CHAPTER IV

RESULTS

This chapter presents the results of the analysis of data as described in the methodology chapter. The results of this study are organized in response to the research questions. Firstly, the vocabulary scores from the immediate test of the text-only group, the picture-only group and the text-plus-picture group were analyzed. Second, the vocabulary scores from the three-week delayed test of the three groups were analyzed. Third, the difference in the immediate and delayed test scores among the three groups was analyzed. Finally, the similarities and differences between the current study and other studies were discussed.

After collecting data, the researcher compared the results from the questionnaire and the students' vocabulary scores in their immediate test. Any subjects who claimed that they knew the meaning of any word (check Column Two of the questionnaire) or they recognized seeing the word but were not sure of its meaning (check Column Three of the questionnaire). If the answer was correct, the subjects would be eliminated from the

current study. But, if the answer was incorrect, the
subjects would be included in the data analysis. For
instance, Student A claimed that she knew the meaning of
this word and she could give the correct answer (old = แก่)
(see Figure 4). Thus, this student was eliminated from
the current study.

vocabulary	I have already seen this word and know its meaning.	I have already seen this word but I am not sure of its meaning.	I have never seen this word and do not know its meaning.
old	✓		

Figure 4. An excerpt from a questionnaire of Student A

Student B claimed that she recognized the word but
was not sure about its meaning. If her answer was
correct (old = แก่), she would be eliminated from the
current study. But, if her answer was incorrect, she
would be included in the data analysis (see Figure 5).

vocabulary	I have already seen this word and know its meaning.	I have already seen this word but I am not sure of its meaning.	I have never seen this word and do not know its meaning.
old		✓	

Figure 5. An excerpt from a questionnaire of Student B

Eventually, the numbers of the subjects, whose scores were analyzed in this study, were 128. Group One (6/1) consisted of 41 subjects, group Two (6/2) consisted of 42 subjects, and Group Three consisted of 45 subjects.

Annotations and L2 Vocabulary Immediate Recall

This section answers the first research question, "What are the effects of the three different types of vocabulary annotations -- text-only, picture-only, and text-plus-picture -- on elementary EFL learners' vocabulary test scores in the immediate test?"

Because the independent variable (the text-only group, the picture-only group, and the text-plus-picture group) was nominal data and the dependent variable (the test scores) was ratio data, a one way ANOVA was used to

analyze the three groups' immediate test scores in order to determine if there were any significant differences among the text-only group, the picture-only group, and the text-plus-picture group.

The results of a one-way ANOVA analysis showed that the mean scores and standard deviations of Group One were 13.63 and 7.39; those of Group Two were 8.69 and 7.57; and those of Group Three were 11.67 and 8.15. The minimum score of Group One was 1 and the maximum score was 34. The minimum score of Group Two was 0 and the maximum score was 30. The minimum score of Group Three was 0 and the maximum score was 30 (see Table 2). Among the three groups, the text-only group got the highest mean score, while Group Two got the lowest mean score. As illustrated in Table 2, the difference between the mean scores of the three groups was statistically significant (p=0.015).

Table 2

One-way ANOVA Analysis of the Text-only Group, the Picture-only Group, and the Text-plus-Picture Group in the Immediate Test

	N	Mean	Std. Deviation	F	Sig. (2-tailed)
Group 1 (text-only)	41	13.63	7.39	4.322	.015
Group 2 (picture-only)	42	8.69	7.57		
Group 3 (text-plus-picture)	45	11.67	8.15		

Because the difference of the mean scores was significant, a further analysis of each pair was conducted. The *t* test was used to compare the mean difference between Group One and Group Two, between Group One and Group Three, and between Group Two and Group Three as shown in Table 3, Table 4, and Table 5 respectively.

As illustrated in Table 3, the difference between the scored of Group One and Group Two had a 2-tailed significance level of .003, which was lower than .05 level. This means that the subjects who acquired vocabulary through text-only vocabulary annotations got significantly higher vocabulary scores in the immediate test than the subjects who acquired vocabulary through picture-only vocabulary annotations.

Table 3

Independent Sample *t* test between Group One (Text-only) and Group Two (Picture-only) Immediate Test Scores

	N	Mean	Std. Deviation	Sig. (2-tailed)
Group 1 (text-only)	41	13.63	7.39	.003
Group 2 (picture-only)	42	8.69	7.57	

As illustrated in Table 4, the difference between the scores of Group One and Group Three had a 2-tailed significance level of .246, which was higher than .05 level. In the immediate test, no significant difference was found between the scores of the subjects who acquired vocabulary through text-only vocabulary annotations and the subjects who acquired vocabulary through text-plus-picture vocabulary annotations.

Table 4

Independent Sample *t* test between Group One (Text-only) and Group Three (Text-plus-Picture) Immediate Test Scores

	N	Mean	Std. Deviation	Sig. (2-tailed)
Group 1 (text-only)	41	13.63	7.39	.246
Group 3 (text-plus-picture)	45	11.67	8.15	

As illustrated in Table 5, the difference between the scores of Group Two and Group Three had 2-tailed significance level of .082, which was higher than .05 level. In the immediate test, no significant difference was found between the scores of the subjects who acquired vocabulary through picture-only vocabulary annotations and the subjects who acquired vocabulary through text-plus-picture vocabulary annotations.

Table 5

Independent Sample *t* test between Group Two(Picture-only) and Group Three (Text-plus-Picture) Immediate Test Scores

	N	Mean	Std. Deviation	Sig. (2-tailed)
Group 2 (picture-only)	42	8.69	7.57	.082
Group 3 (text-plus-picture)	45	11.67	8.15	

Hypothesis One states that among the three groups (text-only vocabulary annotations, picture-only vocabulary annotations, and text-plus-picture vocabulary annotations), the subjects who acquire vocabulary through both text and picture vocabulary annotations will get the highest scores on the vocabulary immediate test.

As the differences in the scores between the text and picture groups and both the text only group and the picture only group were not significant, the finding did not support Hypothesis One.

Hypothesis Two states that the subjects who acquire vocabulary through text-only vocabulary annotations will get higher vocabulary scores than the subjects who acquire vocabulary from picture-only vocabulary annotations in the immediate test.

The result of the analysis supported this hypothesis. The vocabulary test scores of the text-only group were significantly higher than that of the picture-only group in the immediate test.

In conclusion, the text-only group got highest mean scores in the immediate vocabulary test and the text-plus-picture group got higher mean scores than the picture-only group; however, the difference was significant only between the text-only group and the picture-only group.

Annotations and L2 Vocabulary Delayed Recall

This section answers the second research question, "What are the effects of the three different types of vocabulary annotations -- text-only, picture-only, and text-plus-picture -- on elementary EFL learners' vocabulary test scores in the three-week delayed test?"

Because the independent variable (the three types of vocabulary annotations) was nominal data and the dependent variable (the test scores) was ratio data, a one way ANOVA was used to analyze the three groups' delayed test scores in order to determine if there were any significant differences among the text-only group, the picture-only group, and the text-plus-picture group.

The results of a one-way ANOVA analysis showed that the mean scores and standard deviations of Group One were 6.63 and 4.41; those of Group Two were 5.55 and 5.58; those of Group Three were 7.51 and 6.35. The minimum score of Group One was 0 and the maximum score was 15. The minimum score of Group Two was 0 and the maximum score was 23. The minimum score of Group Three was 0 and the maximum score was 24 (see Table 6). Among the three groups, the text-plus-picture group got the highest mean score. In addition, the text-only group got a higher mean score than the picture-only group, and the picture-only group got the lowest mean score. As illustrated in Table 6, the mean difference between the three groups was not statistically significant (p=.258). That means no significant difference among the three groups was found. Since the ANOVA result was not statistically significant, no further analysis was needed.

Table 6

One-way ANOVA Analysis of the Text-only Group, the Picture-only Group, and the Text-plus-Picture Group in the three-week Delayed Test

	N	Mean	Std. Deviation	F	Sig. (2-tailed)
Group 1 (text-only)	41	6.63	4.41	1.370	.258
Group 2 (picture-only)	42	5.55	5.58		
Group 3 (text-plus-picture)	45	7.51	6.35		

Hypothesis Three states that among the three groups (text-only vocabulary annotations, picture-only vocabulary annotations, and text-plus-picture vocabulary annotations), the subjects who acquire vocabulary through text-plus-picture vocabulary annotations will get the highest scores on the vocabulary tests given three weeks after the reading task.

Hypothesis Four states that the subjects who acquire vocabulary through text-only vocabulary annotations will get higher vocabulary scores than the subjects who acquire vocabulary from picture-only vocabulary annotations in the three-week delayed test.

Although the text-plus-picture group had the highest mean scores on the delayed vocabulary test, and the mean scores of the text-only group were higher than that of the picture-only group, the differences were not statistically significant. Therefore, the findings do not support Hypotheses Three and Four.

Vocabulary Immediate and Delayed Recall

This section compares the immediate test scores and the delayed test scores of each group in order to find the reduction rate of the subjects' vocabulary scores.

Text-only Group

As shown in Table 7, the result of a paired sample test revealed that the mean scores and standard deviation of the text-only group in the immediate test was 13.63 and 7.39, while the mean scores and standard deviation of their delayed test was 6.63 and 4.41. The mean scores of the text-only group in the immediate test were significantly higher than the mean scores of their delayed test (p=0.000). This shows that the immediate vocabulary test scores of the text-only group were significantly higher than their delayed test scores.

Table 7

A Paired Sample Test of the Text-only Group in Both Tests

		Mean	Std. Deviation	t	Sig. (2-tailed)
Pair 1 (text -only)	Immediate test	13.63	7.39	6.99	.000
	Delayed Test	6.63	4.41		

Picture-only Group

As shown in Table 8, the result of a paired sample test revealed that the mean scores and standard deviation

of the picture-only group in the immediate test was 8.69 and 7.57, while the mean scores and standard deviation of their delayed test was 5.55 and 5.58. The mean scores of the picture-only group in the immediate test were significantly higher than the mean scores of their delayed test (p=0.000). This shows that the immediate vocabulary test scores of the picture-only group were significantly higher than their delayed test scores.

Table 8

A Paired Sample Test of the Picture-only Group in Both Tests

		Mean	Std. Deviation	t	Sig. (2-tailed)
Pair 2 (picture-only)	Immediate test	8.69	7.57	3.88	.000
	Delayed test	5.55	5.58		

Text-plus-Picture Group

As shown in Table 9, the result of a paired sample test revealed that the mean scores and standard deviation of the text-plus-picture group in the immediate test was 11.67 and 8.15, while the mean scores and standard deviation of their delayed test was 7.51 and 6.35. The mean scores of the text-plus-picture group in the

immediate test were significantly higher than the mean

scores of their delayed test (p=0.000). This shows that

the immediate vocabulary test scores of the text-plus-

picture group were significantly higher than their

delayed test scores.

Table 9

A Paired Sample Test of the Text-plus-Picture Group in
Both Tests

		Mean	Std. Deviation	t	Sig. (2-tailed)
Pair 3 (text-plus -picture)	Immediate test	11.67	8.15	3.99	.000
	Delayed test	7.51	6.35		

The result of the analysis revealed that the

immediate vocabulary test scores of the text-only group,

the picture-only group, and the text-plus-picture group

were significantly higher than their delayed test scores.

Discussion

The first research question addressed in this study

concerned the effects of three different types of

vocabulary annotations -- text-only, picture-only, and

text-plus-picture -- on elementary EFL learners'

vocabulary test scores in the immediate test. The ANOVA analysis revealed that among the three groups, the text-only annotation group got the highest mean scores in the immediate vocabulary test, although it was significant only between the text-only and the picture-only group.

The results of Kost, Foss, and Lenzini (1999) were different from the current study. A combination of the text-plus-picture group were better than the text-only group and the picture-only group in both production (translation) and recognition (multiple-choice) tests. The types of test might affect the results of both studies. The current study measured the subjects vocabulary learning by using only the production test, while Kost, Foss, and Lenzini (1999) used both production and recognition tests. Recognition and production tests are often used to examine students' vocabulary knowledge. However, test and measurement studies indicate that these two forms of testing are quite different and demand separate processing strategies (Cariana & Lee, 2001; Jonassen & Tessmer, 1996). Recognition tests usually involve multiple choice activities which require learners to select or guess the correct response from the alternatives given (McDaniel & Mason, 1985). In contrast, a production test requires learners to respond from memory. It is more demanding than a recognition

test because learners have to search for the correct response within their mental representation of the newly-learned information (Cariana & Lee, 2001; Glover, 1989; McDaniel & Mason, 1985). Moreover, the production test in the current study might be beneficial for the text-only group to get higher vocabulary test scores because the test was similar to the vocabulary annotation that the subjects received.

The second research question addressed in this study concerned the effects of three different types of vocabulary annotations -- text-only, picture-only, and text-plus-picture -- on elementary EFL learners' scores in a three-week delayed test. The results indicated that the text-plus-picture annotation group got the highest mean scores in the delayed test. However, this difference was not statistically significant.

The finding from a study by Chun and Plass (1996a) revealed that words with text-plus-picture annotations were recalled significantly better in the delayed test than words with other types of annotations. The results of the current study revealed that there was a tendency that the text-plus-picture group scored better than other groups in the delayed test although no significant difference was found. Kellogg and Howe (1971) stated that foreign words associated with images or actual

objects are learned more easily than those without additional information. Moreover, Paivio's (1986) dual coding theory was the key explanation. The subjects were able to hold the verbal and visual mental representations simultaneously in verbal and visual working memory, when textual and pictorial units were presented contiguously. The two kinds of representations in memory allowed the subjects to build referential connections between the two systems, which resulted in multiple retrieval routes to the vocabulary items and effects on recall of the information (Paivio, 1991).

Another finding from Yoshii and Flaitz' s study (2002) compared the effects of three multimedia vocabulary annotations -- text annotations, picture annotations, and text-plus-picture annotations -- with ESL university students. The results revealed that the text-plus-picture group outperformed others on both immediate and delayed posttests. These findings contrasted to the current study, which found that the text-plus-picture group tended to perform better than other groups only in the delayed posttest although the difference was not statistically significant. The level of the subjects and the types of test may be the important factors that affected the results of Yoshii and Flaitz (2002) and the current study. In Yoshii and

Flaitz' study (2002), subjects were ESL university students, whereas in the present study, the subjects were elementary EFL students. The proficiency levels of vocabulary background knowledge between these subjects were different. Regarding the types of tests, Yoshii and Flaitz (2002) tested the subjects by using a picture recognition test, a word recognition test, and a definition supply test, while the current study used only a production test. Different types of the tests used may affect the performance of the subjects as discussed earlier. The results of the current study might be statistically significant if the numbers of the subjects increase.

The results of the current study revealed that the picture-only group had the lowest mean scores in both immediate and delayed tests. One of the possible explanation for this was some of the target words were adjective and adverb, [for instance old, young, wet, angry, and quickly] which were difficulty to define by pictures. Unlike concrete words for example "ear" or "tree". The concreteness and abstractness of words might affect the vocabulary test scores of the picture-only group. Some subjects misunderstood and gave the incorrect meaning, for instance, the word "เร็ว" (quickly) was defined as "กระโดด" (jump) or "วิ่ง" (run).

With regard to the comparison between the immediate test and the delayed test, Yoshii and Flaitz (2002) found that the retention scores of all groups declined on the delayed posttest. This is similar to the results of the current study. These findings supported the theory of language acquisition, which is an incremental process (Nagy, Herman, & Anderson, 1985). If L2 learners do not have further exposure to the target words, their memory of the newly-learned information fades away quickly. An incremental process is a reasonable explanation for the subjects whose delayed vocabulary test scores were lower than their immediate test. In contrast, the possible explanation for the subjects whose delayed vocabulary test scores were higher than their immediate test (18 out of 128 subjects) was there might be other factors involved. Some subjects might gain the meaning of the target words from other sources such as asking teachers, searching from the internet, studying in language center or looking up words in the dictionary. These factors could encourage the subjects to get the higher scores in the three-week delayed test.

According to the current study, there were a few subjects who got "o" (2 out of 128 subjects) in both immediate and three-week delayed test. The researcher conducted the informal interview with these subjects in

order to indicate the factors that affected their scores.

Most of them stated that they could not read any target

words so could not give its correct meaning. But when

the researcher read some words for them, they could give

the correct meaning. Thus, it could be concluded that

these subjects had a serious problem in reading skill.

Summary of the Chapter

This chapter described the results of the analysis

of the data to investigate the effects of the three types

of vocabulary annotations -- text-only, picture-only, and

text-plus-picture -- on elementary EFL Thai learners. It

also compared the subjects' vocabulary test scores taken

immediately after the reading task with those taken three

weeks after the task.

In the immediate test, the subjects who read a

passage with text-only vocabulary annotations scored

significantly better than those who read a passage with

picture-only vocabulary annotations. Moreover, the text-

only group also scored higher than the text-plus-picture

and the text-plus-picture scored higher than the picture-

only group, but these differences were not statistically

different.

In the three-week delayed test, there was no significant difference among the three groups and between each pair although the text-plus-picture group got the highest mean scores.

To compare the mean scores of the immediate test and the three-week delayed test of each group, the delayed test' scores of all groups decrease significantly.