

The effect of annotation types on incidental vocabulary learning: The results of vocabulary tests and questionnaires

by
Makoto Yoshii

INTRODUCTION

The use of the technology into language learning and teaching has been increasing with ever-accelerating speed. Our intuition speaks of the great advantage of having multimedia information for teaching, and some studies have pointed us in that direction (Otto and Pusack, 1993; Ruhlmann, 1995). However, we still do not fully understand what kind of multimedia information is most effective for language learning and in what areas of language learning it might be most effective (Bush, 1997: 342). In this study, a special focus is given to one such area, vocabulary learning. Lately, vocabulary learning has been regaining much needed attention (Zimmerman 1997: 5) and the use of technology has been examined for the enhancement for vocabulary learning (Brett, 1998; Duquette et. al, 1998; Ellis, N. C., 1995; Goodfellow, 1995). This research reported herein belongs to this line of studies and looks at incidental vocabulary learning in particular. The study examines how Internet-based reading materials and the use of glosses may enhance incidental vocabulary learning, and how learners perceive reading on computers and the use of glosses.

LITERATURE REVIEW

Vocabulary learning is a crucial element of language learning, and is considered to be a high priority in second language learning and teaching (Candlin, 1988; Crow, 1986; Knight, 1994). There are four major approaches to vocabulary instruction found in the literature (Coady, 1997): 1) context alone (vocabulary development through extensive reading); 2) strategy instruction (strategy training on how to learn from context effectively); 3) development plus explicit instruction (explicit teaching of certain vocabulary); and 4) classroom activities (emphasis on classroom activities of certain vocabulary). Particular attention is given to the "context alone" approach in the present study.

The context alone approach is based on the claim that vocabulary is learned from context through extensive reading (Krashen, 1982, 1989, 1993). It is also called "incidental vocabulary learning" since the main focus of learners is on comprehension, and not on intentional vocabulary retention. Although the efficiency of incidental L2 vocabulary learning has been questioned from time to time (Dubin & Olshtain, 1993; Hulstijn, 1992; Hulstijn, et al., 1996), many agree on the importance of contextual, incidental vocabulary learning (Carter & McCarthy, 1988; Chall, 1987; Qian, 1996). What's more, the effectiveness of incidental vocabulary learning can be enhanced through a variety of methods. One such highly effective method is the use of glosses (Hulstijn, et al., 1996). Glosses are definitions or explanations for unknown or difficult words which appear in the margins. In this study, the terms "glosses" and "annotations" are used interchangeably.

Studies have shown that incidental L2 vocabulary learning is enhanced by the use of glosses in printed materials. For example, Jacobs, Dufon, and Fong (1994) examined the effect of annotations on L2 comprehension recall and L2 vocabulary acquisition. Eighty-five English-speaking university students who enrolled in a fourth semester Spanish course participated in the study. The results showed that the learners under gloss conditions (either L1 or L2) performed better on the immediate vocabulary tests than those who had no gloss. Hulstijn, Hollander & Greidanus (1996) investigated the influence of L1 glosses for incidental vocabulary learning. Seventy-eight Dutch college students who were studying advanced French participated in the study. The results showed that the gloss condition was better than no gloss condition, and that the effect of gloss was better than the use of a dictionary. Watanabe (1997) conducted a study with 231 Japanese university students who were studying English as a foreign language. The results showed that the glosses were more effective than no glosses or appositives. As these studies showed, the provision of glosses does enhance incidental vocabulary learning. Therefore, the question has shifted from whether the glosses help vocabulary learning to what kind of gloss is most effective.

Jacobs, Dufon, and Fong (1994) examined two types of glosses in their study: L1 gloss (English) and L2 (Spanish) gloss. No significant difference was found between the two gloss conditions. The questionnaire revealed that the participants preferred L2 glosses when the glosses were within the students' level of comprehension. The researchers, however, cautioned that those who favored L2 glosses were more proficient in the target language since they were the ones who best comprehended the content of the L2 cues. Watanabe (1997) compared two types of glosses as well: single gloss and multiple-choice gloss. The results showed that there was no significant difference between the two glosses. These studies used textual cues for the glosses. There are also studies using not just textual cues but also visual cues such as pictures. For example, Kost, Foss, & Lenzini (1999) examined three types of annotations: 1) text-only (L1); 2) picture-only; and 3) text and picture cues. A group of 56 American university students who were studying second-semester German participated in the study. The results showed that the combination of text and picture annotation was the most effective among the three types for both immediate and delayed tests.

Annotation studies have also been conducted with computer-based reading texts. Lyman-Hager, Davis, Burnett, & Chennault (1993) examined L2 reading comprehension and vocabulary learning comparing computerized reading with non-computerized reading. The results of a study involving a sample of 262 American university students who were studying intermediate French revealed that the computerized group performed significantly better on the vocabulary retention test than the non-computerized group. Chun & Plass (1996) conducted a series of annotation studies as well. They examined three types of annotation using computer-based reading material—1) text-only; 2) text and picture; and 3) text and video—using 160 American university students of second-year German. The findings revealed that the text and picture combination was the most effective. Plass, Chun, Mayer, & Leutner (1998) continued their examination of the three gloss types with 103 American university students studying second-year German. As in their 1996 study, they found that the text and picture annotation to have the greatest effect. More annotation-type comparisons were made by Al-Seghayer (2001). In this study of text-only, text-and-picture, and text-and-video annotations, the text-and-video annotations produced the most powerful effect of the three, which was contrary to the earlier findings reported (Chun and Plass,

1996; Plass et al., 1998). Additional work by Nagata (1999) examined single- and multiple-choice glosses in a computer-based reading text. Implementing Watanabe (1997)'s research design, she found that the multiple-choice group performed significantly better than the single-gloss group, thereby providing conflicting evidence against previous results (Watanabe, 1997).

NEED FOR THE STUDY

As seen above, annotation studies are a relatively recent phenomenon, and one marked by contradictory findings which demand the further investigation. Following Nagata's example, the results of the studies using printed texts need to be confirmed using computerized texts. Moreover, since most of the studies to date have been limited to homogeneous populations of students with respect to native language and to the intermediate or advanced levels of either French, German, or Spanish, more studies are warranted in which the performance of linguistically heterogeneous groups of ESL learners at the lower proficiency levels are examined. In addition, as many studies have been conducted using a within-subjects design, there is a need for a study using a between-subjects design.

There is also a need for more qualitative data to give a fuller understanding of the effect of annotations. Kost et al. (1999) acknowledge this and recommend further exploration of qualitative aspects. Chun and Plass (1997) mention the importance of investigating the individual differences and their role in their effect on multimedia learning. They list factors to be studied further including affective factors such as motivation, anxiety, interest, etc. (1997:17). Just as Al-Seghayer (2001), in the study of multimedia annotations, included such qualitative data collected through a face-to-face interview and questionnaires, studies supported by both quantitative and qualitative data are needed to understand the effect of annotations better.

THE STUDY

Purpose

The purpose of this study was to examine the effect of annotation types on incidental vocabulary learning using both quantitative and qualitative data. Both vocabulary retention and students' perceptions of different annotation types were analyzed. The study transformed Kost et al.'s (1999) research design into web-based text. It employed a between-subjects design in a multi-lingual setting (ESL learners with different L1s) and considered two novice proficiency levels (beginning and intermediate). L2 cues were used exclusively.

Participants

The participants were 151 ESL learners in beginning and intermediate levels who were studying English in intensive English programs at universities in Florida. There were

69 female and 82 male students representing 38 countries and 18 languages. Their average age was 24.6, ranging from 16 to 47 years old.

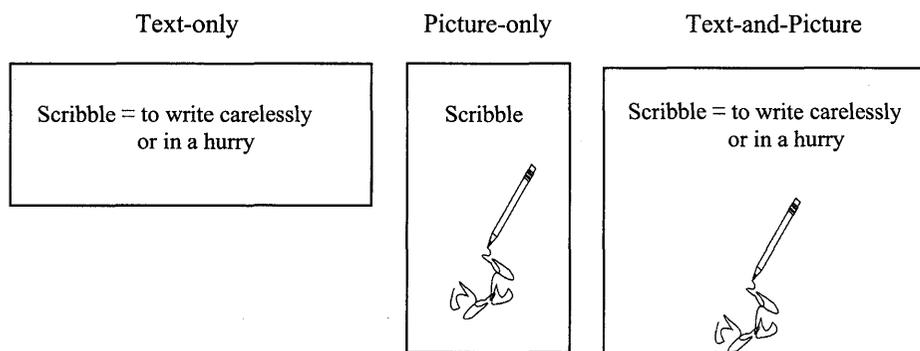
Research Questions

For the quantitative investigation, the study posed three questions regarding the effect of annotation type on vocabulary retention, namely: What are the effects of annotation types on 1) short-term vocabulary retention, 2) long-term vocabulary retention, and 3) vocabulary recall over time? For the qualitative inquiry, the study posed two main questions: 1) what are the students' previous experiences and familiarity with computers? and 2) what are the students' perceptions of the glosses and web-based reading activity?

METHOD

The participants were first stratified according to their L1s and proficiency levels. Then, they were randomly assigned to one of the three annotation conditions: 1) text-only, 2) picture-only, or 3) text and picture. Examples of the annotations can be seen in Figure 1:

Figure 1. Examples of three different annotations



A pre-questionnaire was given to the students before the reading activity and queried them about their previous experiences with and attitudes toward computers. An example of the pre-questionnaire can be found in Appendix A. After filling out the pre-questionnaire, the participants individually read a story on the web for comprehension and took a reading comprehension test. The story can be found in Appendix B. There were 14 target words (all verbs) in the story and they were highlighted so that when the students clicked on the words, they were able to see the annotations on the right side of the screen. Then, they were unexpectedly asked to take vocabulary tests in a variety of formats: definition-supply, picture recognition, and word recognition. The definition-supply test asked students to supply the meaning of the word in either L1 or L2. The picture recognition test required the students to identify from the four drawings the one that best conveyed the meaning of the target word. For the word recognition test students were to choose the best definition out of four choices. Examples of these test items can be found in

Appendix C. Lastly, the participants were asked to fill out the post-questionnaire about the web-reading activity. (See Appendix D.) The whole procedure was to take place in two consecutive 50-minute classes.

RESULTS AND DISCUSSION

The results of the vocabulary tests are summarized below followed by the findings from the two questionnaires.

1. Vocabulary tests

There were two sets of tests used to measure vocabulary retention— immediate tests and delayed tests (two weeks later). Each set consisted of three different types of tests: definition-supply, picture recognition, and word recognition. Fourteen target words were tested in each test type. The results of the definition-supply tests were scored two ways: strict and lenient. For the definition-supply tests, one point was given to a partially correct answer (lenient) and two points were given for the fully correct answer (strict). For both the picture and word recognition tests, one point was given for each correct answer. The results of the immediate tests are seen in Table 1.

Table 1. Means and Standard Deviations of the Immediate Tests by Group

Group	n	Definition-Supply				Picture Recognition		Word Recognition	
		(Strict)		(Lenient)		M	SD	M	SD
Combination	50	1.86	1.80	3.52	2.24	7.46	2.53	7.58	2.6
Picture-only	51	0.96	1.39	3.00	2.39	7.18	2.98	6.41	2.87
Text-only	50	1.38	1.63	2.62	2.46	5.98	2.48	6.12	3.05

Across the tests, the Combination of text and picture annotation group outperformed the other two groups. Significant group differences were found in the Definition-Supply under strict scoring, $F(2, 145) = 3.93, p = 0.0218$, in the Picture Recognition, $F(2, 145) = 4.04, p = 0.0197 < .05$, and in the Word Recognition, $F(2, 145) = 3.28, p = 0.0406 < .05$.

The results of the delayed tests also revealed the overall advantage of the Combination group across the tests except for the Definition-Supply under lenient scoring seen in Table 2.

Table 2. Means and Standard Deviations of the Delayed Tests by Group

Group	n	Definition-Supply				Picture Recognition		Word Recognition	
		(Strict)		(Lenient)		M	SD	M	SD
Combination	50	1.14	1.63	1.98	2.20	6.48	2.67	6.06	3.11
Picture-only	51	0.82	0.95	2.21	2.13	6.31	3.08	5.43	3.23
Text-only	50	0.68	1.04	1.68	1.61	4.92	2.78	4.62	2.42

However, significant group differences were found in the Picture Recognition test only, $F(2, 145) = 4.17, p = 0.0174$, indicating that the group differences became smaller after two weeks. As Tables 1 & 2 indicate, the students' performance on the vocabulary tests declined after two weeks. Table 3 indicates how much vocabulary the learners were able to retain over time.

Table 3. The Retention Rate by Group on the Tests

Group	Definition-Supply (Strict)	Definition-Supply (Lenient)	Picture Recognition	Word Recognition
Combination	61%	44%	87%	80%
Picture-only	85%	74%	88%	85%
Text-only	49%	64%	82%	75%

Although the learners were able to recognize 75-88% of the words they learned (Recognition tests), their recall rate dropped considerably for the production test (Definition-Supply). The Combination group failed to display any particular advantage in recalling what they learned, and, in the case of the Definition-Supply (Lenient) group, they were the ones who declined the most. It was also interesting to find that the Picture-only group was able to sustain around 80% of the words they learned throughout the tests. Their retention rate was the most consistent among the three groups.

Overall, the text-and-picture annotation was the most effective for the immediate and delayed tests. The Picture-only gloss was more effective than the Text-only gloss even though the differences were not significant. The Text-only annotation was the least effective among the three. The details of the vocabulary tests can be found in the study (Yoshii & Flaitz, 2002).

2. Questionnaires

In order to understand the results better, the data gathered from the two questionnaires were examined: one was a pre-questionnaire asking about the students' previous experiences with computers, and the other was a post-questionnaire inquiring into participants' perceptions toward the computer-based reading activity and the use of glosses.

1) Pre-questionnaire

There were three sets of questions in the pre-questionnaire: 1) general questions about students' attitudes toward computers, 2) questions about students' comfort level in using computers, and 3) questions about students' actual daily use of computers. Under the first category, there were three questions: 1a) Do you like to use computers? 1b) Are you good at using computers? and 1c) Do you feel comfortable using computers? In the second category the following three questions were asked: 2a) Is it easy to use computers? 2b) Is it easy to use a "mouse"? and 2c) Is it easy to read something on computers? An additional three questions were asked in the final set: 3a) How many times a week do you use

computers? 3b) How many times a week do you use the Internet? and 3c) How many times in a week do you use email? The format was multiple-choice, and the responses were tallied and converted into percentages. Table 4 provides a summary of the answers to these questions.

Table 4. Summary of the Results of the Pre-questionnaire

a) General questions about students' attitudes toward computes

Question 1a): Do you like to use computers?

	Very much	Yes	Not very much	Not at all
Combination	46%	46%	8%	0%
Picture-only	49%	43%	6%	2%
Text-only	44%	52%	4%	0%

Question 1b): Are you good at using computers?

	Very much	Yes	Not very much	Not at all
Combination	10%	40%	40%	10%
Picture-only	27%	41%	27%	4%
Text-only	20%	38%	40%	2%

Question 1c): Do you feel comfortable using computers?

	Very much	Yes	Not very much	Not at all
Combination	22%	56%	20%	2%
Picture-only	40%	46%	8%	6%
Text-only	26%	54%	18%	2%

b) Questions about students' ease in using computers

Question 2a): Is it easy to use computers?

	Very much	Yes	Not very much	Not at all
Combination	16%	36%	42%	6%
Picture-only	22%	54%	16%	8%
Text-only	14%	46%	38%	2%

Question 2b): Is it easy to use a "mouse"?

	Very much	Yes	Not very much	Not at all
Combination	42%	50%	8%	0%
Picture-only	59%	39%	2%	0%
Text-only	50%	48%	2%	0%

Question 2c): Is it easy to read something on a computer?

	Very much	Yes	Not very much	Not at all
Combination	26%	30%	34%	10%
Picture-only	25%	45%	24%	6%
Text-only	18%	50%	28%	4%

c) Questions about students' actual daily use of computers

Question 3a): How many times a week do you use computers?

%	Almost every day	3~7 times	1~3 times	Never
Combination	46%	20%	30%	4%
Picture-only	58%	24%	16%	2%
Text-only	44%	38%	18%	0%

Question 3b) How many times a week do you use the Internet?

%	Almost every day	3~7 times	1~3 times	Never
Combination	42%	22%	24%	12%
Picture-only	60%	24%	10%	6%
Text-only	34%	34%	30%	2%

Question 3c): How many times a week do you use email?

%	Almost every day	3~7 times	1~3 times	Never
Combination	38%	28%	22%	12%
Picture-only	60%	22%	16%	2%
Text-only	50%	28%	20%	2%

Overall, the Picture-only group expressed greater familiarity, more experience, and higher comfort level with computers. A few differences were found between the Text-only and the Combination groups in terms of their familiarity, experience, and comfort level with computers. For example, for Question 1b), 27% of the Picture-only group felt they were very good at using computers. On the other hand, 20% of the Text-only group responded positively, and the Combination group responded the least confidently with a percentage of 10%. For Question 1c), 40 % of the Picture-only group felt comfortable using computers very much, while 26% of the Text-only group and 22% of the Combination group felt very comfortable. Twenty-two percent of the Picture only group felt it very easy to use computers (Question 2a), while 14% of the Text-only and 16% of the Combination group answered the same way. For Question 2a), 54% of the Picture-only group also answered "yes" to the question. This brings the total percentage of students in the group who felt positive about the use of computers to 76% (22% of "Very much" plus 54% of "Yes") while the Text-only and the Combination group were 60% positive and 52% positive respectively. As for the question asking whether students considered reading something on the computer to be easy (2c), 70% of the Picture-only group answered either "Very much" or "Yes," and 68% of the Text-only group answered favorably, as did 56% of the

Combination group. The actual daily use reflected in responses to Questions 3a), 3b), and 3c) also indicated that the Picture-only group tended to use computers more often than the other groups.

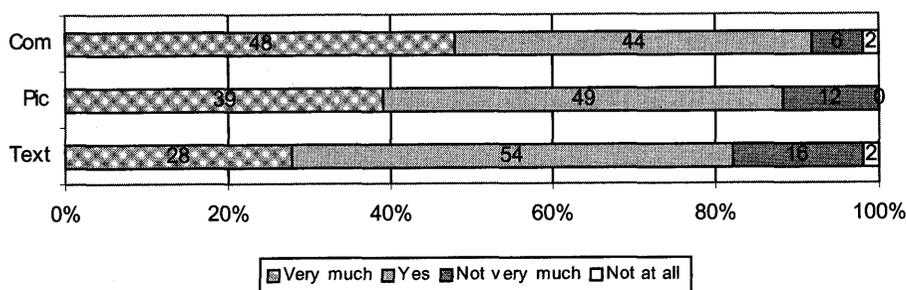
As seen in these results, even though the students were randomly assigned to three groups, the students in the Picture-only group had more familiarity and experience with computers than the other two groups had. There was little difference between the Text-only and the Combination groups although there were some indications that the Combination group had the least. With these results in mind, the fact that the Combination group outperformed the other two on the vocabulary tests seems more convincing.

2) Post-questionnaire

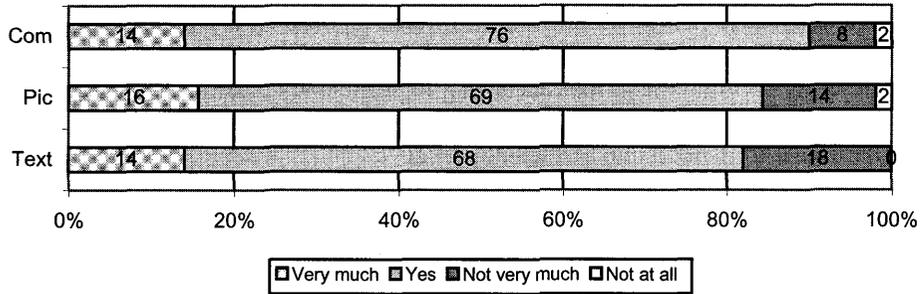
In order to look at the effect of annotation types from another angle, the students' perceptions toward the reading activity on the web and the use of glosses were examined through the post-questionnaire. There were two sets of questions in the questionnaire: 1) questions about the story and reading on computers and 2) questions about vocabulary and glosses. Under the first set, there were 4 questions: 1a) Did you understand the story? 1b) Did you like the story? 1c) Did you like reading on a computer? and 1d) Would you like to read more stories on the computer? Under the second set, there were also 4 questions: 2a) Were the words in the story easy? 2b) Did you understand the glosses? 2c) Did you like the glosses? and 2d) Were the glosses helpful for you? The format was multiple-choice, and, once again, the responses were tallied and converted into percentages. When a large difference (i.e. of more than 20%) was detected in a response among the groups, the data were further examined under two proficiency levels. Therefore, there were three annotation groups and two proficiency levels making up 6 cell groups. The number of each cell is as follows: the Combination-beginning (n = 22), the Combination-intermediate (n = 28), the Picture-beginning (n = 25), the Picture-intermediate (n = 26), the Text-beginning (n = 26), and the Text-intermediate (n = 24). The questionnaire also included an open-ended section for students' comments. The students were asked to write their opinions about the reading activity on the Internet.

The first set of questions inquired the students' perceptions toward the story and reading activity on computers.

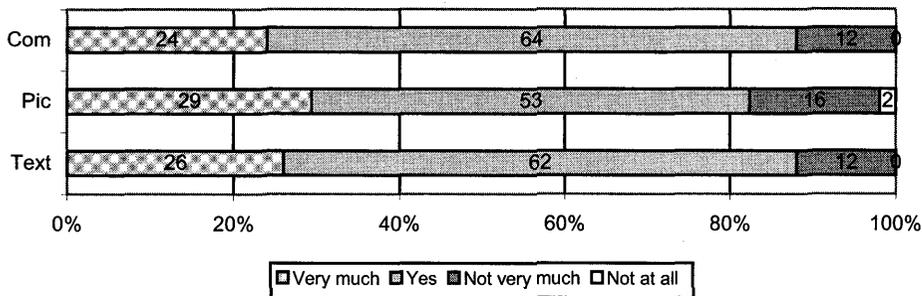
Figure 2. Post-questionnaire results: Perceptions toward the story and reading activity
Question 1a) Did you understand the story?



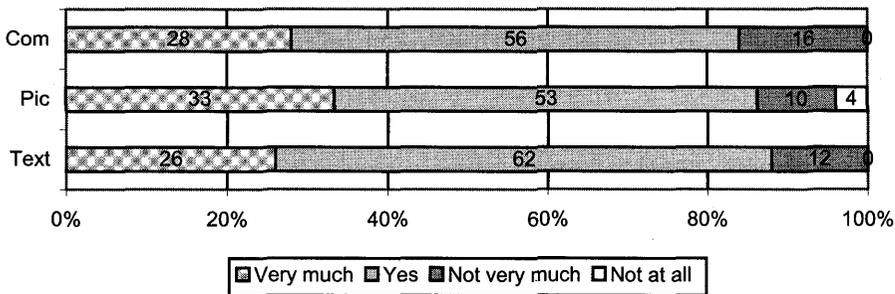
Question 1b) Did you like the story?



Question 1c) Did you like reading a story on the computer?



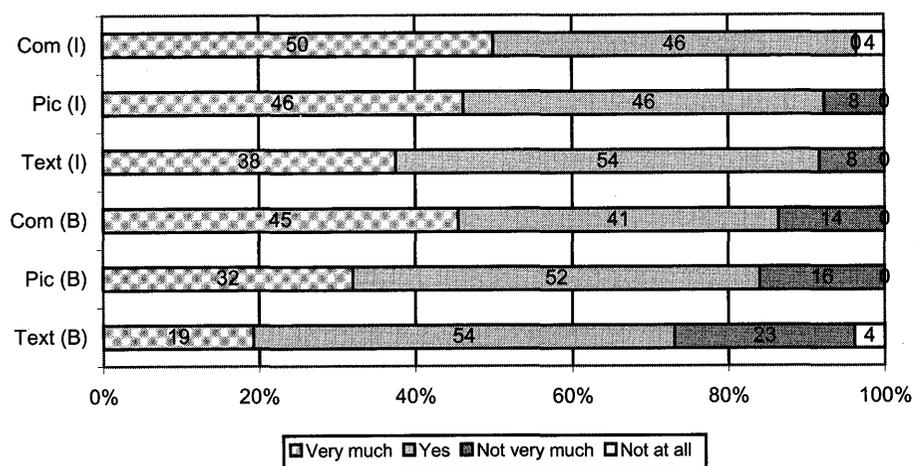
Question 1d) Would you like to read more stories on the computer?



As seen in Question 1b), 1c), and 1d), there was not much difference among the groups in terms of the students' liking for the story and reading activity on computers. Majority of the students (more than 80%) liked the story and the reading activity and responded positively—either “Very much” or “Yes” to these questions.

As for understanding the story seen in Question 1a), the combination group excelled other two groups indicating that the combination gloss seemed to help the students understand the story best. Since the difference in the response “Very much” between the Combination (48%) and the Text-only group (28%) was very large (20%), the results were examined further by dividing the groups according to proficiency level.

Figure 3. Question 1 a) divided into proficiency level
Question 1a) Did you understand the story?

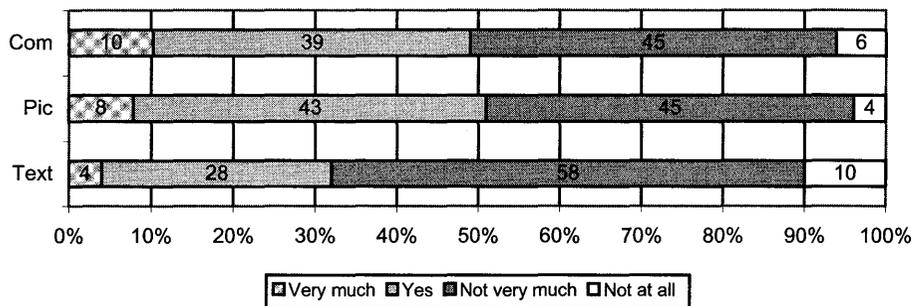


As seen here, in the Beginning level (B), the differences among the groups for the response “Very much” were more evident than those in the Intermediate level (I). These results are a reflection of the students’ perceptions and do not represent their actual reading comprehension. However, the large differences seen here seem worthy of further investigation into the effect of annotations on incidental vocabulary learning. Combination annotations may enhance reading comprehension, and they may have more impact on beginning level students than on intermediate students.

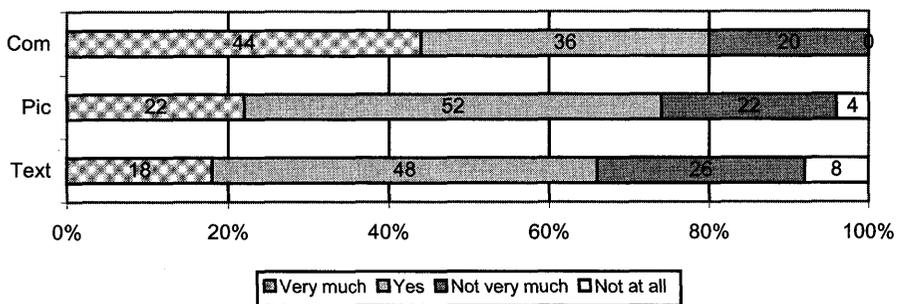
The second set of questions was asked to find out the students’ perceptions toward the vocabulary in the story and the glosses.

Figure 4. Post-questionnaire results: Perceptions toward vocabulary and glosses

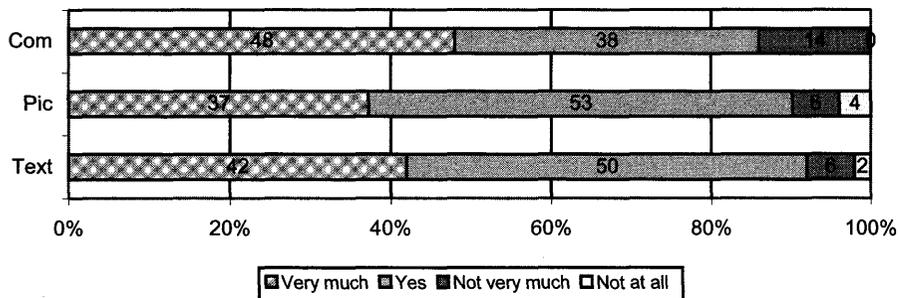
Question 2a) Were the words in the story easy?



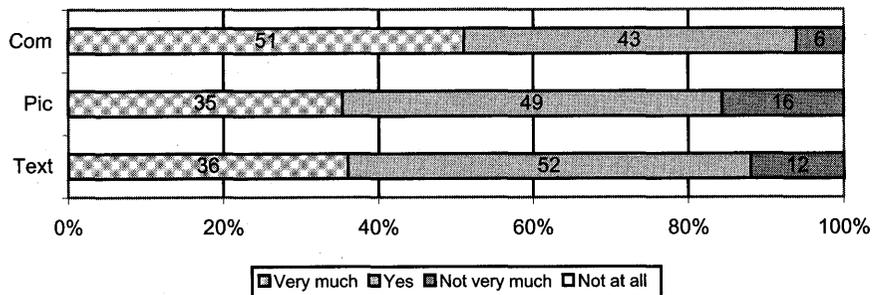
Question 2b) Did you understand the glosses?



Question 2c) Did you like the glosses?



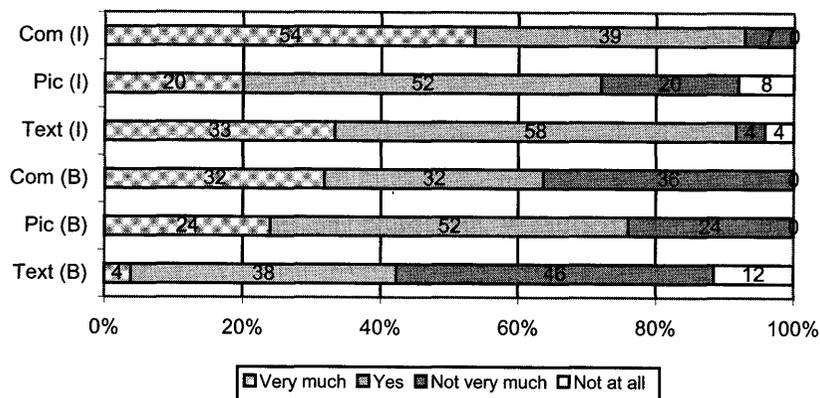
Question 2d) Were the glosses helpful for you?



As seen in Question 2a), the text-only group did not feel the words in the story easy as much as the other two groups felt. The students' pre-knowledge of the target words had been examined through a pretest. The pretest scores did not show any significant differences among the groups, $F(2, 145) = 0.12, p = 0.8908 > .05$. The appropriateness of the word difficulty level for the novice learners had also been examined with a set of novice learners in a pilot study. Even though students' overall pre-knowledge of the words in the story did not seem to differ, there were differences in the responses. This might be related to the gloss types which may influence the students' perceptions and feelings about the difficulty of the words. For example, the Text-only group had to consult the L2 glosses for the difficult words. Even though the glosses were written in simple English for the purpose of aiding the students' understanding of the words, it might have been a burden for them to read and understand the explanations in the L2. This might have left the students in this group with a sense that the words in the story were difficult. The other groups, Picture-only and Combination, had the advantage of access to other types of help such as the pictorial cues, which may have eased the burden of understanding the L2 glosses.

Question 2b), 2c), and 2d) asked the students' perceptions toward the different gloss types. As seen in Question 2 c), there was not much difference among the groups in terms of liking for glosses. More than 80% of the students in each group were in favor of the glosses. This shows the high acceptance and liking of having the glosses no matter what the types were among the students. This is a clear indication of the importance of the provision of glosses for the novice levels of ESL learners. Not only liking for the glosses, the students felt the glosses helpful for the reading activity as seen in Question 2d). Majority of the students said the glosses were either helpful or very helpful. The combination group expressed the most positive response: about the half of the students responded with "Very much." This is another indication that the combination glosses are the most helpful among the three types. This sense of helpfulness may come from how understandable the glosses were as seen in 2b). This shows that the Combination annotation was the most understandable to the students by far. The Text-only annotation was the least clear among the students. Since the difference between the Combination (44%) and the Text-only (18%) for the response "Very much" was 26% exceeding the 20% rule, the results were further examined according to proficiency levels.

Figure 5. Question 2b) divided into proficiency level
 Question 2b) Did you understand the glosses?



In both levels the Combination group had the highest percentage for the response “Very much.” Among the Beginners, 32% of the Combination group answered “very much” and 24% of the Picture-only responded similarly. On the other hand, only 4% of the Text-only group seemed to have understood the glosses. For the overall positive response, the Picture-only had the largest (76%) followed by the Combination group (64%) and the Text-only (42%). This shows how difficult the text-only cues are for the beginning level of students and how helpful the pictorial and combination type cues are for these learners. For the Intermediate level, the Combination group had the highest percentage of “Very much” (54%) followed by the Text-only (33%). The Picture-only had the least of 20%. This indicates that regardless of proficiency level, the Combination was the clearest among the annotation types. This also showed that when the level went up, the percentages also increased.

The effectiveness of the pictorial cues stayed the same across the levels. The pictorial cues were quite effective among the beginners with the responses “Very much” with 24% and “Yes” with 52%, which amounts to 76% of the positive response. The percentages did not differ much for the Intermediate students. They had 20% of “Very much” and 52% of “Yes” making the overall positive response of 72%.

On the other hand, the effectiveness of the textual annotation differed according to the proficiency level. Textual cues were not clear among the beginning level students as seen in the responses “Very much” with only 4% and “Yes” with 38% making the overall positive response of 42%. However, this ineffectiveness changes for the Intermediate students. Among the Intermediate, the percentage of “Very much” jumped to 33% and the percentage of “Yes” also went up to 58% making the overall positive response of 91% which was about the same as the Combination (93%). This indicates how well the Intermediate students were able to understand the L2 glosses while the beginners suffered from not being able to understand the textual cues. This is another indication of the difficulty of the L2 text-only annotations for the beginners and the need for the additional cues such as the visual cues.

3) Results of the Students' Comments

In addition to the multiple-choice format questions, the post-questionnaire also included an open-ended question inviting the students' to comment on the web-based reading activity. A brief summary of the patterns and content that emerged in these comments follows.

About glosses

The students felt that the glosses were helpful for learning new words and for understanding the story. Some specifically mentioned the picture glosses as being particularly helpful.

About web-based reading activity

The students reported that the web-based reading activity was fun, enjoyable, and interesting. Some suggested that they could read the computer-based text faster than they could read the printed text on paper. Others felt that the web text made the reading easier (probably due to the ease and instantaneousness of access to the glosses).

About the activity as a whole

Many students wrote that they enjoyed the activity, believed it was useful, and wanted to have more. Some, however, expressed their anxiety about working with computers.

About learning words

Many of the comments focused on the activity's benefit in terms of vocabulary learning. Nevertheless, a few students lamented the fact that there was insufficient time to learn those words well, and some were overwhelmed by the sheer number of new words introduced in a single activity.

CONCLUSIONS

The results of both the vocabulary tests and the questionnaires indicate that the Combination annotations (text and picture) were the most effective among the annotation types. The Combination group displayed the highest retention of vocabulary for both the immediate and delayed tests. The pre-questionnaire results showed that the Picture-only group was the one which had the edge in terms of previous experience with computers. With this in mind, the advantages for vocabulary learning experienced by the Combination group become convincing. The post-questionnaire revealed that this group expressed the highest level of satisfaction with the web-based reading activity and had the highest percentage of understanding of the glosses. The post-questionnaire also demonstrated that the majority of the students (often times more than 80% of them) liked this reading activity and understood the story. The majority also expressed enthusiasm for the glosses, claiming them to be helpful.

According to the post-questionnaire, as far as the students' perceptions are concerned, the Picture-only glosses were quite effective (or appealing) for the beginners, and their effectiveness stayed the same for the Intermediate level. The Text-only cues were

rather difficult for the Beginners. However, for the Intermediate students, the Text-only cues were very effective since the students understood the L2 glosses better and expressed a greater level of comfort with the glosses than the Beginning level students. These results indicate the importance of providing other types of cues in addition to textual cues for beginning level learners. It is also important to limit the amount of new words introduced at one time, and to recognize the students' need to encounter new words more than once (Hulstijn et al., 1996). Thus, teachers need to seek and explore a variety of ways to expose students to new words in different ways on multiple occasions.

IMPLICATIONS

While it is quite easy to find reading materials on the web for the high-intermediate and advanced learners, similar materials for beginning ESL learners are much harder to come by. Reading activities of the kind used for the present study offer students glosses that are not only appreciated but also increase incidental vocabulary learning. As the web continues to serve as a very important resource for teaching materials, the use of the glosses should be encouraged. Although textual cues are quite effective for intermediate level students, beginning level students may not feel comfortable with such cues. They may, therefore, not be able to understand well, and consequently may feel frustrated and overwhelmed, especially if they feel "flooded" with new vocabulary. When materials developers integrate glosses into reading programs, the use of combination text plus picture glosses should be encouraged.

DIRECTION FOR FUTURE RESEARCH

This study focused on concrete verbs. There is need for further research focusing on abstract words and other parts of speech. Future investigations might also explore the question of "how much is too much:" in other words, what are the limits on the volume of new vocabulary which can be effectively introduced in a single activity. Another direction for future research lies with those students (10-20% overall) who did not prefer reading on the web, did not care for the glosses, and responded negatively. What exactly accounts for this minority opinion? Still another approach to the incidental learning of vocabulary might expand the qualitative dimension of data collection to include interviews, think-aloud protocols, and other ethnographic practices.

Studies reported by Chun and Plass (1996), Plass et al. (1998), Kost et al. (1999), and Yoshii and Flaitz (2002), have indicated that the combination of text and picture glosses is more effective than either text-only or picture-only. The critical question now is "What kind of combination is most effective?" Future studies must also look at the quality of pictures in the glosses as well as at the balance of textual information and visual cues in the combination types and whether the combination should duplicate the information provided or be supplementary to it. The present study has helped initiate that investigation, but further studies are needed to pave the way for producing more effective reading materials on the web.

Acknowledgement

I would like to express my gratitude to Dr. Linda Spetter of Baiko Gakuin University for her editorial assistance and guidance for this paper.

REFERENCES

- Al-Seghayer, K. (2001). The effect of multimedia annotation modes on L2 vocabulary acquisition: A comparative study. *Language Learning & Technology*, 5 (1), 202-232.
- Brett, P. (1998). Using multimedia: A descriptive investigation of incidental language learning. *Computer Assisted Language Learning*, 11 (2), 179-200.
- Bush, M. D. (1997). Implementing technology for language learning. In M. D. Bush & R. M Terry (Eds.), *Technology-enhanced language learning* (pp. 287-349).
- Candlin, C. N. (1988). Preface. In R. Carter & M. McCarthy (Eds.), *Vocabulary and language teaching*. New York: Longman
- Carter, R., & McCarthy, M. (1988). *Vocabulary and language teaching*. New York: Longman.
- Chall, J. (1987). Two vocabularies for reading: Recognition and meaning. In McKeown, M. & Curtis, M. (Eds.), *The nature of vocabulary acquisition* (pp. 7-18). Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.
- Chun, D. M., & Plass, J. L. (1996). Effects of multimedia annotations on vocabulary acquisition. *The Modern Language Journal*, 80 (2), 183-198.
- Coady, J. (1997). L2 vocabulary acquisition: A synthesis of the research. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition* (pp. 273-290). NY: Cambridge University Press.
- Crow, H. T. (1986). Receptive vocabulary acquisition for reading comprehension. *The Modern Language Journal*, 70, 242-250.
- Dubin, F., & Olshtain, E. (1993). Predicting word meanings from contextual clues: Evidence from L1 readers. In T. Huckin, M. Haynes, & J. Coady (Eds.), *Second language reading and vocabulary learning* (pp. 181-202). Norwood, NJ: Ablex.
- Duquette, L., Renie, D., & Laufer, M. (1998). The evaluation of vocabulary acquisition when learning French as a second language in a multimedia environment. *Computer Assisted Language Learning*, 11 (1), 3-34.
- Ellis, N. C. (1995). The psychology of foreign language vocabulary acquisition: Implications for CALL. *Computer Assisted Language Learning*, 8 (2/3), 103-128.
- Goodfellow, R. (1995). A review of the types of CALL programs for vocabulary instruction. *Computer Assisted Language Learning*, 8 (2/3), 205-226.
- Hulstijn, J. H. (1992). Retention of inferred and given word meanings: Experiments in incidental vocabulary learning. In P. J. Arnaud & H. Bejoint (Eds.), *Vocabulary and applied linguistics* (pp. 113-125). London: Macmillan.
- Hulstijn, J. H., Hollander, M., & Greidanus, T. (1996). Incidental vocabulary learning by advanced foreign language students: The influence of marginal glosses, dictionary use, and reoccurrence of unknown words. *The Modern Language Journal*, 80 (3), 327-339.

- Jacobs, G. M., Dufon, P., & Fong, C. H. (1994). L1 and L2 vocabulary glosses in L2 reading passages: Their effectiveness for increasing comprehension and vocabulary knowledge. *Journal of Research in Reading, 17* (1), 19-28.
- Knight, S. (1994). Dictionary use while reading: The effects on comprehension and vocabulary acquisition for students of different verbal abilities. *The Modern Language Journal, 78* (3) 285-299.
- Kost, C. R., Foss, P., & Lenzini, J. J. (1999). Textual and pictorial glosses: Effectiveness on incidental vocabulary growth when reading in a foreign language. *Foreign Language Annals, 32* (1), 89-113.
- Krashen, S. (1982). *Principles and practice in second language acquisition*. Oxford: Pergamon.
- Krashen, S. (1989). We acquire vocabulary and spelling by reading: Additional evidence for the Input Hypothesis. *The Modern Language Journal, 73*, 440-464.
- Krashen, S. (1993). The case for free voluntary reading. *The Canadian Modern Language Journal, 50*, (1) 72-82.
- Lyman-Hager, M., Davis, J. N., Burnett, J., & Chennault, R. (1993). Une Vie de Boy: Interactive reading in French. In F. L. Borchardt & E. M. T. Johnson (Eds.), *Proceedings of the CALICO 1993 Annual Symposium on "Assessment"* (pp. 93-97). Durham, NC: Duke University.
- Nagata, N. (1999). The effectiveness of computer-assisted interactive glosses. *Foreign Language Annals, 32* (4), 469-479.
- Otto, S. K. & Pusack, J. P. (1993). An introduction to foreign language multimedia: the ten most frequently asked questions. (pp. 55-72). in W. Hatfield, ed., *Visions and Reality in Foreign Language Teaching: Where We Are, Where We Are Going*. Lincolnwood, IL: National Textbook Company.
- Plass, J. L., Chun, D. M., Mayer, R. E., & Leutner, D. (1998). Supporting visual and verbal learning preferences in a second-language multimedia learning environment. *Journal of Educational Psychology, 90*(1), 25-36.
- Qian, D. D. (1996). ESL vocabulary acquisition: Contextualization and decontextualization. *The Canadian Modern Language Review, 53* (1), 120-142.
- Ruhlmann, F. (1995) Towards replacement of the teaching process: the emulation of the teaching process with CAL and its implications for the design of a multimedia CAL tutorial. *Computer Assisted Language Learning, 8* (1), 45-61.
- Watanabe, Y. (1997). Input, intake, and retention: Effects of increased processing on incidental learning of foreign language vocabulary. *Studies in Second Language Acquisition, 19*, 287-307.
- Yoshii, M. & Flaitz, J. (2002). Second language incidental vocabulary learning: The effect of text and picture annotation types. *CALICO Journal 20* (1), 33-58.
- Zimmerman, C. B. (1997). Historical trends in second language vocabulary instruction. In Coady, J., & T. Huckin. (Eds.), *Second language vocabulary acquisition*. (p. 5). Cambridge: Cambridge University Press.

APPENDIX A. PRE-QUESTIONNAIRE

Direction: Please answer these questions.

1. What is your level (or class)? _____
2. Where are you from? _____
3. What is your native language? _____
4. How old are you? _____
5. Are you male or female? _____
6. How long did you study English
in your home country? _____
7. How long have you studied English
in the U. S.? _____

Direction: Please read the questions below and check one [] for each question.

1. Do you like to use computers?
[] Very much [] Yes [] Not very much [] Not at all.
2. Is it easy to use computers?
[] Very much [] Yes [] Not very much [] Not at all.
3. Is it easy to use "mouse"?
[] Very much [] Yes [] Not very much [] Not at all.
4. Is it easy to read something on computers?
[] Very much [] Yes [] Not very much [] Not at all.
5. Are you good at using computers?
[] Very much [] Yes [] Not very much [] Not at all.
6. Do you feel comfortable using computers?
[] Very much [] Yes [] Not very much [] Not at all.
7. How often do you use computers in a week?
[] Almost everyday [] 3~7 times [] once ~ 3 times [] Never
8. How often do you use the Internet in a week?
[] Almost everyday [] 3~7 times [] once ~ 3 times [] Never
9. How often do you use email in a week?
[] Almost everyday [] 3~7 times [] once ~ 3 times [] Never

APPENDIX B. STORY: "A SCARY NIGHT"

It's a cold winter night. It's midnight, and is very quiet. I'm still awake and studying. I have a test tomorrow. I need to read two chapters. I finish one chapter and I read the next chapter. It's too difficult. I can't pass the test. What do I do? Shall I keep studying? Can I take the test some other time? Shall I give up? I'm **pondering** many things. I think my head is going to **burst**.

Suddenly, some noise **startles** me. Something **shattered** on the ground. I look at the window. Wait! What is that? I see a light across the street. It is from a new house. It's strange. Mr. & Mrs. Smith are on vacation now. They asked me to **rake** the lawn for them while they're gone. Nobody should be there. Oh, I see the light again.

Then, it **dawns on** me. Someone is **burglarizing** the house. I'm afraid. What do I do now? I have to call the police. I **dash** to the phone and call the police.

After ten minutes, the police arrive. They enter the house. As the police **search** the house, someone **hides** outside the house. The police **yell**, "Stop, right there!" But the man with a black mask runs into the woods near the house. Then, he **tumbles** down the hill in the woods. The police finally catch him. The police take off the mask. He **grins** first, then, starts to **sob**.

Two policemen come to my apartment. The first one looks very serious. He doesn't **greet** me. He just asks for my name. Then, he says, "Thank you for calling us about this problem." The other one is friendlier. He **inquires** about a couple of things. He wants to know when I first saw the light. He **scribbles** some notes.

The policemen are gone, and everything is quiet now. What a strange night! I'm glad this is over, but I am still **shivering** a little. So I **pour** some milk. This might help me. I can't study any longer and can't sleep right away.

I decide to read a book. I got it at a bookstore yesterday. The title is "American Short Stories." I look at the first chapter. And I **gape** at the title. It says, "My Life as a Burglar" by A Man with a Black Mask.

390 words in total

APPENDIX C. EXAMPLES OF VOCABULARY TEST ITEMS

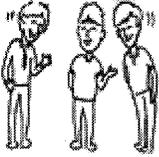
DEFINITION-SUPPLY TEST ITEMS

Directions: Please check any of these words you know. Please put [X] in the box.
Please write the meanings in either in English or your native language.

- [] sob _____
[] shiver _____
[] burst _____
[] grin _____
[] scribble _____

PICTURE RECOGNITION TEST ITEMS

Directions: What does each English word mean? Please choose one matching picture.
Please put [X] in the box.

1. startle [] [] [] []
-    
2. rake [] [] [] []
-    

WORD RECOGNITION TEST ITEMS

Directions: Please match the English word with the correct meaning. Put [X] in the box.

1. ponder [] To study for a test
[] To think very carefully
[] To read something
[] To break open suddenly
-
2. shatter [] To surprise someone
[] To fall suddenly
[] To break something into pieces
[] To look outside

APPENDIX D. POST-QUESTIONNAIRE

Direction: Please read the questions below and check one for each question.

1. Did you understand the story?
 Very much Yes Not very much Not at all.
2. Did you like the story?
 Very much Yes Not very much Not at all.
3. Did you like reading a story in the computer?
 Very much Yes Not very much Not at all.
4. Would you like to read more stories on the computer?
 Very much Yes Not very much Not at all.
5. Were the words in the story easy?
 Very much Yes Not very much Not at all.
6. Did you understand the glosses (help for difficult words)?
 Very much Yes Not very much Not at all.
7. Did you like the glosses?
 Very much Yes Not very much Not at all.
8. Were the glosses helpful for you?
 Very much Yes Not very much Not at all.

What did you think about this reading activity? Could you write your opinions or suggestions?

This is the end of the activities.

Thank you very much for doing these activities.