The Differential Effects of Glossing for Different Parts of Speech on Vocabulary Retention and Reading Comprehension

Shahnaz Zolfagharjooya\textsuperscript{1}  
Manoochehr Jafarigohar\textsuperscript{2}  
Hassan Soleimani\textsuperscript{3}

\textsuperscript{1,2,3} Department of Applied Linguistics, Payame Noor University, Iran

\textbf{Abstract:} The present study considered nouns, verbs, adjectives and adverbs as independent variables, to investigate the combined and simultaneous effects of glossing and word class on retention and comprehension. The sample of the study consisted of 135 participants who were non-English major undergraduates at a university in Iran. After employing a proficiency test, they were selected and assigned randomly to one of four groups: In the first treatment group nouns were glossed, in the second one verbs, for the third group adjectives and for the fourth group adverbs were glossed. After exposure to the L2 reading texts, comprehension and retention were measured through the corresponding multiple choice items. The vocabulary test was repeated four weeks later. Data of the study were collected over a period of 7 weeks. The findings suggested that there were not any significant differences between the mean scores of the four groups on the reading comprehension test but there were differences between the mean scores of the four groups on the immediate and delayed recall tests.

\textbf{Keywords:} Gloss, Glossing, Reading comprehension, Vocabulary Retention

1. Introduction

Reading comprehension skill goes far beyond one single skill in which the readers interact with written materials for various purposes. The extent to which a reader comprehends and learns from L2 texts is a highly complex process depends on a number of factors. Two general factors are (1) those which are internal to readers such as the background knowledge of the readers, and (2) those factors which are related to L2 texts such as structure of sentences, content and organization of text, genre, type of text, lexis, etc. (Alderson, 2000). Of these, the necessity of increasing level of vocabulary knowledge appears to be the most significant one in comprehending a text especially among non-English major students. In such situations, comprehending a reading text elicits moans and groans from students as they spend too much time and effort looking up words’ meanings in the dictionary. Since glosses are assumed to have a vital role in facilitating reading comprehension by proving a translation or brief explanation of difficult or technical words (Segler, 2001), the present study investigated the differential effect of glossing for different parts of speech on retention and comprehension.
2. Literature review

2.1 Glossing and vocabulary retention

To examine the impacts of glossing on incidental vocabulary learning Hulstijn et al. (1996) administered an experiment in which 78 advanced level Dutch learners of French were instructed to read an adapted text (1306 words) under one of three conditions: (1) L1 marginal glosses, (2) bilingual dictionary use, and (3) control condition (neither glosses nor use of dictionary). Related data was collected after examining 16 target words. The results showed that L1 marginal glosses were more effective than bilingual dictionary use because readers relay used the dictionary during their reading.

Gettys et al. (2001) investigated the effects of sentence-level L1 translation and dictionary form glosses on incidental vocabulary learning among second-year college students learning Russian. The results of the study indicated that the dictionary form glosses were more efficient than sentence-level translation glosses and the participants in this group spent more time on the task than the participants who were exposed to the exact translation task.

Chen (2002) conducted a study with 85 college freshmen who were learning English as a second language in Taiwan. The participants in this study were randomly assigned to one of three groups: L1 gloss (Chinese), L2 gloss (English), and No gloss. They had to read an English text containing 193 words with 20 glossed words. Results of the study did not indicate a statistically significant difference between L1 and L2 glosses but the L2 gloss group outperformed the no gloss group.

In another experiment, Lin and Huang (2008) conducted a study with 175 high and low proficiency level English learners from four intact classes in a senior high school in Taiwan. The study examined the effect of meaning-inferred (MI) and meaning-given (MG) glosses on learners' incidental vocabulary learning. As treatments the students of the MI group were given three alternatives of each target word in L1 while those of the MG group were given single L1 translation of each target word. The results revealed that meaning-inferred glosses had a better effect on vocabulary learning and retention than meaning-given glosses but both glosses, were effective in incidental vocabulary learning.

2.2 Glossing and reading comprehension

Some experiments indicated that vocabulary glosses could have positive effects on reading comprehension and some other studies predicted that those lexical glosses could have opposite impacts on comprehension. In the current paper some of the experiments reviewed briefly.
Johnson (1982) examined the effects of vocabulary glosses, prior cultural experience and pre-reading vocabulary instruction on reading comprehension among 72 advanced ESL university students. Comprehension was assessed by a written recall task, a content-based recognition test, and a delayed cloze test. Results indicated that prior cultural experience and pre-reading vocabulary instruction facilitated reading comprehension, but exposure to the different types of vocabulary conditions did not significantly affect the comprehension. Findings of this experiment showed that the top-down processing (attention to the prior knowledge during reading) was more efficient than the bottom-up processing (the availability of vocabulary glosses during reading).

Ko (2005) investigated the effectiveness of the glosses on reading comprehension. 106 Korean undergraduate students read an English text under three conditions: Korean gloss (L1 glossing), English gloss (L2 glossing), and no gloss. After reading the text, they had to take a multiple-choice reading test and respond to a questionnaire. The results indicated that both types of glosses facilitated reading comprehension but the second language (L2) gloss condition significantly affected the students’ reading comprehension and no significant difference was shown between L1 gloss and no-gloss groups. Furthermore, 62% of participants favored L2 glosses for their reading material.

In another study, Al-Jabri (2009) conducted an experiment to examine the impacts of L1 (Arabic) gloss, L2 (English) gloss, and no gloss conditions on comprehension and ideas recall. 90 undergraduate students were given a 470 word English text with 19 glossed words to read. The result of the research indicated that L1 gloss group performed significantly better than the L2 gloss group in text comprehension and no significant difference was seen between no-gloss and L1 gloss and L2 gloss groups. Results of recall protocol revealed that students with L1 glosses and those with no gloss recalled significantly more ideas than learners exposed to L2 glosses. Moreover, more than (94%) of participants expressed their preference to use glosses and (50%) favored L2 glosses for their reading material.

3. Significance of the study and Statement of the problem

Considering the enhanced text literature, the present study is significant because this experiment isolated vocabulary items according to their grammatical roles. Furthermore, most of the previous studies in Iran have examined the effects of glossing among English major students so in order to shed more light on the impacts of glossing on retention and comprehension, the present study aimed to answer the following three research questions:
1. Is there any significant difference in reading comprehension of the four groups exposed to glossed parts of speech?
2. Is there any significant difference in the immediate recall of glossed parts of speech for the four experimental groups?
3. Is there any significant difference in the delayed recall of glossed parts of speech for the four experimental groups?

4. Method

4.1 Participants

To assess the extent to which the use of L2 glossing as decoding aids promotes comprehension and retention, 135 first and second semester students, both males and females, were selected after performing a proficiency test. They were majoring in computer or mechanical or civil engineering. They had studied English for about 7 years at school and their ages ranged from 18 to 23. Thirty five of the students participated only in the vocabulary pretest. The sample of the study did not include students enrolled in other language courses. Participants were informed that their scores would not affect their course grades.

4.2 Materials

4.2.1 Placement Test

The research involved collecting data from intermediate students. To achieve this purpose Oxford proficiency test (Oxford Solution Proficiency Test by Linda Edwards, Oxford University Press 2007) was conducted. The participants (N=135) were selected according to the test results so they were at intermediate level and homogenous in terms of their general English language proficiency.

4.2.2 Reading passages

Reading texts were selected from three web sites. Four passages corresponded to four lexical categories were used. In one of them new vocabulary items were mostly nouns, in the other new lexical items were mostly verbs and in the third and fourth ones they were mostly adjectives and adverbs. To meet the study requirements the passages and the following comprehension items reviewed and manipulated carefully by two experts in the field and the researcher. The manipulations were as follows:

With respect to inferability effect, contextual semantic and pragmatic cues of the passages were taken into account to eliminate the possibility of guessing the TWS from context. With respect to readability, some words were replaced with their simple and most near synonyms based on two criteria: first, words appeared to be difficult for the participants to know or guess. Second, they were belonged to a category other than the targeted one. The texts were also summarized to some degree in a way that the gist, key points, the important details and the TWs remained. As a result of replacement and summarization readability indexes were more close to each other and more appropriate for intermediate level. Flesch-Kincaid Grade level for verbs, nouns, adjectives and
adverbs were 9.27, 9.23, 9.21, 9.28. Furthermore, The TWs occurred only one time in the corresponding passages and the texts were no longer than 380 words. Generally the passages were about pollution, art, education and driving laws. Each reading assignment included four multiple-choice items aimed at assessing comprehension. Overall score for each reading task ranged between 0 and 4.

4.2.3 Vocabulary pre-test

A pilot test was conducted a week prior to the treatment. 35 students (out of a total of 135) were randomly selected to provide us with a representative performance picture. A list of 41 lexical items, including ten adverbs, eleven verbs, ten nouns and ten adjectives, was prepared. The students in this pilot test were instructed to explain any possible meanings, in Persian or in English, of the lexical items they knew and to mark only the words they found difficult or did not know. After examining the results, two lexical items, one adjective and one verb, were excluded, so 39 vocabulary items (10 adverbs, 10 nouns, 10 verbs and 9 adjectives) were remained for glossing.

4.3 Procedure

In the current research the participants were selected after administrating the proficiency test and the TWs were chosen after employing the pretest. The research design involved 4 posttests which were performed over a period of 4 weeks and one week was allocated to the delayed posttests so totally the study was conducted over a period of 7 weeks. The target words for the experimental groups were bold-faced and underlined and the glosses appeared on the right side of the page. Each gloss contained a synonym of 1 word, or a short definition. A total of 100 students participated in the immediate lexical retention and reading comprehension tests. They were randomly assigned to four treatment groups (25 students in each group). Reading comprehension and vocabulary test items were in the multiple-choice format with one stem and four choices. The participants received the texts in traditional paper-based hard copies. Students in each experimental group were instructed to read the passage and after collecting the passages they were asked to answer the immediate comprehension and retention items. Vocabulary posttests were repeated four weeks later. Quite surprisingly, no participants were absent in the delayed posttests. To insure incidental learning the participants did not know that there would be an immediate or a delayed vocabulary test.

5. Results and discussion

5.1 Research Question 1
Is there any significant difference in reading comprehension of the four groups exposed to the glossed parts of speech?

A one-way ANOVA was run. Based on the results displayed in Table 1 (F (3, 96) = .987, P = .402 > .05, $\omega^2 = .0001$ it represent a weak effect size) it can be concluded that there are not any significant differences between the mean scores of the four groups on the reading comprehension test. Thus the first null-hypothesis as there is not any significant difference in reading comprehension of the four groups exposed to the glossed parts of speech is supported.

Table 1: One-Way ANOVA Reading Comprehension by Groups

<table>
<thead>
<tr>
<th></th>
<th>Sum of Squares</th>
<th>Df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>.590</td>
<td>3</td>
<td>.197</td>
<td>.987</td>
<td>.402</td>
</tr>
<tr>
<td>Within Groups</td>
<td>19.120</td>
<td>96</td>
<td>.199</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>19.710</td>
<td>99</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The descriptive statistics for the four groups are as follows; Verb (M = 3.76, SD = .436), Noun (M = 3.84, SD = .374), Adjective (M = 3.68, SD = .476) and Adverb (M = 3.73, SD = .446) (Table 2).

Table 2: Descriptive Statistics Reading Comprehension by Groups

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Std. Error</th>
<th>95% Confidence Interval for Mean</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Lower Bound</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Upper Bound</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verb</td>
<td>25</td>
<td>3.76</td>
<td>.436</td>
<td>.087</td>
<td>3.58</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Noun</td>
<td>25</td>
<td>3.84</td>
<td>.374</td>
<td>.075</td>
<td>3.69</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Adjective</td>
<td>25</td>
<td>3.68</td>
<td>.476</td>
<td>.095</td>
<td>3.48</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Adverb</td>
<td>25</td>
<td>3.64</td>
<td>.490</td>
<td>.098</td>
<td>3.44</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Total</td>
<td>100</td>
<td>3.73</td>
<td>.446</td>
<td>.045</td>
<td>3.64</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>

5.2 Research Question 2

Is there any significant difference in the immediate recall of glossed parts of speech for the four experimental groups?
A one-way ANOVA was run. Based on the results displayed in Table 3 (F (3, 96) = 10.76, P = .000 < .05, \( \omega^2 = .22 \) it represent a large effect size) it can be concluded that there are significant differences between the mean scores of the four groups on the immediate recall test. Thus the second null-hypothesis as there is not any significant difference in in the immediate recall of glossed parts of speech for the four experimental groups is rejected.

Table 3: One-Way ANOVA Immediate Recall by Groups

<table>
<thead>
<tr>
<th></th>
<th>Sum of Squares</th>
<th>Df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>31.760</td>
<td>3</td>
<td>10.587</td>
<td>10.766</td>
<td>.000</td>
</tr>
<tr>
<td>Within Groups</td>
<td>94.400</td>
<td>96</td>
<td>.983</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>126.160</td>
<td>99</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The descriptive statistics for the four groups on the immediate recall test are as follows; Verb (M = 8.72, SD = .980), Noun (M = 8.44, SD = 1.044), Adjective (M = 8.64, SD = .952) and Adverb (M = 7.32, SD = .988) (Table 4).

Table 4: Descriptive Statistics Immediate Recall by Groups

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Std. Error</th>
<th>95% Confidence Interval for Mean</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td></td>
<td></td>
<td>Lower Bound</td>
<td>Upper Bound</td>
<td></td>
</tr>
<tr>
<td>Verb</td>
<td>25</td>
<td>8.72</td>
<td>.980</td>
<td>.196</td>
<td>8.32</td>
<td>9.12</td>
<td>7</td>
</tr>
<tr>
<td>Noun</td>
<td>25</td>
<td>8.44</td>
<td>1.044</td>
<td>.209</td>
<td>8.01</td>
<td>8.87</td>
<td>7</td>
</tr>
<tr>
<td>Adjective</td>
<td>25</td>
<td>8.64</td>
<td>.952</td>
<td>.190</td>
<td>8.25</td>
<td>9.03</td>
<td>7</td>
</tr>
<tr>
<td>Adverb</td>
<td>25</td>
<td>7.32</td>
<td>.988</td>
<td>.198</td>
<td>6.91</td>
<td>7.73</td>
<td>6</td>
</tr>
<tr>
<td>Total</td>
<td>100</td>
<td>8.28</td>
<td>1.129</td>
<td>.113</td>
<td>8.06</td>
<td>8.50</td>
<td>6</td>
</tr>
</tbody>
</table>

Although the F-value of 10.76 (Table 3) indicate significant differences between the mean scores of the four groups on the immediate recall test, the post hoc Scheffe’s tests should be run in order to locate the exact places of differences between any two groups. Based on the results displayed in Table 5 it can be concluded that;
A: There is not any significant difference between the mean scores of the Verb (M = 8.72) and Noun (M = 8.44) groups on immediate recall test (MD = .280, P = .802 > .05).

B: There is not any significant difference between the mean scores of the Verb (M = 8.72) and Adjective (M = 8.64) groups on immediate recall test (MD = .080, P = .994 > .05).

C: There is a significant difference between the mean scores of the Verb (M = 8.72) and Adverb (M = 7.32) groups on immediate recall test (MD = 1.40, P = .000 < .05). The Verb group outperformed the Adverb group on immediate recall test.

D: There is not any significant difference between the mean scores of the Noun (M = 8.44) and Adjective (M = 8.64) groups on immediate recall test (MD = -.20, P = .917 > .05).

E: There is a significant difference between the mean scores of the Noun (M = 8.44) and Adverb (M = 7.32) groups on immediate recall test (MD = 1.12, P = .002 < .05). The Noun group outperformed the Adverb group on immediate recall test.

Table 5: Post-Hoc Scheffe’s Tests Immediate Recall by Groups

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>(I) Treatment</th>
<th>(J) Treatment</th>
<th>Mean Difference (I-J)</th>
<th>Std. Error</th>
<th>Sig.</th>
<th>95% Confidence Interval</th>
<th>Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Immediate Vocab</td>
<td>Verb</td>
<td>Noun</td>
<td>.280</td>
<td>.280</td>
<td>.802</td>
<td>-.52</td>
<td>1.08</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Adjective</td>
<td>.080</td>
<td>.280</td>
<td>.994</td>
<td>-.72</td>
<td>.88</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Adverb</td>
<td>1.400*</td>
<td>.280</td>
<td>.000</td>
<td>.60</td>
<td>2.20</td>
</tr>
<tr>
<td></td>
<td>Noun</td>
<td>Adjective</td>
<td>-.200</td>
<td>.280</td>
<td>.917</td>
<td>-1.00</td>
<td>.60</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Adverb</td>
<td>1.120</td>
<td>.280</td>
<td>.002</td>
<td>.32</td>
<td>1.92</td>
</tr>
<tr>
<td></td>
<td>Adjective</td>
<td>Adverb</td>
<td>1.320*</td>
<td>.280</td>
<td>.000</td>
<td>.52</td>
<td>2.12</td>
</tr>
</tbody>
</table>

* The mean difference is significant at the 0.05 level.

F: There is a significant difference between the mean scores of the Adjective (M = 8.64) and Adverb (M = 7.32) groups on immediate recall test (MD = 1.32, P = .000 < .05). The Adjective group outperformed the Adverb group on immediate recall test.

5.3 Research Question 3

Is there any significant difference in the delayed recall of glossed parts of speech for the four experimental groups?

A one-way ANOVA was run. Based on the results displayed in Table 6 (F (3, 96) = 12.15, P = .000 < .05, $\omega^2 = .25$ it represent a large effect size) it can be concluded that there are significant differences between the mean scores of the four groups on the delayed recall test. Thus the third null-hypothesis as there is not any significant difference in in the delayed recall of glossed parts of speech for the four experimental groups is rejected.
Table 6: One-Way ANOVA Delayed Recall by Groups

<table>
<thead>
<tr>
<th></th>
<th>Sum of Squares</th>
<th>Df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>35.979</td>
<td>3</td>
<td>11.993</td>
<td>12.157</td>
<td>.000</td>
</tr>
<tr>
<td>Within Groups</td>
<td>94.709</td>
<td>96</td>
<td>.987</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>130.689</td>
<td>99</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The descriptive statistics for the four groups on the delayed recall test are as follows; Verb (M = 7.26, SD = .923), Noun (M = 6.68, SD = .930), Adjective (M = 6.76, SD = 1.151) and Adverb (M = 5.61, SD = .951) (Table 7).

Table 7: Descriptive Statistics Delayed Recall by Groups

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Std. Error</th>
<th>95% Confidence Interval for Mean</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Lower Bound</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verb</td>
<td>25</td>
<td>7.26</td>
<td>.923</td>
<td>.185</td>
<td>6.88</td>
<td>6</td>
<td>9</td>
</tr>
<tr>
<td>Noun</td>
<td>25</td>
<td>6.68</td>
<td>.930</td>
<td>.186</td>
<td>6.30</td>
<td>5</td>
<td>8</td>
</tr>
<tr>
<td>Adjective</td>
<td>25</td>
<td>6.76</td>
<td>1.151</td>
<td>.230</td>
<td>6.29</td>
<td>5</td>
<td>9</td>
</tr>
<tr>
<td>Adverb</td>
<td>25</td>
<td>5.61</td>
<td>.951</td>
<td>.190</td>
<td>5.22</td>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td>Total</td>
<td>100</td>
<td>6.58</td>
<td>1.149</td>
<td>.115</td>
<td>6.35</td>
<td>4</td>
<td>9</td>
</tr>
</tbody>
</table>

Although the F-value of 12.15 (Table 6) indicate significant differences between the mean scores of the four groups on the delayed recall test, the post hoc Scheffe’s tests should be run in order to locate the exact places of differences between any two groups. Based on the results displayed in Table 8 it can be concluded that;

A: There is not any significant difference between the mean scores of the Verb (M = 7.26) and Noun (M = 6.68) groups on delayed recall test (MD = .579, P = .243 > .05).
B: There is not any significant difference between the mean scores of the Verb (M = 7.26) and Adjective (M = 6.76) groups on delayed recall test (MD = .499, P = .373 > .05).
C: There is a significant difference between the mean scores of the Verb (M = 7.26) and Adverb (M = 5.61) groups on delayed recall test (MD = 1.64, P = .000 < .05). The Verb group outperformed the Adverb group on delayed recall test.
D: There is not any significant difference between the mean scores of the Noun (M = 6.68) and Adjective (M = 6.76) groups on delayed recall test (MD = -.08, P = .994 > .05).
E: There is a significant difference between the mean scores of the Noun (M = 6.68) and Adverb (M = 5.61) groups on delayed recall test (MD = 1.06, P = .004 < .05). The Noun group outperformed the Adverb group on delayed recall test.

Table 8: Post-Hoc Scheffe’s Tests Delayed Recall by Groups

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>(I) Treatment</th>
<th>(J) Treatment</th>
<th>Mean Difference (I-J)</th>
<th>Std. Error</th>
<th>Sig. 95% Confidence Interval</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>Delayed Vocab Ver</td>
<td>Noun</td>
<td>.579</td>
<td>.281</td>
<td>.243</td>
<td>-.22</td>
<td>1.38</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Adjective</td>
<td>.499</td>
<td>.281</td>
<td>.373</td>
<td>-.30</td>
<td>1.30</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Adverb</td>
<td>1.646*</td>
<td>.281</td>
<td>.000</td>
<td>.85</td>
<td>2.45</td>
<td></td>
</tr>
<tr>
<td>Delayed Vocab Ad</td>
<td>Noun</td>
<td>-.080</td>
<td>.281</td>
<td>.994</td>
<td>-.88</td>
<td>.72</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Adjective</td>
<td>1.067*</td>
<td>.281</td>
<td>.004</td>
<td>.27</td>
<td>1.87</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Adverb</td>
<td>1.147*</td>
<td>.281</td>
<td>.001</td>
<td>.35</td>
<td>1.95</td>
<td></td>
</tr>
</tbody>
</table>

* The mean difference is significant at the 0.05 level.

F: There is a significant difference between the mean scores of the Adjective (M = 6.76) and Adverb (M = 5.61) groups on delayed recall test (MD = 1.14, P = .001 < .05). The Adjective group outperformed the Adverb group on delayed recall test.

6. Conclusion
The current study aimed at investigating the differential effect of glossing for different parts of speech on retention and comprehension. The study found that the treatments did not affect the students’ performance on the reading test but they influenced their performance on the vocabulary tests. As a whole, it seems that among parts of speech adverbs are more difficult to learn. The study has some important pedagogical implications. The first one is making connection between meaning of an unknown lexical item and its grammatical role. According to (Richards, 1976), knowing a word means knowing the syntactic behaviours associated with the word. Another implication of the findings is due to the results of the vocabulary tests. Although the findings are relative, but they revealed that which word class is more difficult to learn and specific attention should be paid to teaching and learning it. The final implication is related to the selection of the TWs. The purpose of the study was enhancing learners’ vocabulary knowledge so any appropriate unknown words, not only the key words, were glossed which is important in EFL countries because students do not usually have enough exposure to the language to learn words incidentally.

Copyright © International Journal of English and Education | www.ijee.org
Several limitations were present in this research. The first one is the proficiency level. The population of current study consisted of four groups at intermediate level of language proficiency. Other researches of this nature should be replicated at other levels of learners such as ESL learners or advanced-level EFL learners. The second one is related to the gloss type. Because only L2 glosses were used in this study further experiments are necessary to examine the effects of other types of glosses. The third limitation of the study is due to its quantitative nature. Since different parts of language are not processed in the same way (VanPatten, 1994), this may motivate other studies to investigate related qualitative analyses.

References

Reading passages were downloaded from the following web sites:
http://www.englishclub.com/reading/levels-test-3c-levels.htm
http://www.cdlponline.org/index.cfm?fuseaction=homepage
http://englishforeveryone.org/Topics/Reading-Comprehension.htm
Appendix: Vocabulary Test

1. encourage:
   A. ask someone politely to do something
   B. advice someone very strongly to do something
   C. **give someone hope or confidence to do something**
   D. guide someone to do something correctly
2. repeal:
   A. cancel
   B. change
   C. rewrite
   D. develop
3. reduce:
   A. make equivalent
   B. **make less**
   C. make suitable
   D. make valuable
4. deter:
   A. make difficult
   B. **stop**
   C. make horrible
   D. follow
5. insist
   A. tell someone to do something legally
   B. decide freely on something
   C. **demand something forcefully or strongly**
   D. refuse to accept something strongly
6. spot
   A. assist
   B. arrest
   C. sue
   D. **notice**
7. apprehend:
   A. predict
   B. threat
   C. shoot
   D. **arrest**
8. constrain
   A. differ
   B. define
   C. **limit**
   D. select
9. infringe:
   A. act to limit or undermine someone's legal rights
   B. act to harm people intentionally
   C. take the trouble to do something
   D. behave in an illogical way
10. benefit:
    A. receive financial aid
11. authority:
A. legal documents you have because of your inquiries
B. the power you have because of your official position
C. the wealth you have because of your job
D. the popularity you have because of your knowledge

12. freshman
A. a first-year student at a university or high school
B. a practitioner at a university or high school
C. a second-year student at a university
D. a second-year student at a university or high school

13. attorney
A. adviser
B. supervisor
C. director
D. lawyer

14. alibi
A. documentary hypothesis
B. Documentary evidence
C. evidence of blameless
D. evidence of absence

15. administrator
A. manager
B. observer
C. spectator
D. teacher

16. prosecutor
A. a person who determines legal punishments exactly
B. a person who suggests legal activities to do
C. a person who examines legal proceedings against someone
D. a person who checks various documents thoroughly

17. track
A. blame
B. footprint
C. defect
D. document

18. transcript
A. a written version of material taught by a particular teacher
B. an original biography of students presented legally
C. an official record of a student’s work at a university or high school
D. a written record of material presented at a university or high school

19. incident
A. event
B. story
C. deceiver
D. misconduct

20. trial
A. council
B. congregation
C. court
D. department

21. premature:
   A. Too strange
   B. Too early
   C. common
   D. Painful

22. toxic
   A. Nuclear
   B. Poisonous
   C. Mysterious
   D. Bacterial

23. hazardous
   A. difficult to control
   B. difficult to recognize
   C. beneficial to health and the environment
   D. harmful to health of plants, animals or humans

24. infectious:
   A) an illness that can be passed from one person to another
   B) an illness that can be led to one or more types of cancer
   C) an illness that can be led to gradual damage to human health
   D) an illness that can be led to sudden death

25. indoor:
   A. smoky
   B. inside
   C. household
   D. burning

26. proper:
   A. suitable
   B. powerful
   C. modern
   D. chilly

27. leading:
   A. most obvious
   B. most dangerous
   C. most important
   D. most common

28. domestic:
   A. existing or happening under a particular condition, carefully controlled
   B. existing or happening inside a particular country, not foreign or international
   C. existing among people with certain characteristics
   D. existing or happening in most developing countries

29. accountable:
   A. deserved
   B. thoughtful
   C. confirmed
   D. Responsible

30. delightfully:
   A. carefully
31. exclusively:
   A. exactly
   B. uniquely
   C. definitely
   D. completely

32. meticulously:
   A. exactly
   B. reasonably
   C. absolutely
   D. really

33. daintily:
   A. uniquely
   B. carefully
   C. responsibly
   D. patiently

34. scarcely:
   A. shortly
   B. slowly
   C. hardly
   D. irregularly

35. frankly:
   A. honestly
   B. largely
   C. freely
   D. happily

36. exceedingly:
   A. entirely
   B. surely
   C. satisfactorily
   D. extremely

37. despairingly:
   A. doubtfully
   B. hopelessly
   C. sadly
   D. shakily

38. disgracefully
   A. shockingly unacceptable
   B. extremely violent
   C. strongly
   D. bitterly

39. eventually:
   A. gradually
   B. relatively
   C. finally
   D. entirely