

Exploring the Impact of Interlinear vs. L1 Marginal Glosses on Iraqi EFL Learner's L2 Vocabulary Learning

Esmaeil Bagheridoust¹ & Hiba Adnan Jawad²

¹ Assistant Professor at Islamic Azad University, South Tehran Branch, Iran

² Islamic Azad University, Iran

Correspondence: Hiba Adnan Jawad, Islamic Azad University, Iran. E-mail: nononesreen7@gmail.com

Received: April 12, 2023

Accepted: August 1, 2023

Online Published: August 8, 2023

doi: 10.5539/elt.v16n9p13

URL: <https://doi.org/10.5539/elt.v16n9p13>

Abstract

This study investigates the impact of interlinear vs. L1 marginal glosses on the Iraqi Learning of vocabulary in EFL instructional setting. Along with all the efforts made during past 3 decades to improve the teaching of vocabulary via various strategies and techniques countywide, the classical grammar-translation method and its offspring are still widely used in most schools in Iraq. By analyzing the elicited data out of the management, treatment, and assessment of 92 pre-university students studying in two high schools in Iraq, the inquirer observed a reliable improvement of vocabulary learning in participants. To examine the effectiveness of the two glossing strategies, i.e., the implantation of L1 marginal and interlinear glosses, the results of pre and post-test of three groups were compared using One-Way ANOVA. We found reliable difference between the two groups in post-test vocabulary learning. Consequently, L1 marginal and Interlinear glosses both have come to make the learners understand L2 texts well, in comparison with the other group. Ultimately, Marginal glossing is more effective than the interlinear (text) glossing in the improvement of vocabulary development of Iraqi EFL learners.

Keywords: interlinear glosses, L1 marginal glosses, vocabulary learning, pre-university EFL learners

1. Introduction

1.1 The Problem

Although much effort has been made to improve the teaching of vocabulary by using different strategies and techniques as well as methods and approaches countywide, the traditional grammar-translation method and its related offspring are still widely used in many schools all over the country. For less proficient students, learning vocabulary can be an ordeal, often because of the great amount of unknown forms that makes it difficult or even impossible to get the main idea of the texts. As Paulston (1976) asserts "Reading is the most important skill for most students of English through the world" (P. 157).

Despite such specific need for reading ability, it is a common experience, at least of EFL teachers that most students fail to learn to read adequately in the foreign language because of not knowing how to deal with vocabulary. Very frequently, EFL students seem to read with less understanding that one might expect, all again due to lack of familiarity with sufficient vocabulary. In fact, vocabulary learning in a language which is not the learner's first language is a source of considerable difficulty. Since much of school learning especially the fourth grade of high school focuses on reading texts, it is important to find ways to develop EFL learner's stock of vocabulary, or at least detect influencing factors to increase the probability of students' success in reading comprehension. Researchers have always been investigating methods of facilitating vocabulary learning. Recently, however, scholars have begun to question the ability of schema theory alone to provide a complete understanding of comprehension processes (Alderson 2000). That is why more attention is paid to the role of vocabulary knowledge in reading comprehension and vocabulary teaching. (Lewis, 1993; Coady and Huckin, 1973; Alderson, 2000; Nation, 2001). For example, Alderson (2000) states that "measures of readers' vocabulary comprehension tend to be highly correlated with measures of reading comprehension and are often indeed the best single predictor of text comprehension."

One of the techniques that can be regarded as an effective way to increase stock of vocabulary in less amount of time is "glossing". Considering the mentioned problem in high school level, and the importance of vocabulary

knowledge, this study tries to focus on this technique and it is hoped that it can help high school students to improve their vocabulary learning and teachers to teach more efficiently.

1.2 Significance of the Study

Reading skill has been in the locus of attention in various research settings, hence for which multifarious strategies have been developed both for teaching and learning reading. The same story is true for the three areas of knowledge (Vocabulary, Pronunciation, and Grammar), Vocabulary cannot be fully learned or acquired unless there learners avail from suitable techniques and strategies that enables them to develop their stock of vocabulary more effectively and efficiently. This is very common among EFL learners to stop over each unknown vocabulary in reading passages and ask for help from the teacher or look it up in the dictionary. The students find vocabulary learning difficult and boring, because it demands that they utilize many cognitive and linguistic strategies of which they are uncertain. As it is reported by many teachers both in high schools and institutes, even most of them are often perplexed by different problems in their reading classes because of vocabulary deficiency. In this research, the researcher intends to find the impact of Interlinear vs. L1 Marginal Glosses on high school students' reading comprehension. In-text and marginal glosses help EFL learners to become independent and autonomous and enable them to read and comprehend more in a shorter time.

The first people who can benefit from this research will be the researcher's own students and further, all students at different educational institutions may enjoy this new way of having vocabulary classes. Moreover, English language instructors can save their time and energy and teach online while doing handling their other activities. This study is designed for one major purpose. It is designed to address the direct effect of the glosses on students' vocabulary learning and whether such a perception correlates with students' improved reading performance.

1.3 Review of the Related Literature

There is a growing consensus among L2 researchers and practitioners that providing marginal glosses facilitates this. L2 vocabulary learning (e.g. Abraham, 2008; Ko, 2012; Mohsen & Balakumar, 2011; Yanguas, 2009; Yoshii, 2006; Yoshii & Freats, 2002; Yun, 2011, etc.). Although it is claimed that L2 teachers can use real texts, without simplification, using marginal gloss makes it easier to read L2, Vocabulary Learning (Abraham, 2008). In addition, it is believed that it is possible to provide an edge gloss effect L2 learners are encouraged to read more authentic texts (O'Donnell, 2013). Look up unknown vocabulary in a dictionary (Nagata, 1999). Reading breaks make it easier to make precise connections between form and meaning (Huang & Lin, 2014) By providing accurate meanings for lexical items that L2 learners may not be able to guess the correct meaning of Meaning out of context (Nation, 2001). Similarly, lexical inference from context is not seen as an effective method for quickly learning a large number of her L2 words, as our research findings highlight. Number of variables involved in lexical reasoning. To date, L2 scholars and practitioners have investigated the effects of providing marginal glosses on L2 vocabulary learning from various perspectives, especially multimedia glosses (e.g., Al-Seghayer, 2001; Chun and Plass, 1996a, 1996b; Jones and Plass, Turk, and 2002). Ercetin, 2014; Yanguas, 2009; Yoshii and Flaitz, 2002), paper and computer glosses, i.e. e-glosses (e.g. Abraham, 2008; Abuseileek, 2008; Kim, 0 Gilman, 201, 1 Yun); - and multiple-choice glosses (e.g., Hulstijn, 1992), the role of individual differences in the use of multimedia glosses (e.g., Rouhi and Mohebbi, 2013; Yeh and Wang, 2003), and L1 and L2 marginal glosses. (e.g. Cheng and Good, 2009; Jacobs, Dufon, and Hong, 1994; Rouhi and Mohebbi, 2012; Taylor, 2006; Yoshii, 2006). In more recent studies, Hu, Vongpumivitch, Chang, and Liou (2014), Huang and Lin (2014), and Yoshii (2014) investigated the effectiveness of glosses in L2 word learning and reading comprehension from new perspectives, namely affect. L1 and L2 e-gloss on incidental word learning with upper and lower secondary learners of English, the effects of gloss matching inference or meaning retrieval on word learning, and the effects of checking glosses and glosses on L2 word learning through reading.

Yoshii (2006) compared the effects of L1 and L2 glosses on incidental word learning in a multimedia environment. Participants in this study were 195 EFL university students. The study was conducted under four conditions: L1 lightness, L2 lightness, L1 text lightness and image lightness and L1 brightness to textual brilliance and pictorial brilliance.

The study material was a story of 390 words, of which 20, were glossed, 14 target words and 6 known words. The glosses L1 and L2 were found to be effective for casual learning and useful for improving vocabulary learning, but in terms of long-term retention they were they suggested that L1 textual glosses led to better retention results compared to L2 textual glosses or L2 textual glosses and images. Lin and Huang (2008) compared the effect of inferred meaning and glosses of meaning on casual vocabulary learning by students. The study participants were 175 high and low achievers in Taiwan. The results confirmed the supportive effect of

glosses on vocabulary learning, in the sense that bold glosses promoted students' attention to new words. They also concluded that although both inferred meaning and meaning-given glosses can lead to accidental vocabulary learning in meaning-oriented reading, inferred meaning glossing turned out to be more effective in meaning-oriented reading and Dwyer (2008) compared the effects of 'vocabulary support' in different positions. Participants were 76 college students in the United States learning Spanish. They were asked to read a newspaper article in Spanish with vocabulary support either before reading, during reading, before reading and during reading and without help.

They concluded that students who received vocabulary help while reading performed better than those who received it before reading. Alessi and Dwyer offer three reasons to help with vocabulary reading. One is the empirical support for activities while reading as opposed to the theoretical support for reading aloud activities. The other reason is better adaptation to individual differences. And the third important practical advantage of the reading aid is "just in time". Regarding individual differences, they declare "fit", claim that different readers have different prior knowledge. As a result, the pre-reading activities may be useful for some readers and a waste of time for others. As for just in time, they argue that vocabulary support before reading can be forgotten over time and may not be available when really needed. But while reading vocabulary help is available at all times while reading. Al-Jabri (2009) examined the effects of L1 and L2 glosses on reading comprehension and memory of ideas. Nineteen non-native students were assigned to three groups of L1 gloss, L2 gloss, and no Gloss groups. The three groups were offered a 470-word text in English with 19 words annotated. Although the L1 group performed times better than the L2 group on the immediate multiple-choice reading comprehension test, there were no significant differences between the non-gloss group and of the non-gloss group -gloss groups. In terms of recall of ideas, the L2 gloss group was the least successful group in the retrieval protocol. As already mentioned, empirical L2 studies on the effects of marginal glosses on L2 vocabulary learning have confirmed the benefit of providing marginal glosses in improving L2 vocabulary learning. However, aspects of fringe glosses have been little studied. To our knowledge, to date, almost no L2 studies have been conducted to investigate the differential effects of giving and receiving marginal glosses on L2 vocabulary and reading comprehension. This study examines the effect of giving and receiving repetition of an L1 marginal raw and target lexical item. To promote L2 vocabulary in English as a foreign language (EFL). Also, as Fu, Vongpumivic, Chang, and Liou (2014) correctly point out, there are previous studies that have compared level one and two performances. Fringe results vary. This study is thus an attempt to fill a perceived gap in the research field. The next section of this article provides a brief overview of his recent L2 research investigating L1 effects. L2 marginal gloss improves L2 word learning and reading comprehension.

1.4 Research Questions and Hypotheses

As far as I am concerned, research on glossing has left a couple of questions unanswered. This study focuses on the following questions:

Do L1 marginal glosses have a significant effect on high school students' vocabulary learning?

Do interlinear glosses have a significant effect on high school students' vocabulary learning?

Is there any significant difference between the effectiveness of using the two strategies developing vocabulary learning in Iraqi EFL learners?

Here are the formulated research null-hypotheses:

H01: L1 marginal glosses have NO significant effect on high school students' vocabulary learning.

H02: Interlinear glosses have NO significant effect on high school students' vocabulary learning.

H03: There is NO significant difference between the effectiveness of using the two glossing strategies in developing vocabulary learning in Iraqi EFL learners.

2. Method

2.1 Participants

At the outset, the researcher started study by recruiting 120 participants, Iraqi EFL learners, pre-university students studying in two high schools in Iraq. They were all female with an English background of 5 years in junior high school. To screen and sift the participants on the grounds of their language proficiency for the purpose of homogeneity, the inquirer employed and administered a reliable and valid version of Nelson Language Proficiency test including grammar (30 items), vocabulary (10 items), and reading comprehension (10 items) in a testing session ran for to all the participants. Nelson Test reliability and validity were computed through a pilot study. The researcher did not select the students randomly or omit them from their classes, since

they were in the intact groups. She chose the students whose scores ranged between +1 and -1 standard deviation for her study and did not take into account other students who were not in that range. Since the number of the participants were low in each class, two classes were selected for each experimental group and control group. So, the study began with 6 classes. They were then randomly assigned to three groups, two experimental groups and one control group. The first experimental group consisted of two classes of vocabulary by using L1 marginal glosses. The second experimental group received interlinear glosses in the process of learning vocabulary. The control group class was taught using the conventional form of vocabulary learning and retention practiced normally in Iraqi English classes by their teachers. In fact, they faced non-glossed texts. As a result, 92 students were selected to partake in all phases of the study. It should be also mentioned that 20 students, in addition to the main participants were invited to take part in the pilot study for the purpose of calculating the measures of reliability and validity of Nelson test. Needless to say that this group of students were matching with the main participants of the study. Detailed process of homogeneity is explained later in the result section where data analysis is reported.

2.2 Instruments and Materials (Data Collection)

For the purpose of the study and in order to investigate the hypothesis, four instruments were utilized. There was proficiency test of Nelson, pre and post-test and high school vocabulary tests extracted out of the high school books. The instruments applied are explained below:

2.2.1 Nelson Test

The homogeneity of the participants was controlled through their scores in proficiency test of Nelson. It included 30 grammars, 10 vocabularies and 10 reading comprehension items. Participants are given 40 minutes to answer the test (or) The time allotted to the test is 40 minutes. Before administering the test to the students, the test was piloted with 20 students who were studying in fourth grade in another high school in Iraq. Getting the results, the researcher came to this conclusion that the test was appropriate for the fourth grade of high school since the test reliability and validity were computed by SPSS. The test reliability was 0.87.

2.2.2 Teacher-made Pre-test

As the goal of this study was to determine the effect of two kinds of glossing on reading comprehension, four reading passages followed by 22 multiple-choice questions were given to high school students in all three groups. The participants took this test one week before the treatment. The purpose of administering the pre-test was to assess the participants' reading comprehension ability, and also to ensure that there was no significant difference among members before the experiment. The passages were taken from a book provided for the fourth grade teachers of high school as a teaching guide and a supplementary book called ASAN. The researcher chose the texts based on syntactic complexity, text length, and content. Then, the test was piloted with a group in another high school and its reliability and validity were computed and necessary revisions were done by the researcher.

2.2.3 Teacher-made Post-test

One week after the treatment, the researcher administered a posttest. Its texts were chosen from high school fourth grade teachers' guide and ASAN supplementary book. Doing this the researcher wanted to capture the degree of improvement in the experimental and control groups and determine the difference between them after the treatment. The test had four passages followed by 22 multiple choice reading comprehension items. It should also be mentioned that the test was piloted before the administration to the main groups in order to approve of the appropriateness of the questions. The time allocation was 30 minutes.

2.2.4 Reading Comprehension Passages Including Vocabulary Glosses

The passages used for both experimental and control groups were selected from fourth grade of high school English text book, which their construction procedure will be explained below: To ensure which words were better to be glossed, the students were asked to underline the words that were unknown to them prior to the treatment. So the words that most of the students underlined were decided by the researcher and an experienced teacher who was supposed to teach the reading passages under the researchers' supervision and observation were selected to be glossed either in the text next to the glossed words in the parenthesis (L1 in-text glosses) for one of the experimental groups or in the margin of the reading texts (L1 marginal glosses) for the second experimental group. Additionally, the words which were glossed were bold faced to be noticed by students.

2.3 Design of the Study

In this research, the researcher intended to apply quasi-experimental design to investigate the effect of L1 marginal versus in-text glosses on reading comprehension ability of high school students. One of the subcategories of quasi-experimental design is “intact group design”. For this research, the researcher used this design which was the most appealing according to Seliger and Shohamy (1989) for foreign language experiments conducted in school environment because it requires the least amount of change in school routines. It does not require the reassignment of subjects to groups different from those in which they were already found. Also based on Hatch and Farhady (1982), in this kind of design both experimental and control group will receive a pre-test and post-test. Experimental group receives treatment while placebo is offered to control group. The presence of control group eliminates problems related to internal validity. Whereas since the students have not been randomly assigned to groups, we can’t generalize the results as much and don’t have external validity.

The schematic representation for the research design is illustrated below:

Table 1. Schematic representation of the design of the study

Groups	Pre-Test	Treatment	Post-Test
Experimental Group A	T1	X(marginal glosses)	T2
Experimental Group B	T1	X (in-Text glosses)	T2
Control group	T1	placebo	T2

A: Exposure of experimental group A to marginal glosses

B: Exposure of experimental group B to in-text glosses

T1: Pre-test, T2: post-test, X: treatment

In this study there were two independent variables, in-text and marginal glosses, and one dependent variable, namely, reading comprehension.

2.4 Procedure

In order to conduct the research and to verify the research hypothesis the following steps were taken by the researcher. Since the number of the students was low in their classes, two classes were chosen for each group. As a result, the study began with six classes. The number of the students in three groups, two experimental and one control group was the following: 20 and 18 students in L1 in-text gloss group, 20 and 20 students in L1 marginal gloss group and 20 and 22 in control group. Then a piloted Nelson language Proficiency test was administered to all 6 classes to find out the homogeneity of the groups in reading comprehension ability. After analyzing the data, the students whose scores ranged between +1 and -1 standard deviation were chosen to participate in this study. As it was explained in participant’s section, 92 students were within the appropriate range. The rest of the students whose scores were not at this modified range, were dropped from the study. The next step was to administer the pre-test. It was administered in order to obtain the subjects’ proficiency level of reading comprehension before the treatment and compare its results with the posttest to find out whether the independent variables -L1 marginal and in-text glosses- have any significant effect on the dependent variable -reading comprehension- or not. It was done with Word Creator software 2010 and the mean score was calculated. Therefore, the texts, which were chosen for comprehension test, had the readability between 51.5 and 71.5. In order to pilot the test, the researcher administered it to a group of 20 students. Next, to ensure the validity of the test, the groups answered PET questions. After that, the students reading scores were correlated with their scores in PET using Pearson Product Correlation Coefficient. One week after the pretest, the treatment began. The students in experimental groups received 10 sessions of treatment, 2 sessions for each reading text. They were taught the same way in both experimental groups; just the difference was in the way of text presentation. It is necessary to mention that the glossed texts were prepared by the researcher. In one of the groups the new words in the text were presented using L1 marginal glosses and in another one L1 in-text glosses, while control group faced with non- glossed texts in the same period of the time. The teaching procedure in the experimental groups is explained below:

The teacher entered the class. After greeting, she introduced the researcher, her sister. The researcher had explained the way of teaching to the teacher before coming to the class. It should be mentioned that each reading text in fourth grade of high school has a picture related to the text. Therefore, the teacher focused students’ attention on the picture as a warm up. She asked some questions to involve them in the warm up and introduced some of the words in this phase; she even used TPR to show the meaning of the words that would be seen later in

the reading texts. There were also 5 pre-reading questions that the students were supposed to answer. The teacher allotted a time of 5 minutes to these questions. The students were assigned to groups to do these exercises. The teacher observed what they were doing and helped them with the answers. In addition to the questions there were also some sentences that the students were to check whether they agree or disagree. The time allocation for this activity was 3 minutes. Also the students should have written two other sentences related to the reading text.

After doing pre-reading exercises, the teacher went through the reading texts. First, she herself read the first paragraph to model the pronunciation of the words. Then, she asked one of the students to read from the text to check pronunciation. Next, she instructed the students to read the text for comprehension in their previously assigned groups. She also asked the students to pay attention to the bold faced words and their meanings either in the margin or next to the words. She also observed them and answered their questions if any. After reading each paragraph, the students were asked to say whatever they comprehended in their first language. When the students were incomplete in presenting the meaning, the teacher asked other students to do that and as the last resort, she herself presented the complete meaning. They followed the same procedure for other paragraphs until the end of the text. Finishing working on the paragraphs, they started post-reading activities which were the following: True/ False, multiple choice and completion questions. Because the reading passages are long in fourth grade of high school, each reading passage took two sessions.

The teacher taught the same texts in control group too, but the students were faced with non-glossed text. They were assigned to groups and there was one head for each group who was stronger than others. They tried to get the meaning of the paragraphs by the head's help and the teacher observation. At the end, they went through post-reading activities like experimental groups' classes.

After working on reading passages through glossed words in experimental groups and non-glossed texts in control group, the researcher administered a posttest one week after the treatment. The test contained 4 reading passages followed by 22 multiple choice items. Further, the researcher found the readabilities of the texts through Fog index of readability. To ensure the validity of this test, the scores of pilot group were correlated with student's PET scores using Pearson Product Moment Correlation Coefficient. Finally, the researcher collected all scores including Proficiency test of Nelson, pretest and post-test to analyze. Chapter four presents more on data analysis and the results of the tests in all groups.

3. Results of Data Analysis

3.1 Statistical Results on Nelson Test

In order to make sure that participants did not differ significantly or in other words, they were homogeneous in the advent of the study, the researcher administered a reliable and valid Nelson test to 120 participants in all 6 classes in two high schools in Iraq. Those participants who scored within the range of one standard deviation above and below the mean were selected as homogeneous. Other participants whose scores were not in that range were discarded from the study. As a result, the researcher selected totally 92 pre-university participants for this study. They were assigned to three groups; two experimental and one control group. All the groups were equal regarding their teaching materials, time of teaching, number of class sessions, and pre and post-tests. The reliability of Nelson test was computed using Cronbach Alpha in the pilot study which was 0.76. Since this number was close to one, the researcher concluded that the test was reliable.

Table 2. Nelson Test Reliability Index by Cronbach Alpha

Reliability Coefficients	
N of Cases = 120.0	N of Items = 50
Alpha = .7677	

Table 3. Case Processing of the Nelson Proficiency Test

Case Processing Summary						
	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
COUNTER	120	100.0%	0	.0%	120	100.0%

Table 4. Descriptive Statistics of the Nelson Proficiency Test

		Statistic	Std. Error
Mean		29.167	.63799
95% confidence	Lower bound	27.8534	
Interval for mean	Upper bound	30.3799	
5% trimmed mean		29.3704	
Median		30.0000	
Variance		48.843	
St. Deviation		6.9881	
Minimum		10.843	
Maximum		45.00	
Range		35.00	
Interquartile range		7.0000	
Skewness		-.687	.221
Kurtosis		-.603	.438

Table 4 shows that the mean of the Nelson test was 29.167 and its standard deviation 6.98. The following histogram shows the Nelson Proficiency Test with the normal curve imposed on it.

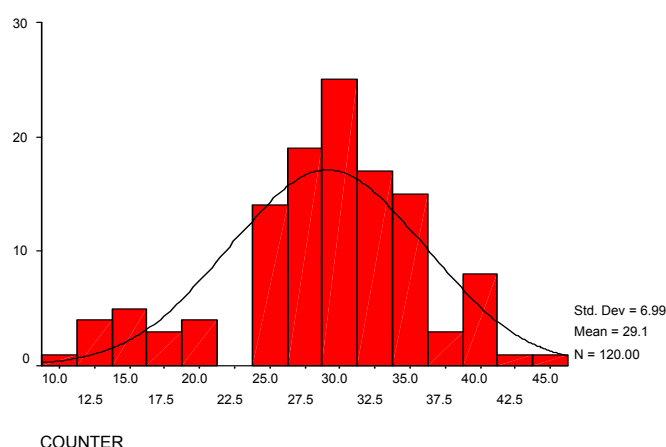


Figure 1. Histogram of the Nelson Proficiency Test

3.2 Readability and Text Selection for Pretest and Post Test

Mousavi (1991) defines readability as "a measure of understandability of written text as giving by an analysis of a variety of factors including syntax complexity, vocabulary, thematic expression and continuity of themes" (p.310). Since the pretest and posttest were developed by the researcher and she, herself, chose the reading texts for the tests the readability of the participants' course book was considered as the basis of choosing the texts according to the level of participants' reading proficiency. To do so, the researcher assessed the course book texts' readabilities using Fog index of readability and computed the mean. She selected the pre and post-test's texts within -1 and +1 standard deviation of the course book text's readability. Table 5 above shows the participants' English course book texts' readabilities.

Table 5. Readability indexes of participants' English course book

Text	Exercise	Speech	Global Warming	Earthquake	Child Labor
Readability	19.9	21.75	16.87	15.67	23.01

Table 6. Descriptive Statistics of participants' English course book

Mean	SD	Min	Max
19.40	1.56	15.67	23.01

As the results show, the average readability score of the participants' textbook turned out to be 19.40 and the standard deviation was 1.56, so pre and post-test texts' difficulty in this experiment was considered to be 17.84 ≤ text difficulty ≤ 20.96. Table 7, 8 and 9, 10 show the readabilities of pre and post-test respectively.

Table 7. Readability of selected texts for pre-test

Text	1	2	3	4
Readability	23.63	21.6	15.8	17.76

Table 8. Descriptive Statistics of selected texts for Pre-test

Mean	SD	Min	Max
19.69	2.05	15.8	23.63

Table 9. Readability of selected texts for post-test

Text	1	2	3	4
Readability	15.21	24.7	19.04	20.48

Table 10. Descriptive Statistics of selected texts for Post-test

Mean	SD	Min	Max
19.85	2.26	15.21	24.7

3.3 Reliability Indices (Pilot Study)

The third statistical procedure was to estimate the reliability of vocabulary learning tests (pre and post- test) administered to the participants. As it is stated by Farhady et al., (2003): "on a reliable test, one's score on its various administrations would not differ greatly" (p. 125). In fact, reliability can be defined as the extent to which a test produces consistent results when administered under similar conditions. Therefore, in order to compute the reliability of both pre and post-test, they were piloted with 20 students in another high school.

Table 11. Statistical analysis of pretest in the pilot study

	Scale	Scale	Corrected	
	Mean	Variance	Item-	Alpha
	if Item	if Item	Total	if Item
	Deleted	Deleted	Correlation	Deleted
Q1	17.6000	16.3579	.5858	.8618
Q2	17.5500	17.8395	.1883	.8757
Q3	17.5000	17.2105	.4329	.8673
Q4	17.5500	16.9974	.4417	.8671
Q5	17.5000	16.1579	.8042	.8554
Q6	17.5500	17.6289	.2505	.8737
Q7	17.5000	17.2105	.4329	.8673
Q8	17.5500	16.9974	.4417	.8671
Q9	17.4500	16.9974	.6180	.8626
Q10	17.6500	17.2921	.2934	.8734
Q11	17.5000	18.1579	.1180	.8768
Q12	17.5500	15.5237	.9179	.8498
Q13	17.5000	17.2105	.4329	.8673
Q14	17.4000	17.8316	.4125	.8687
Q15	17.4500	16.9974	.6180	.8626
Q16	17.5000	17.3158	.3970	.8684
Q17	17.5500	17.3132	.3452	.8705
Q18	17.5000	16.7895	.5785	.8627
Q19	17.4500	16.9974	.6180	.8626
Q20	17.4500	17.2079	.5318	.8649
Q21	17.5000	17.1053	.4689	.8662
Q22	17.6000	17.2000	.3428	.8710
Reliability Coefficients				
N of Cases = 20.0			N of Items = 22	
Alpha = .8720				

They were also studying in fourth grade of high school and were female too. Tables 11 and 12 show the statistical analysis of the reliability of the pre and post-test respectively.

Table 12. Statistical analysis of posttest in the pilot study

	Scale	Scale	Corrected	
	Mean	Variance	Item-	Alpha
	if Item	if Item	Total	if Item
	Deleted	Deleted	Correlation	Deleted
Q1	17.7000	18.1158	.3747	.8840
Q2	17.6500	17.9237	.5372	.8800
Q3	17.7000	17.9053	.4448	.8821
Q4	17.7500	17.7763	.4258	.8829
Q5	17.7000	16.8526	.8084	.8716
Q6	17.7500	18.3026	.2698	.8877
Q7	17.7000	17.9053	.4448	.8821
Q8	17.7500	17.7763	.4258	.8829
Q9	17.6500	17.7132	.6216	.8779
Q10	17.8500	18.0289	.2926	.8883
Q11	17.7500	17.9868	.3629	.8849
Q12	17.7500	16.1974	.9241	.8666
Q13	17.7000	17.9053	.4448	.8821
Q14	17.6000	18.5684	.4151	.8833
Q15	17.6500	17.7132	.6216	.8779
Q16	17.7000	18.0105	.4096	.8831
Q17	17.7500	17.9868	.3629	.8849
Q18	17.7000	17.4842	.5875	.8781
Q19	17.6500	17.7132	.6216	.8779
Q20	17.6500	17.9237	.5372	.8800
Q21	17.7000	17.8000	.4801	.8811
Q22	17.7500	17.7763	.4258	.8829
Reliability Coefficients				
N of Cases = 20.0		N of Items = 22		
Alpha = .8859				

3.4 Pretest of Vocabulary Learning

Since the researcher was very cautious to see that this homogeneity was not just due to some specific aspects of the administered test, she gave the three groups a teacher-made pretest. The descriptive statistics of pretest for all three groups is presented below:

Table 13. Descriptive statistics of pretest of vocabulary learning for interlinear gloss group

		Statistic	Std. Error
Mean		13.4138	.46162
95% confidence	Lower bound	12.4682	
Interval for mean	Upper bound	14.3594	
5% trimmed mean		13.4981	
Median		13.0000	
Variance		6.180	
St. Deviation		2.48592	
Minimum		8.00	
Maximum		17.00	
Range		9.00	
Interquarile range		3.5000	
Skewness		-.347	.434
Kurtosis		-.679	.845

As the above tables show the mean of two experimental groups - interlinear and marginal-and the control group are 13.41, 13.36 and 12.90 respectively. In order to check the amount of significance of difference, the researcher compared the means using One-Way ANOVA. As the Table 14 shows the F value equals 0.673. Since the F value is bigger than 0.5 level of significance and the means differ slightly, participants in three groups were approximately of the same proficiency in vocabulary learning. This means that there was no significant difference between the scores of the two groups in pretest at the beginning of the research and their homogeneity is proved.

Table 14. One-Way ANOVA Pretest of Vocabulary learning by groups (Pretest)

	Sum of Square	df	Mean Squares	F	Sig
Between Groups	4.924	2	2.462	.398	.673
Within Groups	550.728	89	6.188		
Total	555.652	91			

3.5 Posttest of Vocabulary Learning

The researcher took the following steps after the treatment was over. She administered a piloted posttest. It consisted of four reading passages followed by 22 multiple choice questions. In order to examine the effectiveness of the two strategies used in this study -the implantation of L1 marginal and interlinear glosses- and probe the first and second research questions the results of pre and post-test were compared using One-Way ANOVA. As the Table 15 shows the F observed value is 0.000 which is $0.00 \leq 0.05$, so there is a significant difference between groups. So, the first and second hypothesis have been rejected. This means that the new treatment has positive effect and leads to the improvement of vocabulary learning ability in the experimental groups.

Table 15. One-Way ANOVA Post-test Vocabulary learning by Groups

	Sum of Square	df	Mean Squares	F	Sig
Between Groups	144.737	2	72.369	71.433	.000
Within Groups	90.165	89	1.013		
Total	234.902	91			

In order to examine the third hypothesis that is to check which strategy was more effective Scheffe Post-test was used. We use this test to determine exactly which group or groups differ with others. The results show that L1 marginal gloss group's mean was 4.66, L1 interlinear gloss group 2.93, and finally the mean of control group was 1.63. This means that marginal gloss group surpassed interlinear gloss group and both surpassed control group. Therefore, the third hypothesis was also rejected. The following Table shows the descriptive analysis of the

Scheffe:

Table 16. Descriptive statistics of posttest of Scheffe

Subset for alpha			
N	1	2	3
Groups			
Control	33	1.6364	
Interlinear	29	2.9310	
Marginal	30	4.6667	
Sig	1.000	1.000	1.000
Means for groups in homogeneous subsets are displayed			
Uses harmonic mean sample size = 30.575.			
The group sizes are unequal. The harmonic mean of			
The group sizes are used. Type 1 error level are not guaranteed.			

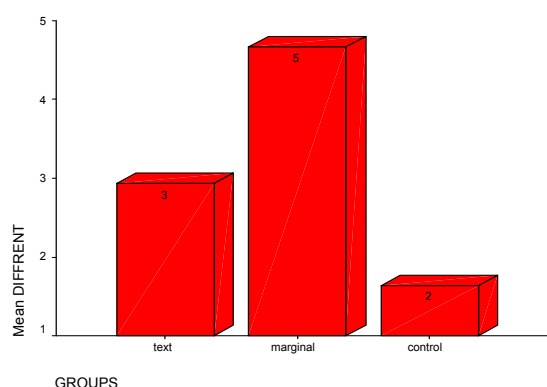


Figure 2. The Bar Graph of the groups comparison after the post test

4. Discussion

In this study, the researcher tried to show the impact of glossing strategy on Iraqi EFL high school students' vocabulary learning. To do so, she administered a Nelson Proficiency Test to determine the homogeneity of the participants. Ninety-two homogenized participants were assigned to two experimental groups and one control group. In order to measure their initial vocabulary learning, all the participants in the control group and two experimental groups took a teacher-made vocabulary learning test as a pre-test. The test had been piloted with same level English learners in another high school to determine its reliability. After conducting the pre-test, the data were processed and afterward One-Way ANOVA was run which revealed that there was no significant difference in the reading ability of all three groups before the treatment. After 10 sessions of treatment, teaching experimental groups using L1 marginal gloss in one group and L1 interlinear gloss in another one in reading sessions, a post-test of vocabulary learning was administered. This was done to clarify the impact of the reading strategy instruction via L1 marginal and interlinear glosses on the vocabulary learning of Iraqi pre-university students. The first and second questions, on the effectiveness of L1 marginal and interlinear glosses on vocabulary learning, received positive answers. Learners in EXP class demonstrated a better understanding of the texts than learners in NO gloss CON group. As for the comparison of both EXP and CON Groups, Post test results indicate reliable differences. No doubt that L1 marginal and interlinear glosses appeared to have helped learners in EXP class perceive L2 texts better, as compared with NO gloss treatment lass (CON Group).

In contrast, the participants in the no gloss group demonstrated less success in vocabulary learning, as they were generally not able to develop and grow a good stock of vocabulary. If meaning is at learner's reach, they are exposed to a shaded version of the second language text were able to decode and understand sentences better than learners exposed to a shaded version of the same text. For those participants, decoding a text at the sentence level without gloss proved to be particularly difficult. In addition, empirical studies of glosses conducted to improve understanding of the reading process from a bottom-up perspective support the theoretical proposition

that glosses are effective vocabulary learning aids. (e.g. David, 1989; Jacobs, Dufon, and Hong, 1994; Roby, 1991, 1999; Hulstijn, 1993;), therefore, this study supported these claims.

5. Conclusions and Findings

We found reliable positive effect from the party of L1 marginal glosses on vocabulary learning process. After receiving the treatment and comparing the results of performance of each group on the pre-test with their performance on the post-test, it was revealed that there is a significant increase in the performance of participants in the experimental groups; that means that the participants in experimental groups benefitted significantly from the treatment, L1 marginal glosses and interlinear glosses, conducted. In reality, this research has created a bond between the L2 learners' deficiency of vocabulary understanding and L2 passages (Birch, 2002). Put is another way, L2 reading in a bottom-up way, glosses are typically regarded as operative strategies that make vocabulary more accessible to L2 learners. L1 marginal glosses in the hand of the learners in EXP group creates a gain over their CON cohorts. For these participants learning vocabulary at the sentence level without glosses proved difficult.

To conclude, the results of this study confirmed some researchers' studies on the effectiveness of glossing on reading comprehension ability and vocabulary learning. His study reported a detrimental effect for using glosses; since the learners focus on linguistic aspect of the text and approach the text from a bottom-up view that prevent them from guessing or using their background knowledge. Considering the language of glosses, this study confirmed Aljabri's (2009) study in which the participants recalled more ideas from reading and vocabulary learning, using L1 glosses rather than L2 ones.

Due to the fact that, high school students and their teachers speak their first language in their English classes, and when working on vocabulary learning, they try to recall the vocabulary ideas and their propositions in their first language, hence, using L1 glosses seems more effective than L2 ones. Based on what we can read and understand from the ANOVA table reflected in chapter 4 and the results obtained, the researcher has come up with the following findings: The use of L1 marginal glosses definitely DOES play a significantly positive role in promoting the vocabulary learning skills of Iraqi EFL high school students. The use of interlinear glosses definitely DOES play a significantly positive role in promoting the vocabulary learning skills of Iraqi EFL high school students.

References

- Alderson, J. C., & Urquhart, A. H. (Eds.). (1984). Introduction: What is reading? *Reading in foreign language*. (pp. 14-17). <https://doi.org/10.4236/ce.2012.33055>
- Alderson, J. C. (2000). *Assessing Reading*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511732935>
- Al-Jabri, S. (2009). The effect of L1 and L2 glosses on reading comprehension and recalling ideas by Saudi students. *Umm Al-Qura University Journal of social Sciences*, 1(1), 11-27. <https://doi.org/10.47750/jett.2022.13.02.028>
- Al-seghayer, K. (2001). The effect of multimedia annotation modes on L2 vocabulary acquisition: A comparative study. *Language Learning & Technology*, 5(1), 202-232. Retrieved July 10, 2005, from <http://lilt.msu.edu/vol5num1/alseghayer/default.html>
- Bell, F., & LeBlanc, L. (2000). The language of glosses in L2 reading on computer: Learners' preferences. *Hispania*, 83, 274-285. <https://doi.org/10.2307/346199>
- Birch, B. (2002). *English L2 reading, getting to the bottom*. Mahvash, NJ: Lawrence Erlbaum.
- Bowels, M. A. (2004). L2 glossing: To CALL or no to CALL. *Hispania*, 87(3), Interactive reading in French. In F. B-& E. M. T. Johnson (Ed). <https://doi.org/10.2307/20063060>
- Cheng, Y., & Good, R. L. (2009). L1 glosses: Effects on EFL learners' reading comprehension and vocabulary retention. *Reading in a Foreign Language*, 21(2), 119-142. From <http://nflrc.hawaii.edu/rfl>. <https://doi.org/10.14483/calj.v19n1.10032>
- Chun, D. M., & Plass, J. L. (1996). Effects of multimedia annotations on vocabulary acquisition. *Modern Language Journal*, 80(2), 183-198. <https://doi.org/10.1111/j.1540-4781.1996.tb01159.x>
- Coady, J. (1979). A Psycholinguistic Model of ESL Reader. In R. Mackay, Berkman, & R. R. Jordan (Eds.) *reading a second language*, 5-12. Rowly, Mass: Newbury House. <https://doi.org/10.4236/jss.2016.46018>
- Farhady, H., Birjandi, P., & Jafarpour, A. (2003). *Testing Language skills: From theory to practice*. Tehran:

SAMT.

- Hulstijn, J. H. (1992). Retention of inferred and given word meanings: Experiments in incidental vocabulary learning. In P. J. L. A. a. H. Bejoint (Eds.), *Vocabulary and applied linguistics* (pp. 113-125). London: McMillan. <https://doi.org/10.4304/jltr.1.2.117-120>
- Hustijn, J. H. (1993). When do foreign language readers look up the meaning of unfamiliar words? The influence of task and learner variables. *The Modern Language Journal*, 77(2), 139-147. <https://doi.org/10.4304/jltr.1.2.117-120>
- Jacobs, G. (1994). What lurks in the margin: Use of vocabulary glosses as a strategy in second language reading? *Issues in Applied Linguistics*, 5(1), 115-137. <https://doi.org/10.5070/L451005174>
- Jacobs, G., Dufon, P., & Hong, F. C. (1994). L1 and L2 glosses in reading passages, their effectiveness for increasing comprehension and vocabulary knowledge. *Journal of Research in Reading*, 17(1), 19-28. <https://doi.org/10.1111/j.1467-9817.1994.tb00049.x>
- Johnson, P. L. (1982). Effects on reading comprehension of building background knowledge. *TESOL QUARTERLY*, 16(14), 503-516. <https://doi.org/10.2307/3586468>
- Lewis, M. (1993). *The lexical Approach*. Hove, England: *Language teaching linguistics*. Rowly, Mass: Newbury House. Pp. ix + 200. ISBN 0-906717-0-99-X
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511667190>
- Nation, P. (1995). Best practice in vocabulary and learning. *EA Journal*, 13(2), 7-15. In Richards & Renandya, *Methodology in language teaching: An anthology of current practice*. London: Cambridge University Press. <https://doi.org/10.1017/CBO9780511667190>
- Otto, W., & White, S. (1982). Editor's epilog: Look to the interaction. In W. Roby, *What's in a gloss* (Ed.). Japan: Dokkyo University. Available at <http://ilt.msu.edu/vlo2num2/robby/index.html>. <http://ilt.msu.edu/vol2num2/robby/>
- Roby, W. B. (1999). What is in a gloss? A commentary on Lara L. Lomicka's "To gloss or not to gloss": An investigation of reading comprehension online. *Language Learning & Technology*, 1(2), 94-101.
- Seliger, H.W., & Shohamy, E. (1989). *Second language research methods*. Oxford: Oxford University Press. <https://doi.org/10.4236/ce.2017.815167>
- Yanguas, I. (2009). Multimedia glosses and their effect on L2 text comprehension and vocabulary learning. *Language Learning & Technology*, 13(2), 48-76. <https://doi.org/10.1017/S095834401100005X>
- Yoshii, M. (2006). L1 and L2 glosses: Their effects on incidental vocabulary learning. *Language learning & technology*, 10(3), 85-101. <https://doi.org/10.1016/j.lindif.2015.11.018>

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).