and meaning-guessing of new lexicon in context versus non-context vocabulary presentation. *The Reading Matrix*, 9(1), 56-66.
- Zamanian, M., & Heydari, P. (2012). Readability of texts: State of the art. *Theory and Practice in Language Studies*, 2(1), 43-53.

### Author Biographies

**Tayebeh Danesh** holds an MA in TEFL from the Khouzestan Science and Research Branch, Islamic Azad University, Ahvaz, Iran. She teaches at different English language institutes. Her research interests include second/foreign language vocabulary learning and reading in a second/foreign language.

**Mohammad Taghi Farvardin** is an assistant professor of TEFL at the Department of English Language Teaching, Islamic Azad University of Ahvaz, Iran. His areas of interest include psycholinguistics, second/foreign language vocabulary learning and teaching, and reading in a second/foreign language. He teaches at both undergraduate and postgraduate levels, and also supervises postgraduate students.