Story Animation: Helping Students Build Connections between Words and Pictures in Multimedia Learning

Kuo, Feng-lan & Chiang, Heien-kun
National Changhua University of Education
laflkuo@cc.ncue.edu.tw, hkchiang@cc.ncue.edu.tw

Abstract

Previous research has shown that multimedia annotations can facilitate L2 vocabulary acquisition. The vocabulary annotations being investigated include textual definition, still pictures, video clips, and auditory input. Based on Paivio’s 1986 dual-coding theory, Mayer and Sims (1994) emphasized the interconnectedness of two distinct cognitive systems—visual and verbal. Although the two systems were independent, the interconnectedness of the two systems was the important determinant of conceptualization. Given verbal and visual input simultaneously might assist the construction of referential connection, most previous studies have centered on exploring the effects of multimedia annotations on vocabulary learning of adult ESL/EFL learners (Bell, & LeBlanc, 2000; Chun & Plass, 1996a, b; Jones, 2004; Lomicka, 1998; Yeh & Wang, 2003). In contrast, this study aims to investigate the effects of three types of multimedia annotations: glossing, animations, and glossing plus animations on both vocabulary recognition and reading comprehension of 4 intact classes from an elementary school in central Taiwan. Subjects were given a pre-test to establish the baseline of their vocabulary knowledge. A narrative story was used as the instrument to implement different versions of the vocabulary and contextual annotations. Post-tests results showed that the version with glossing plus animations was the most effective one. Implications and suggestions for further studies will also be provided.

Key words: multimedia annotations, vocabulary acquisition, reading comprehension, dual-coding theory

Many studies have been conducted to explore the effect of multimedia annotations on L2 vocabulary acquisition and/or reading comprehension and positive effects have been found. Regarding subject population, most previous studies lay their focus on the effect of multimedia annotations on adult ESL/EFL learners, instead of young learners (Aweiss, 1994; Bell & LeBlanc, 2000; Jones, 2004; etc.). In addition, subjects of most prior studies come from an alphabetical L1 language background. Learners of such language background tend to rely more on audio channel to process the language information. In contrast, Chinese EFL learners come from a logographic L1 language background. Huang and Hanley (1995), Kroonenberg (1995), Shih (1995), Chikamatsu (1996), and Chen (1999) found that Chinese EFL learners rely heavily on visual information to learn English vocabulary. Therefore, it remained to be shown whether young learners of English with a Chinese L1 background can benefit from glossing delivered through the computer. This study thus aims to investigate the effectiveness of three types of multimedia annotations: glossing, Flash animations, and glossing plus Flash animations on vocabulary recognition and reading comprehension of EFL elementary school students in Taiwan. Two research questions will be answered:

1. Are multimedia annotations useful for second language vocabulary acquisition? If so, which type of multimedia annotation is more useful for Chinese EFL elementary school students’ word recognition and retention?

2. Which type of multimedia annotation is more useful for Chinese EFL elementary school students’ reading comprehension and retention?

Literature Review

In this section, studies related to the theoretical foundations on benefits of dual channels of information processing will be discussed. Empirical findings of the effects of multimedia

---

1 This study is supported in part by National Science Council grant NSC 94-2411-H-018-005.
annotations on incidental vocabulary learning and reading comprehension will be reviewed. Recent studies concerning appropriate reading materials for children will also be presented. In the following sections, it will be shown the findings of each area of studies.

Dual-coding Theory

Paivio’s 1986 dual-coding theory emphasized the interconnectedness of two distinct cognitive systems—visual and verbal. The verbal system contains word-like codes including visual, auditory, articulatory, and other modality-specific verbal codes. In contrast, the nonverbal representations (the visual system) include modality-specific images for shapes, environmental sounds, actions, skeletal or visceral sensations related to emotion and other nonlinguistic objects and events. Based on the dual-coding theory, Mayer and Sims (1994) emphasized the interconnectedness of two distinct cognitive systems. Although the two systems were independent, the referential connection between them was the important determinant of conceptualization. Given verbal and visual input simultaneously might assist the construction of referential connection, most previous studies have centered on exploring the effects of the combination of these two modes of multimedia annotations on vocabulary learning and/or reading comprehension of adult ESL/EFL learners (Bell, & LeBlanc, 2000; Chun & Plass, 1996a; Jones, 2004; Kost, Foss, & Lenzini, 1999; Leutner & Plass, 1998; Lomicka, 1998; Yeh & Wang, 2003). Chun and Plass (1996a), Leutner and Plass (1998), and Kost et al. (1999) found that two modes of information results in better comprehension than only one.

On the other hand, mixed results have been found for different implementation of the combination of visual and verbal information. Chun and Plass (1996a) compared the effects of two types of two-modes presentations. Better recall was found for the group receiving vocabulary annotation with picture plus text than that with picture plus video clip. Al-Seghayer (2001) conducted a similar study on a group of 30 ESL subjects. However, the results showed that the printed text definition with video clips version produced significantly better results on both vocabulary recognition test and production test than the version implemented with printed text definition and still pictures. A possible factor affecting the results of the above studies might be whether the visuals bear relevant context or content information that complements the verbal (written) portion of the stimulus (Ginther, 2002). In Taiwan EFL context, Yeh and Wang (2003) compared the effects of three types of multimedia annotations for enhancing vocabulary acquisition: text only, text plus picture, and text plus picture and sound. It was found that the text plus picture version was the most effective for L2 vocabulary learning. Results of these studies consistently provided support for the dual-coding theory and the generative theory of multimedia learning (Mayer, 1997; Paivio, 1986). The interaction of textual source of information and graphic has led to a complementary effect and thus produces better comprehension and vocabulary learning outcome. However, it remained to be shown which type of visuals will produce the largest effect on vocabulary retention and facilitate reading comprehension.

Retention Effects of Multimedia Annotations

Learning styles of learners have been shown to affect the learning outcomes. Dunn, Dunn, and Price (1989) categorized secondary students in the United States into four sub-types of learners: auditory learners, visual learners, tactile learners, and kinesthetic learners. Among them, the auditory and visual learners comprised the majority at 70%, with only 30% for the other two types of learners. In a different learning environment, Wallace (1995) used the Learning Style Inventory developed by Dunn et al. (1989) to probe the learning styles of a group of 450 Philippine sixth and seventh graders. It was found that 8.4% of the students were classified as auditory learners, 41.4% as visual learners, 20.4% as tactile
learners, whereas the rest 29.8% were identified as kinesthetic learners. Though the results of the two studies differ to a certain degree, both studies have revealed that visual learners indeed comprised a majority of all learner types and the researchers further suggested that students learn best when teaching method matches their learning styles.

Regarding effects of retention, prior research mostly used an immediate posttest to measure effectiveness of multimedia annotations on language learners’ vocabulary acquisition or post-reading comprehension. The empirical findings of Rusted and Coltheart (1979a) revealed that pictures can provide a major cue to retrieve the meanings of words and thus facilitate recall. In a follow-up study, Rusted and Coltheart (1979b) further pointed out that visual aids have the largest effects on delayed task recall.

In recent studies, Chun and Plass (1996a), Kost et al. (1999), and Jones (2004) explored not only the immediate effects but also the delayed effects of multimedia annotations on adult ESL/EFL learners’ vocabulary acquisition. Chun and Plass (1996a) found that in both the immediate and 2-week delayed vocabulary production test the picture plus text group performed better than the text only group and the video plus text group. However, the video plus text group was found to perform best in the immediate vocabulary recognition test while it is the picture plus text group performing the best in the 1-week delayed vocabulary recognition posttest. Chun and Plass (1996a) attributed the results to the dual-coding effect since learning of a vocabulary item is best achieved when both verbal and visual information are presented (Mayer & Sims, 1994; Paivio, 1986). Kost et al. (1999) compared the effects of three types of gloss conditions: text only, picture only, and text plus picture on incidental vocabulary growth. The results revealed that in the picture recognition task the text plus picture group significantly outperformed the text only group on the immediate as well as the 2-week delayed posttest. In contrast, in the word recognition task subjects under textual gloss condition significantly outperformed the subjects utilizing pictorial gloss in the immediate posttest. In the delayed task, though not significant enough the textual group still performed better than the pictorial group. The finding indicated that there is some effect of correspondence between the gloss condition and the test format. However, there were no significant differences among the three groups on the production task in either the immediate or the delayed posttest. This seemed to indicate that there might be some interactions between task types (recognition task or production task) and glossing conditions.

Furthermore, in contrast to other studies using a written text, Jones (2004) examined the influence of multimedia annotations on vocabulary learning from an aural text. The results of the study showed that subjects in the three treatment groups receiving written, pictorial, or both written plus pictorial annotations outperformed the control group receiving no annotations during listening in both the immediate posttest and the 3-week delayed posttest on written vocabulary recognition task as well as the pictorial vocabulary recognition task. However, in the vocabulary production task, it was found that subjects performed best when the mode of testing matched the treatment. In other words, the pictorial plus written annotations group and the written annotations group performed as good in the written vocabulary test whereas the pictorial plus written annotations group and the pictorial annotations group performed the best in the pictorial vocabulary test. Results of Jones’s study provided partial support for the benefit of dual channels of information processing. Nevertheless, the results also pointed to an opposite direction in that the pictorial mode of information did not increase the efficiency of vocabulary learning which is counter to the findings of Kost et al’s (1999) study. According to Jones (2004), this might result from the fact that images may have provided too much information than the more precise information provided by direct translations similar to the claim of Sakar and Ercetin (2005). Sakar and Ercetin (2005) studied the effects of two types of annotations on 44 intermediate-level Turkish EFL learners’ reading comprehension. The textual annotations included three
subtypes of glossing: text, audio, and graphics. In the extratextual annotations, video is added as a fourth type of glossing. Effects of overall annotation use indicated that performance on the reading comprehension test degraded as annotation use increased. Regarding the effect of particular type of annotations on reading comprehension, it was further found that audio annotation and video annotation in fact had a deleterious impact on comprehension possibly due to the overload on a single channel of information processing. According to the suggestion of Chun and Plass (1997), a second explanation might be because the participants of Sakar and Ercetin’s (2005) study have not reached a certain language proficiency level to take advantage of the information given in the various annotations to facilitate reading comprehension.

However, results of previous studies must be interpreted with caution since some studies focused on investigating the effects of annotations on vocabulary learning (Al-Seghayer, 2001; Jones, 2004; Kost et al., 1999; Yeh & Wang, 2003) or reading comprehension alone (Aweiss, 1994; Bell & LeBlanc, 2000; Lomicka, 1998; Omaggio, 1979), while others centered on exploring their effects on both vocabulary acquisition and reading comprehension (Chen, 2002; Chun & Plass, 1996a, b; Leutner & Plass, 1998). For instance, Chen (2002) compared the effects of L1 and L2 glosses on reading comprehension and vocabulary retention of Chinese EFL learners. The results showed that L2 glosses had the largest effect on reading comprehension while L1 glosses had the best effect on vocabulary retention. In addition, measurements were found to vary across studies on the effect of annotation types on vocabulary acquisition and reading comprehension. For example, Omaggio (1979) used a recall and a vocabulary recognition test to measure effects of visuals on reading comprehension. Aweiss (1994) used only recall protocol for reading comprehension assessment. In contrast, Chun & Plass (1996a) used a vocabulary recognition test, a vocabulary production test as well as a recall protocol to measure effects of multimedia annotations on vocabulary acquisition and reading comprehension.

Story Books as Appropriate Reading Material for Children

Prior research has shown that children, especially younger and lower proficiency learners, were found to rely more on visual aids to process information than adults (Bornens, 1990; Goldstein & Underwood, 1981; Paris & Paris, 2003; Rusted & Coltheart, 1979a, b). Paris and Paris (2003) found that in the past 10 years studies conducted in the United States mostly focused on exploring children’s decoding, instead of their reading comprehension skills. Based on the story grammar framework of Mandler and Johnson (1977), Paris and Paris (2003) further emphasized that reading story books with illustrations can provide a vocabulary referent, a context, and can stimulate students’ reading interest, and thus facilitate reading comprehension. Since story books contain the 5 major narrative elements: characters, setting, initiating event, problem/episodes, and solution, and words and sentences used in narratives are a form of language commonly found in everyday life, both in interactions with others and in educational and recreational media, they are considered by Paris and Paris (2003) to be authentic experiences in young children’s daily lives and thus are recommended as suitable reading materials for children. In an L1 context, findings of numerous studies have revealed that storytelling as well as reading stories can help developing the language skills of children and can motivate active learning (Cooper, 1989; King & Ippolito, 2001; Koki, 1998; Wilson, 1997). On the other hand, Bishop and Edmondson (1987) claimed that preschool narrative performance appears to predict children’s later language development as well as reading comprehension. Results of these studies suggested that story books are appropriate reading material for children.

In an EFL context, scholars have constantly emphasized the importance of integrating story-reading into the elementary English curriculum (Chang, 2002; Chien & Huang, 2000;
Chuang, 1999). However, only a few studies have been conducted to evaluate the effectiveness of storytelling instruction on children (Hsieh, in press; Liao, 2002; Tsou, 2003; Tsou, Wang, & Tzeng, in press). By integrating story books into a whole language classroom in an elementary school in central Taiwan, Liao (2002) showed that after 13 months of English instruction, the 37 5th graders were found to make improvements in phonics skills, vocabulary ability, and story-reading ability. Furthermore, the class becomes more homogeneous regarding their overall English ability. Tsou (2003) compared teacher’s and students’ behaviors in the storytelling classroom versus those in the non-storytelling English classroom. Subjects of her study were 29 fifth graders respectively from two intact classes in Southern Taiwan. Analysis of the teacher talks showed that the teacher in the storytelling classroom used more open questions, prompts and student volunteers than in the non-storytelling classroom. From the students’ perspective, it was found that the experimental subjects’ oral participation was enhanced through storytelling. Furthermore, a more positive classroom atmosphere was promoted by storytelling. Hsieh (in press) compared the effectiveness of storytelling instruction and traditional dialogue instruction on the two intact classes of 65 fourth graders in Southern Taiwan based on six post-measures: listening comprehension, vocabulary matching, unscrambling the story, reading comprehension, filling in the blanks, and completing the dialogue. The results revealed that the storytelling group performed significantly better than the dialogue group on the first four measures. On the other hand, Tsou et al. (in press) developed a multimedia Storytelling Website to demonstrate the effectiveness of web-based technology in facilitating teacher’s storytelling and children’s story recall processes. Posttest results showed that the Storytelling Website group outperformed the regular storytelling group (the control group receiving the same teaching procedures as the experimental group but without additional support from the Storytelling Website) in story sentence complexity and post-instructional language proficiency test. Though the two groups did not differ significantly regarding story comprehension, the Storytelling Website group was found to produce more details in story recall. Furthermore, questionnaire results indicated that subjects in the Storytelling Website group tended to be more confident and enjoy the story recalling process than their control counterparts.

In summary, prior studies have shown that two modes of information integrating both verbal and visual presentation results in better comprehension than only one. A survey of previous studies on the effects of multimedia annotations on vocabulary acquisition and reading comprehension of second language learners has further shown that subject population of most studies are adult learners. With the exceptions of Aweiss’s (1994) study on native speakers of English learning Arabic as a foreign language, Sakar and Ercetin’s (2005) study on native speakers of Turkish learning English as a foreign language, and Chen’s (2002) and Yeh and Wang’s (2003) study on Chinese subjects learning English as a foreign language, the subject population of prior research mostly are native speakers of English with an alphabetical language background learning another alphabetical second language such as French, German, and Spanish. Furthermore, multimedia annotations are found to have different effects on L2 vocabulary learning and reading comprehension. In addition, most of the previous research only centers on investigating the immediate effects of multimedia annotations. On the other hand, narrative stories were shown to be effective reading material for facilitating the development of children’s language skills. Therefore, further research is needed to examine the immediate as well as delayed effectiveness of CALL-based annotations of narrative stories on Chinese EFL elementary school students’ vocabulary learning and reading comprehension.
Method

Subjects

Ninety-one 5th graders from 4 intact classes in one elementary school in central Taiwan voluntarily participated in the study during their self-study periods (between 7:30am and 8:30am). Officially these students started learning English at 3rd grade. They received 1 hour of English instruction per week at 3rd and 4th grade, and they started to receive 2 hours of English instruction per week at 5th grade. The average age of these subjects are 11 years and 5 months old, and their average starting age for English learning was 7 years and 10 month old. This indicates that some of the subjects began studying English at a private sector before they received formal English education at school. The control group receiving no special treatment was composed of 23 subjects. The experimental group 1 receiving textual annotations included 20 students; the experimental group 2 receiving pictorial (Flash) annotations consisted of 24 subjects, whereas the experimental group 3 receiving both textual and pictorial annotations contained 24 students. Before the experiment, both the control group and the treatment groups completed a 20-item vocabulary recognition pretest to determine their prior knowledge of the target vocabulary used in this study. The twenty unfamiliar words were first chosen by the first author, then verified by the students’ English teacher and later marginally glossed (see Appendix 1 for the list of target words). These 20 words are content words including concrete nouns, adjectives, and verbs. As shown in Table 1, all subjects were found to have low prior knowledge of the vocabulary with an average score of 7.65, 10.10, 8.08, and 10.08, respectively out of a maximum score of 20. The ANOVA test results on the homogeneity of the 4 groups were found to be non-significant, $F(3, 87) = 2.60, p > .05$.

<table>
<thead>
<tr>
<th>Groups</th>
<th>n</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>23</td>
<td>7.65</td>
<td>3.34</td>
</tr>
<tr>
<td>Textual annotations</td>
<td>20</td>
<td>10.10</td>
<td>4.83</td>
</tr>
<tr>
<td>Pictorial (Flash) annotations</td>
<td>24</td>
<td>8.08</td>
<td>3.05</td>
</tr>
<tr>
<td>Textual plus pictorial annotations</td>
<td>24</td>
<td>10.08</td>
<td>3.40</td>
</tr>
</tbody>
</table>

Note: Maximal score = 20

Materials

The book used as the reading text was The Outing, a 24-page story book written by Hunt and Lui (1998) and illustrated by Brychta, A. It tells the story of a group of children who first went to the zoo for an outing. Because of the rain, one of the boys suggested the whole class went to the museum instead. The children finally had a good time in the museum. The reason for selecting this book as the reading material is that it has a clear story line with an obvious sequence of events and the five main elements of stories as suggested by Paris and Paris (2003). This narrative story was judged by the researchers to be interesting and of approximately suitable level of difficulty for the subjects of the present study. An additional criterion for selecting this reading passage was that the book was designed to provide multiple exposures of target words to facilitate incidental vocabulary learning. A text analysis revealed that there were a total of 512 words and 151 different words in the story. This leads to a low type-token ratios of 30% indicating it should be easily understood by readers. Vocabulary difficulty analysis (checking against Jeng’s (2001) CEEC SAET word list used frequently in elementary and secondary school level in Taiwan) further showed that most of the vocabulary were high frequency words with 85% of them as level 1 vocabulary and 7% of them as level 2 vocabulary. Analysis on the content words of the text revealed that 15 of the
words occurred more than 5 times in the selected text. There are another 35 lexical items which appeared between 2 to 4 times in the text. This is in line with the suggestion of Nation (1990) that a range from five to sixteen exposures may be needed for vocabulary acquisition. Two bilingual girls studying in a junior high school and a senior high school respectively were recruited to narrate the story. They were asked to dramatically enunciate the lines of the characters as distinctly as possible. Their narration lasted about 6.5 minutes long, thus their reading speed was around 80 words per minute. In addition, the words were articulated clearly with enough pauses at phase and sentence boundary.

Four different multimedia versions of the same narrative were implemented: a text only version, a textual glossing version, a pictorial version (Flash animation), and a textual plus pictorial version (Flash animation). All four versions of the text were provided with the oral narration.

Figure 1. Screen shot of the no glossing version of the online narrative

Figure 2. Screen shot of the textual glossing version of the online narrative
Figure 3. Screen shot of the pictorial glossing version of the online narrative

Figure 4. Screen shot of the textual plus pictorial glossing version of the online narrative

Figure 1 shows the screen shot of the no glossing version of the narrative. As shown in Figure 2, the textual glossing version provides learners with equivalent Chinese translation as well as a sample English sentence containing the target vocabulary. In contrast, the pictorial version shown in Figure 3 was implemented using the multimedia authoring program Macromedia Flash MX 2004. This type of annotations provides learners with animations which may directly or indirectly contain the target vocabulary (content visuals) and provides the context visuals (Ginther, 2002). On the other hand, Figure 4 shows the screen shot of the textual plus pictorial version including both the verbal information provided in the textual glossing version and the visual information provided in the pictorial (Flash) version. The four
multimedia treatments were presented to the four groups of students respectively using a 40-station personal computer lab. The participants were asked to listen to the story content through headsets and view the reading text from their own computer screens. In addition, they were allowed to go back and forth the web pages according to their own pace. After completing the story reading, they were given the immediate posttest without accessing the story web pages.

Instruments
The instruments administered to the subjects in the present study were a background questionnaire, a pretest, an immediate posttest, and a 2-week delayed posttest. The pretest is a 20-item multiple choice test of vocabulary recognition knowledge. The maximum score for this test was 20. In contrast, the immediate posttest contains a 6-item multiple choice test of reading comprehension in addition to the 20-item test of vocabulary recognition knowledge (see Appendix 2 for the list of comprehension questions). The maximum score for the immediate posttest was 26. Items included in the delayed posttest are exactly the same as those in the immediate posttest.

Procedures
A pre/posttest control group design was used to observe the effects of multimedia annotations on 4 groups of elementary school students’ L2 vocabulary learning and reading comprehension. All activities were completed in two self-study sessions. During the first self-study period, students were allowed 10 minutes to complete the written vocabulary recognition pretest to measure their initial knowledge of the 20 unfamiliar words. Then each of the four groups was randomly assigned to read one of the 4 different versions of the story respectively. All of the students were given up to 10 minutes to read the story while the three experimental groups were allowed to access the prescribed annotations respectively. After reading the story, the subjects were allowed up to 10 minutes to take the immediate posttest. Next the subjects were given a questionnaire in order to understand their English learning experience. Two weeks later, without any additional experience with the reading text and without any prior notice, following the same test procedure as the immediate posttest the subjects took the delayed posttest during the second self-study period.

Results and Discussion
In order to answer the first research question, in the first analysis, paired t-tests were used to compare the gain of each group between the immediate posttest and the pretest. It was found that the control group receiving no glossing annotations did not gain any improvement in the immediate vocabulary recognition posttest, \( t = 2.04, p > .05 \). In contrast, all three experimental groups performed significantly better in the immediate vocabulary post measure. The results showed that the textual glossing group performed significantly better in the immediate vocabulary recognition posttest, \( t = 6.37, p < .001 \); the pictorial annotations group had significant performance, \( t = 2.95, p < .01 \); while the textual plus pictorial annotations group also had significant performance, \( t = 8.33, p < .001 \).

In the second analysis, the correct answers on the immediate posttest among the four groups then were compared. The descriptive statistics are presented in Table 2. The ANOVA test results for the immediate vocabulary recognition posttest were found to be significant, \( F(3, 87) = 25.80, p < .001 \). Scheffe multiple group comparisons showed that the textual and textual plus pictorial annotations group alike outperformed the control group and the pictorial annotations group. However, the two latter groups did not differ significantly from each other. Similarly, there was no difference on the performance between the textual and the textual plus pictorial annotations group. This supports the finding of Jones (2004) that textual...
glossing provides precise information through direct translation and thus facilitates vocabulary learning.

Table 2. Mean group scores and standard deviations on the immediate vocabulary recognition posttest results

<table>
<thead>
<tr>
<th>Groups</th>
<th>n</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>23</td>
<td>8.57</td>
<td>3.33</td>
</tr>
<tr>
<td>Textual annotations</td>
<td>20</td>
<td>15.95</td>
<td>4.17</td>
</tr>
<tr>
<td>Pictorial (Flash) annotations</td>
<td>24</td>
<td>9.71</td>
<td>4.27</td>
</tr>
<tr>
<td>Textual plus pictorial annotations</td>
<td>24</td>
<td>15.96</td>
<td>2.30</td>
</tr>
</tbody>
</table>

Note: Maximal score=20

In the third analysis, the correct answers on the 2-week delayed posttest among the four groups were compared. The descriptive statistics are presented in Table 3. Similar to the results of the immediate posttest on vocabulary recognition, a one-way ANOVA yielded a significant $F$-statistic for the delayed vocabulary recognition posttest, $F(3, 87) = 11.10$, $p < .001$.

Table 3. Mean group scores and standard deviations on the delayed vocabulary recognition posttest results

<table>
<thead>
<tr>
<th>Groups</th>
<th>n</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>23</td>
<td>8.52</td>
<td>3.63</td>
</tr>
<tr>
<td>Textual annotations</td>
<td>20</td>
<td>13.65</td>
<td>4.60</td>
</tr>
<tr>
<td>Pictorial (Flash) annotations</td>
<td>24</td>
<td>10.33</td>
<td>4.05</td>
</tr>
<tr>
<td>Textual plus pictorial annotations</td>
<td>24</td>
<td>14.38</td>
<td>3.62</td>
</tr>
</tbody>
</table>

Note: Maximal score=20

Scheffe multiple group comparisons showed that the textual and textual plus pictorial annotations group outperformed the control group. Similar to the results of the immediate vocabulary posttest, the performance of the pictorial annotations group and the control group did not differ significantly. Nevertheless, differing from the immediate posttest results, only the textual plus pictorial annotations group, but not the textual annotations group, was found to perform significantly better than the pictorial annotations group. Although the textual annotations group had higher scoring on the delayed vocabulary test, the textual glossing group did not outperform the pictorial annotations group ($t = 3.32$, $p = .06$). Paired $t$-test results further revealed that the performance of the textual annotations group indeed degraded significantly on vocabulary recognition in the delayed posttest ($t = 3.52$, $p < .005$). This provides partial support to the findings of Rusted and Coltheart (1979b) that visual aids, but not verbal aids, have the largest effects on delayed task recall. In contrast, the results strongly complied with the findings of earlier studies on dual-coding theory that combination of two modes of information will lead to better comprehension than only one mode of information (Chun & Plass, 1996a; Leutner & Plass, 1998; Kost et al., 1999). Similar to the findings of the immediate posttest, there was no difference on the performance between the textual and the textual plus pictorial annotations group. This seemed to comply with the findings of Shohamy and Inbar (1991) that “local” information on factual details and the meaning of lexical items were easier to answer than “global” questions on main ideas and inferences. The results again support the finding of Jones (2004) that textual glossing provides precise information through direct translation and thus facilitates vocabulary learning. Though the latter two groups performed equally well, paired $t$-test results showed that the performance of the textual plus pictorial annotations group also degraded significantly on vocabulary
recognition in the delayed posttest \( (t = 2.45, p < .05) \). Comparing the results of the immediate posttest with those of the delayed posttest, it seemed to indicate that vocabulary annotations provided more “local” information on factual details and the meaning of lexical items, thus it is easier to correctly answer the test questions in the immediate posttest. However, the retention effects in the delayed task were not as clear as those found in the immediate posttest. A possible explanation is that too many textual annotations might create an overload on a single channel of information processing (e.g., Sakar & Ercetin, 2005). According to Kost et al. (1999) and Jones (2003), a second possible explanation is that pictures demanded deeper processing than did L1 verbal translation and therefore subjects receiving both textual and pictorial annotations demonstrated greatest incidental vocabulary retention.

In order to answer the second research questions, in the fourth analysis, the correct answers on reading comprehension of the immediate posttest among the four groups were computed. Mean group scores and standard deviations are presented in Table 4. The ANOVA test results for the immediate reading comprehension posttest were found to be significant, \( F(3, 87) = 4.53, p < .005 \).

<table>
<thead>
<tr>
<th>Groups</th>
<th>n</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>23</td>
<td>2.17</td>
<td>1.40</td>
</tr>
<tr>
<td>Textual annotations</td>
<td>20</td>
<td>3.40</td>
<td>1.88</td>
</tr>
<tr>
<td>Pictorial (Flash) annotations</td>
<td>24</td>
<td>3.38</td>
<td>1.74</td>
</tr>
<tr>
<td>Textual plus pictorial annotations</td>
<td>24</td>
<td>3.92</td>
<td>1.67</td>
</tr>
</tbody>
</table>

Note: Maximal score=6

Scheffe multiple group comparisons revealed that only the textual plus pictorial annotations group significantly outperformed the control group. However, there were no significant differences among the other three groups. In comparing the results on the vocabulary test and the reading comprehension test, this seemed to comply with the findings of Shohamy and Inbar (1991) that “global” comprehension questions on main ideas and inferences were harder to answer than vocabulary tests related to “local” information on factual details and the meaning of lexical items. It also further supports the findings of the dual-coding theory that two modes of information processing are more effective for enhancing L2 reading comprehension.

In the last analysis, the correct answers on reading comprehension of the delayed posttest among the four groups were computed. Mean group scores and standard deviations are presented in Table 5. The ANOVA test results for the delayed reading comprehension posttest were found to be significant, \( F(3, 87) = 8.28, p < .001 \).

<table>
<thead>
<tr>
<th>Groups</th>
<th>n</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>23</td>
<td>1.74</td>
<td>1.32</td>
</tr>
<tr>
<td>Textual annotations</td>
<td>20</td>
<td>2.80</td>
<td>1.70</td>
</tr>
<tr>
<td>Pictorial (Flash) annotations</td>
<td>24</td>
<td>3.42</td>
<td>1.67</td>
</tr>
<tr>
<td>Textual plus pictorial annotations</td>
<td>24</td>
<td>4.00</td>
<td>1.77</td>
</tr>
</tbody>
</table>

Note: Maximal score=6

Differing from the results of the immediate posttest, Scheffe multiple group comparisons revealed that both the pictorial group and the textual plus pictorial annotations
group significantly outperformed the control group. However, there were no significant differences among the other groups. It was found that the performance of the control group and the textual glossing group degraded in the 2-week delayed posttest. In contrast, the performance of the pictorial group and the textual plus pictorial annotations group remained stable. The results further extend the finding of Kost et al. (1999) that there are some interactions between test formats and glossing conditions. That is, textual glossing and textual plus pictorial annotations are more effective for answering vocabulary recognition test, whereas pictorial annotations and textual plus pictorial annotations are more effective for answering reading comprehension test. This seemed to indicate that visuals can facilitate children’s understanding of the story and promote information retention. This further extends the finding of Kost et al. (1999) and Jones (2003) that pictures demanded deeper processing than did L1 verbal translation and therefore subjects receiving both textual and pictorial annotations not only demonstrated greater incidental vocabulary learning but also exhibited deeper reading comprehension than those receiving other types of annotations. The results are partially consistent with the findings of Rusted and Coltheart (1979b) that visual aids have the largest effects on delayed task recall. Unlike the short retention effect found in the delayed posttest on vocabulary recognition, the retention effect of visuals on reading comprehension seemed to be stronger as shown in the results of delayed posttest. Nevertheless, the combination of two modes of information processing was found to be better than just single mode of information processing.

Conclusion and Future Studies

The results of the present study extend the finding of Jones (2004) that textual annotations as well as textual plus pictorial annotations can facilitate L2 vocabulary acquisition since direct translation provides precise information for a particular vocabulary. Regarding reading comprehension, it was found that the textual plus pictorial annotations group outperformed the other three groups in the immediate posttest. In contrast, both the pictorial annotations group and the textual plus pictorial annotations group performed significantly better than the other two groups in the delayed task. This further extends the finding of Kost et al. (1999) and Jones (2003) that pictures demanded deeper processing than did L1 verbal translation and therefore subjects receiving pictorial annotations and textual plus pictorial annotations demonstrated better performance in delayed tasks requiring global understanding of the text than those receiving other types of annotations. On both vocabulary recognition and reading comprehension task, the textual plus pictorial annotations group was found to perform the best among the four groups. The results also provide strong support for the dual-coding theory (Paivio, 1986) that textual plus pictorial annotations are most effective for enhancing L2 development. In future studies, the effects of multimedia annotations on subjects’ productive vocabulary knowledge should be investigated. Furthermore, the performance of subjects of different language proficiency levels should be explored. Other test formats such as pictorial testing should also be examined (Jones, 2004; Kost et al., 1999).

References
156-173.


莊坤良 (1999), 兒童文學與英語教學: 兒童歌謠、圖畫童書、簡易故事書與分章故事書。國民小學英語科教材教法。戴維揚主編。台北: 文鶴。頁 283-301。

鄭恆雄 (2001)，大考中心學科能力測驗英文參考詞彙表。第五屆國際多媒體語文教學研討會論文集。台北：文鶴。頁 208-216。

廖美玲 (2002)，以童書教英語之成效研究。第六屆兒童文學與兒童語言學術研討會論文集。台北：富春文化公司。頁 68-89。

張湘君 (2002)，聽故事學英文，「粉」另類!英文工廠 I 。頁 4。

鄒文莉 (2003)，說故事英語教學與國小英語學習。師大學報:人文社會類, 48*(1)*。頁 53-67。
Appendix 1
Words included in the pretest, immediate posttest, and delayed posttest

1. glowing  2. push  3. sets off  4. shouted
5. frightened  6. got off  7. toilet  8. kicked
9. flew  10. adventure  11. splash  12. silly
13. crocodile  14. museum  15. dinosaur  16. apatosaurus
17. camera  18. photograph  19. bought  20. model

Appendix 2
Comprehension questions included in the immediate posttest and the delayed posttest

1. _____ Which of the followings is not a character of the story?
2. _____ What was the first outing place the children go to?
3. _____ What was the second outing place the children go to?
4. _____ Why did the children go to the second outing place?
5. _____ Who suggested the children go to the second outing place?
6. _____ What did the children not do in the second outing place?