Using Contextualized Reading Tasks for Improving Incidental Vocabulary Retention among Iranian EFL Learners

Shahrazad Hosseini and Hadi Salehi*

English Department, Najafabad Branch, Islamic Azad University
Najafabad, Iran

*Corresponding author’s email: hadisalehi1358 [AT] yahoo.com

ABSTRACT--- Vocabulary is central to English language teaching because without sufficient vocabulary students cannot understand others or express their own ideas. Vocabulary knowledge also plays an important role in almost all areas of language learning. So, the purpose of this study was to identify the relationship between the incidental vocabulary learning, task types, time devoted on tasks, and vocabulary retention of EFL learners. To achieve the aim, this research employed a quasi-experimental design to investigate the possible effects of various contextualized reading tasks and timing on the L2 incidental vocabulary acquisition of Iranian EFL learners. Sixty intermediate EFL learners in six intact classes in Zabansara language institute in Isfahan were accessed. Three classes were respectively exposed to dictionary-based, fill-in-the-blank, and sentence-making tasks with time restrictions, and the other three classes were respectively exposed to dictionary-based, fill-in-the-blank, and sentence-making tasks without time restrictions. Having experienced their relevant treatments, the participants were asked to take a delayed vocabulary test. Since task type and timing were two independent variables and vocabulary learning/recall was the dependent variable of the study, two-way ANOVA was used to compare the learners’ vocabulary scores on the delayed vocabulary test. The results revealed that time restriction exerted a significant effect on the performance of the learners in all the three groups on the delayed tests. Also, learners in the groups without time restriction significantly outperformed those in the groups with time restriction.

Keywords--- Incidental Vocabulary Learning, Task Type, Time Devoted on Tasks, Delayed Vocabulary Test, Vocabulary Retention

1. INTRODUCTION

Vocabulary is central to English language teaching because without sufficient vocabulary students cannot understand others or express their own ideas. Stern (1992) wrote that “. . . while without grammar very little can be conveyed, without vocabulary nothing can be conveyed” (pp. 111–112). Skehan (1998) went further to argue, “lexis is the core or heart of language” (p. 89). Particularly as students develop greater fluency and expression in English, it is significant for them to acquire more productive vocabulary knowledge and to develop their own personal vocabulary learning strategies. Students often instinctively recognize the importance of vocabulary to their language learning. As Skehan (1998) noted, “learners carry around dictionaries and not grammar books” (p. 4). Ian S P Nation (2001) stated that the vocabulary of a language is huge and its acquisition takes time even for a native speaker. Some practitioners believe that vocabulary learning is simple. Majority of students learning a second or foreign language consider vocabulary as their major priority and agree that their difficulties arise from their lack of vocabulary knowledge. However, in second language research, it is only in recent years that a growing concern with vocabulary research has been raised.

Because there is a definite gap between what is taught and what is learned, the issue of incidental vocabulary learning comes into a great significance. Teachers have to ensure that learners know the basis of the target language, its grammar, phonetics, spelling, and vocabulary. Once this threshold is reached, learners are sufficiently autonomous to expand their vocabulary by engaging in different tasks like reading comprehension and writing. According to this view, teachers have a sufficient role to help and encourage students to read and write more and more during which they learn many vocabularies by the process known as incidental learning, that is students learn words even when they are not paying any particular attention to them. Therefore, an attempt has been made in this study to investigate whether incidental vocabulary learning through different tasks, e.g., dictionary-based reading comprehension task, fill in the blanks tasks, and sentence making with the target words, establishes and strengthens the lexical form-meaning connections in the student's mental lexicon and increase the likelihood of vocabulary retention or not. So, the following question is expected to be answered in this research: are there any significant differences among vocabulary scores of learners in dictionary-based, fill-in-the-blank, and sentence-making groups on the delayed vocabulary test with and without a time restriction?
2. LITERATURE REVIEW

Recently, the importance of vocabulary acquisition has been emphasized by researchers and commentators, since vocabulary acquisition plays a significant role in learning a second language (Hogben & Lawson, 1994; S P Nation, 1990). However, there is a lack of agreement regarding the conceptualization of the process concerning what contributes to vocabulary acquisition. They stress out that another determining factor concerning vocabulary acquisition is the importance of context and the value of reading (Mondria, 2003).

Different scholars define vocabulary learning strategies differently, some of which are: 1) memorizing strategies, 2) repeating strategies, 3) association strategies, 4) key word method, 5) inferencing strategy, 6) dictionary use, 7) semantic grid strategies, 8) word lists (Farhady & Delshad, 2007). According to Farhady and Delshad (2007), applying certain types of strategies forms an approach to vocabulary learning that influences the level of foreign language proficiency. In other words, appropriate strategy use results in improved achievement in specific skills or sub-skills. In addition, language proficiency also affects the use of particular vocabulary strategy use.

2.1. The Role of Contextualization

Moltz (2010) believes that contextualization is a form of “deep learning” which occurs via linking ideas and concepts across courses. In language teaching context, it refers to putting the target language in a realistic setting in order to make the learning process meaningful to the learners. As for the contextual approach, Mayer (2003) points out that instructional method cannot be used out of context. Nunan (1999, p.304) gives another definition of context that he terms as “the linguistic and experiential situation in which a piece of language occurs”. N. C. Ellis (1994) emphasizes the efficiency of contextualized tasks, suggesting that contextualization strategies work better for those learners who have an acceptable level of second language knowledge. According to Mazzeo, Rab, and Alssid, (2003, p.3-4), contextualization refers to a set of instructional strategies that focus on “concrete applications in a specific context that is of interest to the student”. Different terminologies used for contextualization among them are the followings; including contextualized instruction (Parr, Edwards, & Leising, 2008), content-area literacy (McKenna & Robinson, 2013) contextual teaching and learning (Johnstone & Shanks, 2001) embedded instruction (Simpson, Hynd, Nist, & Burrell, 1997) functional context education, etc. Contextualization includes several key terms: teaching skills through real world events and practices, content courses instruction in basic skills (Perin, Hare, Pevery, & Mason, 2010). It also takes place in the basic skills and subject-area instruction (Paquette & Kaufman, 2008).

2.2. Significance of Reading Instruction

Smith (2004) declares that reading is a complex cognitive activity, almost a miraculous one, in fact, since it involves the secondary uses of cognitive skills in relatively new ways, at least in terms of evolutionary development. Reading is not a natural process inherently like that speaking and listening are in a first language (L1). Nothing is free with respect to reading; unlike our first spoken language, which one might say "comes for free.” Considerable cognitive effort and a long learning process are required for learning to read, whether one is learning to read in the L1 or in a L2. If a person is not taught to read, in one way or another (e.g., by a teacher, a parent, a sibling), that person will not learn to read (Smith, 2004). Smith (2004) concludes that as a consequence, the teaching of reading is also a complex matter. Obvious variables such as student proficiency, age, L1/L2 relations, motivation, cognitive processing factors, teacher factors, curriculum and materials resources, instructional setting, and institutional factors all impact the degree of success of reading instruction.

3. METHODOLOGY

The present study, therefore, employed a quasi-experimental design to investigate the possible effects of various contextualized reading tasks and timing on the L2 incidental vocabulary acquisition of Iranian EFL learners. Sixty intermediate EFL learners in six intact classes in Zabansara language institute in Isfahan were accessed. Three classes were respectively exposed to dictionary-based, fill-in-the-blank, and sentence-making tasks with time restrictions, and the other three classes were respectively exposed to dictionary-based, fill-in-the-blank, and sentence-making tasks without time restrictions. Having experienced their relevant treatments, the participants were asked to take a delayed vocabulary test. Based on the research objective, the following hypothesis is formulated: there is no significant difference among vocabulary scores of learners in dictionary-based, fill-in-the-blank, and sentence-making groups on the delayed vocabulary test with and without a time restriction. Since task type and timing were two independent variables and vocabulary learning/recall was the dependent variable of the study, two-way ANOVA was conducted to capture any possible differences among the learners in different groups on the delayed vocabulary test, and thus provide an answer to the research question.
3.1. Research Participants

The participants of the study were sixty male EFL learners from Zabansara institute in Isfahan. They were within 18-20 age-range, of the same nationality (Iranian) and shared Persian as their native language. They were at an intermediate level of proficiency based on the institute’s continuous assessment criteria, which was more realistic than a summative test. Their teachers also were of the opinion that the participants were capable of carrying out the required tasks which, thus further corroborates the institute’s assessment of the learners’ level. Then, the participants were randomly divided into two groups with having three subgroups in each.

3.2. Research Instruments and Materials

For the purpose of this study, the following instruments and materials were used:

3.2.1. Reading Text

The reading passage was an article selected from a reading-comprehension book, *Reading Master* (Liu et al., 2002) which had already been used by Scholfield (1997) in a study with a similar purpose. The passage is about the suppression of emotions and the potential threats of such behavior to the mental and physical health of human beings. The reason for the selection of this topic was that it was of a general nature and was understandable to the participants, who could relate it somehow to their own personal experiences.

The text consisted of 331 words and was selected on two grounds: (i) participants are supposed to have some general ideas of the topic yet little knowledge of the words relevant to the issue; (ii) learners are capable of writing about their personal experience pertinent to this topic.

Five multiple-choice reading comprehension questions were adopted from the same reading material. The understanding of the ten target lexical items is generally relevant to the completion of the comprehension questions. The criterion for modifying the text is the number of occurrences for each target word. The passage was revised in such a way that all the target words and their roots would appear only once.

3.2.2. Target Words

Ten target words were selected from the reading text, based on three criteria: (i) assumed unfamiliarity to the participants, (ii) ease of incorporating the words into a narrative gender of writing describing personal experiences, and (iii) ease of supplying a synonym or a definition in English as well as an appropriate translation in Persian in the immediate vocabulary test. Each criterion was judged mainly by two lecturers who have extensive experience in teaching English to university students.

The participants’ own teachers, who had wide experience in teaching English to Iranian students in institutes, were consulted regarding the above criteria. They assured the researcher of their appropriateness. The unfamiliarity of the target words to the participants was also ascertained by checking the target words and their respective word families against the list of each of the previously taught as well as their current textbooks. Besides, these ten words were presented to a group of students of the same level who were not supposed to attend the study to ensure the participants’ lack of knowledge of these words.

The ten target words in the reading were annoy, hostile, conflict, unfortunately, suppression, maintain, determination, circumstance, grit, and endure (four nouns, four verbs, one adjective, and one adverb).

3.2.3. Tasks

Three tasks with varying involvement loads were used to serve the purpose of the present study. The two main groups of the study shared identical tasks, along with the factor of time-on-task.

*Task 1:* Dictionary-based reading comprehension. It was performed by two of the six groups. Learners performing task were provided with a text and five multiple-choice comprehension questions based on the reading passage. These questions either incorporated some target words or paraphrased the original sentences in which these target words occurred. Accordingly, the successful completion of the questions entailed the understanding of the target lexical items. In the reading passage, the ten target words were highlighted in bold print. Students’ task was to read the text and answer the five comprehension questions. It had to be noted that all the participants in these two main groups had already been trained by their teachers how to use a dictionary.
Task 2: Fill in the blanks reading comprehension task. Students of other two subgroups performing this task were given the same text and the same questions as those performing Task 1. In Task 2, the ten target words were omitted from the text, leaving ten gaps to be filled in. The ten target lexical items, along with five distracters that did not appear in the original passage, were provided for the students with their English explanations. The task was to read the article, fill in the ten gaps with the words from the list of 15 words, and answer the same comprehension questions as the first group.

Task 3: Sentence making with ten target words. The participants in the last two subgroups performing this task were required to make sentences to express their feelings about the incidence in the passage using the target words given to them. The students were informed that grammaticality was of secondary importance, and that the clarity of the main idea of the sentences as well as the incorporation of the ten target words would account for the most part of the scoring criteria. The same ten words and their respective glosses were given as in Task 1, but with a sample phrase for each target word. The provision of a phrase rather than a sentence was to minimize the possibility that students might simply copy the sentences to their compositions, and thus reduce the need to elaborate processing of the words.

3.2.4. Vocabulary Tests

Delayed Test: The participants in this study sat for an unexpected delayed test two weeks after performing the tasks to determine the retention of the target words. This vocabulary test was composed of 15 multiple choice items in which the participants were asked to choose the best answer for each item. This vocabulary test measured receptive knowledge only, as this study aimed to investigate learners’ retention of the meaning of these target lexical items.

3.3. Research Procedures

The six sub-groups in the two main groups (three groups each) were randomly assigned to perform one of the three tasks during regular English class sessions. The students followed their English teachers’ instructions to perform the tasks and were not informed of an on-coming retention test of the target lexical items in the reading passage. None of the tasks was presented as a vocabulary-learning task, with the first two tasks being introduced as reading activities and the third task as a writing exercise. As is delineated in the section on the tasks, time-on-task was not controlled in the first experiment. Consequently, the three sub-groups in the first main group spent 16, 28, and 30 minutes accomplishing their assigned task, respectively.

The time limit in the second experiment, nevertheless, was kept constant across different tasks (35 minutes). In both experiments, the work sheets were collected after the completion of the task, and the students were given a vocabulary test sheet with a list of twenty five lexical items, for which they were requested to provide meanings (either in Persian or English). The participants were also required to indicate whether they had known the words prior to the tasks. This practice was an additional check for the pre-knowledge of the ten target lexical items on the part of the learners. The test sheets were then collected and not returned to the students. Two weeks later, the students received a delayed test in the form of multiple choice questions. The time limit for the immediate test was 10 and for the delayed test 15 minutes.

The vocabulary tests were scored by the researcher. A word that was not glossed (either in English or Persian) or was wrongly glossed was assigned the score of zero. A correct response received 1 point. A semantically approximate response was awarded 0.5 point. If a learner gave a correct response but had also marked the target word as known prior to the experiment, the response was scored as zero. In delayed vocabulary test, the question which was marked correctly in the answer sheet scored one.

4. RESULTS AND FINDINGS

The research question was formulated to see whether there were any significant differences among vocabulary scores of learners in dictionary-based, fill-in-the-blank, and sentence-making groups on the delayed vocabulary test with and without a time restriction. The results obtained through two-way ANOVA are presented in this part.
Table 1: Descriptive Statistics for Comparing the Delayed Vocabulary Scores of the Learners in Different Groups with Different Timing Conditions

<table>
<thead>
<tr>
<th>Groups</th>
<th>Timing</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>DBG</td>
<td>WTR</td>
<td>5.60</td>
<td>.84</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>WOTR</td>
<td>5.50</td>
<td>1.08</td>
<td>10</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>5.55</td>
<td>.94</td>
<td>20</td>
</tr>
<tr>
<td>FITBG</td>
<td>WTR</td>
<td>6.20</td>
<td>.42</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>WOTR</td>
<td>6.70</td>
<td>.82</td>
<td>10</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>6.45</td>
<td>.68</td>
<td>20</td>
</tr>
<tr>
<td>SMG</td>
<td>WTR</td>
<td>6.80</td>
<td>.63</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>WOTR</td>
<td>7.90</td>
<td>.87</td>
<td>10</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>7.35</td>
<td>.93</td>
<td>20</td>
</tr>
<tr>
<td>Total</td>
<td>WTR</td>
<td>6.20</td>
<td>.80</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>WOTR</td>
<td>6.70</td>
<td>1.34</td>
<td>30</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>6.40</td>
<td>1.12</td>
<td>60</td>
</tr>
</tbody>
</table>

As for the time restriction condition, the mean scores of the DBG, FITBG, and SMG members were 5.60, 6.20, and 6.80, respectively. Regarding no time restriction condition, the mean scores of the DBG, FITBG, and SMG learners turned out to be 5.50, 6.70, and 7.90, respectively. In addition, the total mean score of the time restriction condition groups was less than that of no time restriction condition groups (6.20 < 6.70). To figure out whether the differences among the groups and between timing conditions were of statistical significance or not, one should check the p values in front of Groups and Timing under the Sig. column in Table 2.

Table 2: Results of Two-Way ANOVA for Comparing the Delayed Vocabulary Scores of the Learners in Different Groups with Different Timing Conditions

<table>
<thead>
<tr>
<th>Source</th>
<th>Type III Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
<th>Partial Eta Squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corrected Model</td>
<td>39.75</td>
<td>5</td>
<td>7.95</td>
<td>12.23</td>
<td>.000</td>
<td>.53</td>
</tr>
<tr>
<td>Intercept</td>
<td>2496.15</td>
<td>1</td>
<td>2496.15</td>
<td>3840.23</td>
<td>.000</td>
<td>.98</td>
</tr>
<tr>
<td>Groups</td>
<td>32.40</td>
<td>2</td>
<td>16.20</td>
<td>24.92</td>
<td>.000</td>
<td>.48</td>
</tr>
<tr>
<td>Timing</td>
<td>3.75</td>
<td>1</td>
<td>3.75</td>
<td>5.76</td>
<td>.020</td>
<td>.09</td>
</tr>
<tr>
<td>Groups*Timing</td>
<td>3.60</td>
<td>2</td>
<td>1.80</td>
<td>2.76</td>
<td>.072</td>
<td>.09</td>
</tr>
<tr>
<td>Error</td>
<td>25.10</td>
<td>54</td>
<td>.65</td>
<td>.55</td>
<td>.76</td>
<td>.76</td>
</tr>
<tr>
<td>Total</td>
<td>2571.00</td>
<td>60</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>74.85</td>
<td>59</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2 shows that there was a statistically significant difference in the delayed vocabulary scores of the learners who had learned vocabulary through dictionary-based, fill-in-the-blank, and sentence-making tasks due to the fact that the p value under the Sig. column in front of Groups was less than the alpha level of significance (i.e. .000 < .05); the exact location of the differences among these groups will be presented in the post hoc Scheffe table (Table 3). In addition, the p value in front of Timing was smaller than the significance level (.000 < .05), which means that the groups without time restriction (M = 6.70) outperformed those in the time restriction condition (M = 6.20). Nevertheless, the interaction between task type and timing did not reach statistical significance since the p value in front of Groups*Timing was greater than the significance level (.072 > .05). The results obtained here are also graphically shown in figure 1.

Table 3: Results of Scheffe Post Hoc Test

<table>
<thead>
<tr>
<th>Groups</th>
<th>Mean Difference</th>
<th>Std. Error</th>
<th>Sig.</th>
<th>95% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Lower Bound  Upper Bound</td>
</tr>
<tr>
<td>DBG</td>
<td>FITBG</td>
<td>-.90*</td>
<td>.25</td>
<td>.004</td>
</tr>
<tr>
<td>SMG</td>
<td></td>
<td>-.1.80*</td>
<td>.25</td>
<td>.000</td>
</tr>
<tr>
<td>FITBG</td>
<td>DBG</td>
<td>.90*</td>
<td>.25</td>
<td>.004</td>
</tr>
<tr>
<td>SMG</td>
<td></td>
<td>.90*</td>
<td>.25</td>
<td>.004</td>
</tr>
<tr>
<td>SMG</td>
<td>FITBG</td>
<td>1.80*</td>
<td>.25</td>
<td>.000</td>
</tr>
<tr>
<td>FITBG</td>
<td></td>
<td>.90*</td>
<td>.25</td>
<td>.004</td>
</tr>
</tbody>
</table>

Asian Online Journals (www.ajouronline.com) 131
The difference between DBG ($M = 5.55$) learners and FITBG ($M = 6.45$) learners on the delayed test of vocabulary was statistically significant because the relevant $p$ value under the Sig. column was .004, which is lower than the significance level. Likewise, there was a statistically significant difference between DBG and SMG ($M = 7.35$) learners since the corresponding $p$ value was .000. Finally, the difference between FITBG and SMG members reached statistical significance since the $p$ value was less than the alpha level (i.e. .004 < .05). These obtained results are graphically represented in Figure 1.

![Figure 1: The Mean Scores of the Different Groups with Different Timing Conditions on the Delayed Vocabulary Test](image)

Figure 1 shows that for time restriction condition (shown here by the blue line), the mean score of the SMG learners was greater than the mean score of the FITBG learners, which in turn was greater than the mean score of the DBG learners. Moreover, the differences among these three groups of learners were statistically significant. The same was also true for the no time restriction condition, which was shown by the green line in this figure: the mean score of the SMG group was larger than that for the FITBG members, whose mean score was in turn greater than the mean score of the DBG learners, and the differences among these three groups were statistically significant as well. Finally, in terms of timing, there was a statistically significant difference between the learners exposed to different tasks in the time restriction condition, and those who underwent the same tasks in no time restriction condition. In a nutshell, on the delayed vocabulary test, task type and timing were both shown to exert significant effects on the learners’ performances.

5. DISCUSSION AND CONCLUSION

The null hypothesis of the study was formulated as follows: there is no significant difference among vocabulary scores of learners in dictionary-based, fill-in-the-blank, and sentence-making groups on the delayed vocabulary test with and without a time restriction. In order to examine this hypothesis, another two-way ANOVA was used to compare the learners’ vocabulary scores on the delayed vocabulary test across the three groups with and without the time restriction factor.

The results of data analysis revealed that there was a statistically significant difference in the delayed vocabulary scores of the learners who had learned vocabulary through dictionary-based, fill-in-the-blank, and sentence-making tasks. More specifically, it was found that for time restriction condition, the mean score of the SMG learners was greater than the mean score of the FITBG learners, which in turn was greater than the mean score of the DBG learners. It should be noted that the differences among these three groups of learners were statistically significant. The same results were obtained for the no time restriction condition: the mean score of the SMG group was larger than that for the FITBG members, whose mean score was in turn greater than the mean score of the DBG learners, and the differences among these three groups were statistically significant as well. Moreover, in terms of timing, there was a statistically significant difference between the learners exposed to different tasks in the time restriction condition, and those who underwent the same tasks in no time restriction condition. Consequently, it can be claimed that task type and timing both exerted a significant effect on the learners’ performance on the delayed vocabulary test.
The findings regarding the null hypothesis of this study are in accordance with those of other researchers who have reported the benefits of long-term retention of vocabulary learned through incidental learning (Lafer & Hulstijn, 2001; Lee, 2003; Moradian, 2005; Paribakh & Wesche, 1997; Swamborn & De Glocker, 2002). Moreover, the findings of this study lend further support to Lafer and Hulstijn (2001) whose findings came out to be in favor of the sentence writing task compared to the other two tasks. In other words, in compliance with our results, they also reported that the participants revealed better performance on vocabulary retention test after performing the output productive task with a higher involvement load—namely, sentence writing. Partial explanation for the significant differences among the three groups on the delayed test in this study comes from Lafer and Hulstijn’s (2001) Involvement Load Hypothesis, asserting that retention scores are related to the amount of task-induced involvement load; that is, retention was highest in the sentence writing task, lower in Fill-in-the-blank task, and lowest in the dictionary-based task because sentence writing entails a deeper cognitive processing level, leading to better retention. Nevertheless, the findings of this study oppose the findings of other researchers who have rejected the positive effects of incidental learning tasks. For instance, Huckin and Coady (1999) have reported the limitations of incidental learning at the practical level. According to them, incidental learning is susceptible to certain serious limitations such as the need for a tactical use of inferences as well as substantial prior vocabulary knowledge on the part of the learner. With regard to time restriction, as stated above in the discussion of the first hypothesis of the study, our findings are in accordance with those of Hulstijn and Trompetter (1998) and Lafer and Hulstijn (2001) who claim that the factor of time-on-task has a significant impact on vocabulary retention. In other words, the group without time restriction had a better performance than the group with time restriction on the delayed vocabulary test.

6. REFERENCES


