The Effects of glosses and tasks on L2 vocabulary acquisition

: A pilot study

SLS 620: Second Language Reading

Instructor: Dr. Thom Hudson

Spring 2007

Jung Min Lee
Almost all of the second or foreign language learners become aware of the necessity of having rich vocabulary knowledge while reading. The awareness of the essential role of vocabulary knowledge on reading drives many researchers to be primarily interested in how vocabulary learning occurs and accordingly investigate what factors would facilitate vocabulary acquisition in terms of initial learning and retention.

With regard to the issue of how vocabulary learning happens, there are two trendy views; intentional learning and incidental learning. Intentional learning involves deliberate committing of memory of thousands of words, whereas incidental learning refers to “picking up” of words, simply engaging in a various communicative activities and focusing on the meaning rather than the form of the words (Hulstijn, 2003, p. 349). This is in line with Schmidt’s interpretation on incidental learning, referred as “learning of one thing (e.g., grammar) when the learner’s primary objective is to do something else (e.g., communicate).” (1994a, p. 16). Taken their views, for example, readers can learn unknown words while reading for the meaning. Moreover, there is a general agreement that little vocabulary is acquired in an intentional mode and much vocabulary is incidentally, incrementally learned through reading. In the experiment by Nagy and Herman (1987), the students cannot have learned a large amount of words exclusively by means of explicit vocabulary instruction. Rather, they must have learned many words through frequent encounters during extensive reading.

Regarding learning-vocabulary-through-reading, however, it has fundamentally been argued on whether simply exposing to the extensive reading could guarantee acquisition of all the words the learners need (Krashen, 1989) or whether the inclusion of cooperating tasks and/or explicit strategies like using glosses or dictionaries facilitates vocabulary acquisition (Hulstijn, 1992; Hulstijn, Hollander, & Greidanus, 1996; Rott, William, & Cameron, 2002; Watanabe,
Hulstijn et al. (1996) presented some reasons for that not always readers succeed in learning vocabulary incidentally during reading for meaning: (a) the readers sometimes fail to notice the unknown words or believe that they know a word, but in reality they do not; (b) the readers sometimes decide to ignore the unknown words even though they notice their presence; (c) the context may be so redundant that the readers fail to connect the form and meaning of the unknown words; and (d) the context does not give sufficient clues to infer the meaning.

Based on the findings, it is shown that there are such crucial factors as attention and noticing, need, or contextual clues to determine whether or not vocabulary learning occurs more successfully while reading for the meaning. Schmidt (1994a) argues that although there may be some forms of learning without attention, it is necessary to have attention and noticing to input to improve encoding system since many features of input are likely to be infrequent, non-salient, and communicatively redundant. Need is a motivational concept, which has already been well-known on promoting achievement in second language (L2) learning. Laufer and Hulstijn (2001) interpreted need in its positive sense, that is, need to achieve, and take account of it as one of task-induced involvement constructs which substantially affect on successful vocabulary learning. For example, readers want to refer to an unknown word when they recognize its necessity for comprehension. Otherwise, they conveniently skip over the unneeded words without paying attention and attempting to guess the meaning; therefore, the likelihood of vocabulary learning will be comparatively declined (Hulstijn, 1993). Inferring the meaning of unknown word is largely contingent on the context. Since not all contexts are reader-friendly, however, they have been modified by the name of elaboration to become eligible enough to provide sufficient clues (Kim, 2006; Urano, 2000). Such devices as glosses (Hulstijn, 1992;
Watanabe, 1997) or a dictionary (Knight, 1994; Luppescu & Day, 1993) in order to compensate for lack of helpful context have also been investigated by many researchers.

One other essential factor conducive to vocabulary acquisition is elaboration. When readers process new lexical information more elaborately, they are more likely to have higher retention than by doing so less elaborately. Hulstijn (1992) proposed mental effort hypothesis in a series of his experiments, holding the same view that the retention of inferred meaning will be better than the retention of given meaning for words. This concept was originated from depth of processing hypothesis (Craik & Tulving, 1975), wherein depth would mostly be compliance with greater degree of semantic involvement. They argued that semantic processing of lexical items would result in higher retention than phonological or orthographical processions.

Recently, Laufer and Hulstijn (2001) introduced the construct of a task-induced involvement for incidental L2 vocabulary acquisition. According to them, “retention of words when processed incidentally, is conditional upon three factors in a task: need, search, and evaluation.” (p. 14). The need component is the motivational, non-cognitive dimension of involvement, which is based on a drive to comply with the task requirements. Search and evaluation components are cognitive dimensions of involvement, conditional upon noticing and consciously assigning attention to the form-meaning relationship. In this sense, when learners need to infer the word, attempt to search the meaning of the unknown words, and evaluate whether the inferred meaning fits in the context or not, the possibility of learning the words and its retention will be amplified as a result. Thus, other factors being equal, words with higher involvement load will be retained better than ones processed with lower involvement load (p. 15). Also, other factors being equal, carefully designed tasks with a higher involvement load will be
more effective for learning the words and its retention than ones with a lower involvement load (p. 17).

This study investigates the effects of different types of glosses and task inclusion on incidental L2 vocabulary acquisition. The following research question guides the study: how do different types of glosses and inclusion of a task affect vocabulary acquisition in terms of initial learning and retention?

**METHOD**

**Participants**

24 L2 English learners enrolled in a U.S. university participated in the study. They were undergraduate and graduate students from different countries such as Japan, China, Korea, Taiwan, Bulgaria, Mongolia, and Vietnam. All the participants were assigned to one of four different subgroups: (a) Definitional Glosses (DG) and a fill-in-the-blank task, (b) only DG, (c) Contextual Glosses (CG) and a fill-in-the-blank task, and (d) only CG. In this 2X2 factorial multivariate analysis of variance (MANOVA) design, six students were assigned for each cell. All the participants were provided with seven dollars for compensation.

**Materials**
**Reading Passage** The reading passage was selected from the articles in a reading textbook, *Exploring Content 1*, which aims for college-level L2 English learners. The article was about the tendency of interpreting people’s behaviors without considering the power of social influence. The modified passage contained 446 words and its readability by Flesch-Kincaid grade level was 10.9. Ten target words were chosen from the passage on the basis of three criteria: (a) basically assumed unknown words to the study participants, (b) words which did not occur once again to exclude the frequency effect of words, (c) content words which functioned as nouns, verbs, or adjectives both for the ease of supplying a synonym or a definition and for clarity of scoring the answers in either English or native tongue. Two versions of reading passage were developed: (a) a passage with DG and (b) a passage with CG, both of which were located in the side margin of the passage. A synonym or a definition of the target words was provided for DG and a sentence including the target words was for CG (see Appendix A).

**Reading Comprehension Question** Asking the main ideas of the passage was employed for reading comprehension question to turn participants’ main attention to an understanding of the passage as a whole, not to make them solely focus on the target words. Using a reading question as a distractor was believed to provide the study with more reliable results on investigating incidental vocabulary learning.

**Procedure**

The study was conducted in a pretest-(treatment)-posttest-delayed posttest design throughout two meetings with individual or group. In the first meeting, the participants were
asked to do a pretest, a biodata, a reading comprehension question, a fill-in-the-blank task for treatment groups, and a surprise posttest. In the second meeting, a week later, they were asked to do an unexpected delayed posttest. There was no strict time limit for the participants to accomplish the required tasks. Detailed procedures of the study are presented in the following.

**Step 1: Pretest and collecting Biodata from Research Participants**  In order to measure the depth of vocabulary knowledge of target words, Vocabulary Knowledge Scale (VKS) adapted from Wesche and Paribakht (1996) was assigned to participants (see Appendix B). The pretest asked participants to check the statement which applies to their vocabulary knowledge scale and to provide the meaning of the word or a sentence including the target word, if applicable. The items in pretest were composed of ten target words and ten additional words, both of which appeared in the reading passage, and were positioned in a random order.

In order to collect background information of the research participants, a questionnaire was developed and delivered to the twenty four students. The items in the questionnaire asked such general information as the students’ age, gender, institution, duration of English study, duration of stay in English-speaking countries, and so forth. Also, to minimize the frequency effect of target words, all participants were asked to write down and talk about some interesting experiences requiring to trace their memory such as studies, majors, future plans, travel, or the like.

**Step 2: Treatment**  In this session, all the participants were asked to answer comprehension question after reading the passage with either DG or CG. Then, half of the participants conducted a fill-in-the-blank task for the treatment. The task asked the treatment group to fill in the blanks using an appropriate word among the target words in the passage (see Appendix C).
**Step 3: Posttest** The third section consisted of vocabulary posttest, wherein only ten target words were contained but the same directions were provided.

**Step 4: Delayed Posttest** The fourth section was the same as the third except for being conducted one week after initial posttest.

The following Table 1 displays the general scheme of the study.

<table>
<thead>
<tr>
<th>Procedure</th>
<th>DG, +T ( n = 6 )</th>
<th>DG, –T ( n = 6 )</th>
<th>CG, +T ( n = 6 )</th>
<th>CG, –T ( n = 6 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st Meeting</td>
<td>Biodata and pretest</td>
<td>Pre-VKS and questionnaire (10-15 min)</td>
<td>Reading comprehension question (15-20 min)</td>
<td>No task</td>
</tr>
<tr>
<td>Treatment</td>
<td>Fill-in-the-blank task (10 min)</td>
<td>No task</td>
<td>Fill-in-the-blank task (10 min)</td>
<td>No task</td>
</tr>
<tr>
<td>Posttest</td>
<td>No task</td>
<td>Post-VKS (5-10 min)</td>
<td>Delayed post-VKS (5-10 min)</td>
<td>Delayed post-VKS (5-10 min)</td>
</tr>
<tr>
<td>2nd Meeting</td>
<td>Delayed posttest</td>
<td>No task</td>
<td>No task</td>
<td>No task</td>
</tr>
</tbody>
</table>

**Scoring**

**Vocabulary Knowledge Scale (VKS)** Scoring was performed with 0-5-point scale ranging from the first statement which refers to complete unfamiliarity to the word to the last one which requires making a sentence using the target word. Each category of the VKS was added up individually. A correct response received one point and an incorrect one zero point.
**Reading Comprehension Question**  Since this question intentionally served as a distractor to turn participants’ main attention to an understanding of the passage as a whole, not to make them solely focus on the target words; there was no scoring for this question.

**Data Analysis**

This study investigated the effects of two types of glosses and a task on incidental L2 vocabulary acquisition; involving a pretest-(treatment)-posttest-delayed posttest design. The research question was addressed by employing a two-way factorial MANOVA, with types of glosses (two levels: DG, CG) and presence or absence of a fill-in-the-blank task (two levels: a task, no task) as between-subjects factors. This was to determine whether or not two categorical independent variables (and their interactions) significantly would affect the performance of the study participants measured by two dependent interval variables (i.e., gain scores of posttest and of delayed posttest from pretest). All the statistical analyses were conducted using SPSS, with the alpha level set at 0.05.

**RESULTS**

To measure the depth of vocabulary knowledge, the gain scores of VKS in terms of posttest and delayed posttest were used as independent variables for a two-way MANOVA. Table 2 shows that the descriptive statistics of the gain scores of posttest and of delayed posttest from pretest.
Table 2. *Descriptive statistics for gain scores from pretest of the four treatment groups*

<table>
<thead>
<tr>
<th>Group</th>
<th>Posttest</th>
<th></th>
<th>Delayed posttest</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>DG, +T (n = 6)</td>
<td>16.83</td>
<td>8.57</td>
<td>11.67</td>
<td>7.63</td>
</tr>
<tr>
<td>DG, –T (n = 6)</td>
<td>7.67</td>
<td>3.44</td>
<td>7.50</td>
<td>2.81</td>
</tr>
<tr>
<td>CG, +T (n = 6)</td>
<td>13.33</td>
<td>9.09</td>
<td>9.50</td>
<td>9.44</td>
</tr>
<tr>
<td>CG, –T (n = 6)</td>
<td>5.33</td>
<td>4.80</td>
<td>3.33</td>
<td>3.14</td>
</tr>
<tr>
<td>Total (N = 24)</td>
<td>10.79</td>
<td>7.94</td>
<td>8.00</td>
<td>6.76</td>
</tr>
</tbody>
</table>

As shown in the Table 2, the participants with DG and a task scored best and those with CG and no task did worst in cases of both posttest and delayed posttest. The students with DG performed better than those with CG, and the students with a task did better than those without a task. A two-way MANOVA showed that the main effect of a task was statistically significant, Wilks’ Lambda = 0.67, $F(2, 19) = 4.71$, $p < 0.05$, whereas the main effect of glosses did not show statistical significance, Wilks’ Lambda = 0.93, $F(2, 19) = 0.70$, $p = 0.51$. There was no interaction effect of two variables, Wilks’ Lambda = 0.96, $F(2, 19) = 0.43$, $p = 0.66$. Post hoc univariate analysis, further, revealed that task inclusion was statistically significant on the posttest, $F(1, 20) = 9.26$, $p < 0.05$, $R^2 = 33\%$, but not on the delayed posttest, $F(1, 20) = 3.88$, $p = 0.06$, $R^2 = 16\%$.

The following Figure 1 graphically illustrates, regarding the gain scores of posttest from pretest, that the two Confidence Intervals (CIs) for the DG and CG groups substantially overlaps with each other, presenting that the effect of two types of glosses is not statistically significant. To the contrary, there is no overlapping of the two CIs for the +T and –T groups, clearly indicating that there is statistical significance between the two groups. With regard to the gain scores of delayed posttest from pretest, Figure 2 suggests that there are considerable portions of overlapping of the two CIs for the DG and CG groups and for +T and –T groups respectively.
This is to say that the effect of glosses is not statistically significant. The same interpretation applies to the groups of +T and –T.

**Figure 1.** 95% CI for glosses and a task on the gain scores of posttest

**Figure 2.** 95% CI for glosses and a task on the gain scores of delayed posttest
DISCUSSION

Glosses and Vocabulary Acquisition

Many researchers already provided the positive effect of glosses on L2 vocabulary acquisition in their empirical studies (Hulstijn, 1992; Hulstijn et al., 1996; Rott et al., 2002; Watanabe, 1997). Thus, with the above premise in mind, this study did not employ a control group wherein the passage had no gloss. Rather, the study attempted to further investigate significant effects of two different types of glosses (DG and CG) on vocabulary acquisition in terms of initial learning and retention. In addition of the empirical findings, the depth of processing hypothesis (Craik & Tulving, 1975) and later mental hypothesis (Hulstijn, 1992) initially guided the following supposition: CG will promote vocabulary learning better than DG because CG entails deeper mental processing in order to infer the meaning of target words with the help of contextual clues.

This study adopted CG as one of devices like MC glosses in the hypothetical respect of profoundly activating the students’ mental efforts to infer the meaning of unknown words and of consequently resulting in better vocabulary learning and retention. Contrary to the expectation, in fact, there was no statistical significance between two types of glosses on vocabulary learning. In the observed means on the gain scores of posttest and of delayed posttest, the treatment groups with DG scored better than the ones with CG. This result totally runs to counter to the findings by Hulstijn (1992) and gets in similar line with the ones by Watanabe (1997). Unlike Hulstijn, Watanabe reported in his study that there was no statistically significant difference between
single and multiple-choice (MC) glosses. One possible reason assumed by him was that the students chose the wrong alternatives from the multiple choices. His assumption was provided with the evidence that the students in the MC gloss condition scored much lower than those in the gloss condition. Similar to the explanation by Watanabe, the study assumed that the sentences prepared for CG failed to provide sufficient contextual clues enough to infer the correct meaning of the words. In other words, the students either guessed the incorrect meaning of the words or simply failed to guess their meaning.

A task and Vocabulary Acquisition

The main effect of a fill-in-the-blank task reached statistical significance on initial vocabulary learning, but not on retention of the learned words. This finding can be explained with the proposal of task-induced involvement by Laufer and Hulstijn (2001): constructs of need, search, and evaluation are conditional on retention of vocabulary learning. The fill-in-the-blank task induced need construct with respect to triggering the necessity of knowing the meaning of unknown words to accomplish the task and evaluation construct with respect to deciding whether or not the inferred meaning of the words fit into the context. However, the fill-in-the-blank task failed to exert its influence on delayed posttest, suggesting that the task was not powerful enough to substantially induce motivational and/or cognitive constructs resulting in better retention of the learned words.

CONCLUSION
Summary of the Findings

The findings of the study are summarized as follows: (a) the students who read the passage with DG considerably performed better on vocabulary posttest and delayed posttest than those who read the same passage with CG, but there was no statistical significance between two types of glosses on L2 vocabulary acquisition; (b) the students who conducted a fill-in-the-blank task after reading the passage scored better on vocabulary posttest and delayed posttest than those who did not practice the task after reading, but statistical significance of the task did not apply to delayed posttest; (c) there was no statistical interaction effect of glosses and a fill-in-the-blank task on L2 vocabulary acquisition.

Pedagogical Implications

The findings of the study might enhance the understanding of L2 vocabulary acquisition and accordingly facilitate to develop materials or activities especially regarding glosses and tasks. With the assumption that inclusion of glosses themselves augment the potentialities of learning vocabulary and retention, there are still many rooms for adopting or adapting glosses while considering such types as definitional, contextual, or pictorial glosses, such numbers as single or multiple-choice glosses, or such language as first language (L1) or L2 glosses.

Limitations
The main drawback of this study was a small number of participants. In other words, performance of each participant might be highly likely to skew the results of the study. Given that the large number of students participated in the study, the study might have different results. Also, the fact that the participants conducted an initial posttest after a pretest on the same day can bring out the issue of frequency effect of words. At least three interval sessions with a pretest, initial posttest, and delayed posttest could draw more reliable data and even a longitudinal study would be the best recommended. The criteria of constructing DG or CG can be controversial as well. Some target words were provided with only single word and some were with more than one word for their definitions in DG. That is, the number of definitional words in DG could have a slight impact on vocabulary learning. Also, the participants might have been unable to understand its definition rather than the target word itself. Similarly, the study did not prove to adjust the power degree of contextual clues in CG; thus, some sentences presented could be less helpful than others to infer the correct meaning of the target words and vice versa. Lastly, the study did not consider the level of English proficiency of the participants. Even though the study tried to investigate the depth of vocabulary learning of each student regardless of their English proficiency level, it would be more interesting to be capable of exploring distinctive features of learning vocabulary according to the level of English proficiency.

REFERENCES


Nagy, W. & Herman, P. (1987). Breadth and depth of vocabulary knowledge: Implications for acquisition and instruction. In M.G. McKeown and M. Curtis (Eds.), *The Nature of*


APPENDIX A
Sample of reading passage

<Reading Passage with DG>

When trying to convince people that their behavior is greatly influenced by the social environment, the social psychologist is up against a formidable barrier. This barrier is known as the fundamental attribution error. It is the tendency to explain our own and other people’s behavior totally in terms of personality traits. formidable: fearful

<Reading Passage with CG>
When trying to convince people that their behavior is greatly influenced by the social environment, the social psychologist is up against a formidable barrier. This barrier is known as the fundamental attribution error. It is the tendency to explain our own and other people’s behavior totally in terms of personality traits.

A pair of scissors can be a formidable weapon to kill someone.

APPENDIX B
Sample of VKS

► Directions: Please check any of those which apply to your vocabulary knowledge scale and fill in the blanks, if applicable.

1. ____________________
   I. I don’t remember having seen this word before.
   II. I have seen this word before but I don’t know what it means.
   III. I have seen this word before and I think it means ____________________. (Either L1 or L2)
   IV. I know this word. It means ____________________. (Either L1 or L2)
   V. I can use this word in a sentence in English: ____________________________.
      (if you do this section, please also do section IV)

APPENDIX C
Sample of fill-in-the-blank task

► Direction: Please fill in the blanks using the best appropriate words.

1. In spite of his ____________________________, he did not quit his job and was trying to take his share in the work.

2. The notes help to ____________________________ the difficult parts of the text.