

L1 glosses: Effects on EFL learners' reading comprehension and vocabulary retention

Ying-Hsueh Cheng
Ohio State University
United States

Robert L. Good
National Kaohsiung First University of Science and Technology
Taiwan

Abstract

The present study examines the effects of 3 kinds of glosses—first-language (L1) Chinese glosses plus second-language (L2) English example sentences, L1 in-text glosses, and L1 marginal glosses—in comparison with a no-gloss condition in reading an English passage, to explore whether providing glosses can facilitate reading comprehension and vocabulary acquisition. A total of 135 undergraduate business and engineering students at 4 English proficiency levels studying at a technical university in Taiwan completed 1 vocabulary pretest, 1 reading session, 1 posttest, and 2 delayed vocabulary recall tests. The study found that L1 glosses helped subjects learn new words and review learned words. Learners' retention declined between the immediate and the 1st delayed recall tests. However, between the 1st and 2nd delayed recall tests, a slight increase in retention was observed for all groups. Unexpectedly, reading comprehension did not improve significantly. Additionally, a questionnaire queried learners' experience using glosses during reading.

Keywords: reading comprehension, vocabulary retention, Chinese glosses, language forgetting patterns, vocabulary gloss questionnaire

For many university students in Taiwan who study English as a foreign language as part of their general education requirements, reading has long been considered an essential skill. Even though their courses in their major are usually taught in their first language (L1), which is Mandarin Chinese in the environment of this study, some of their textbooks for these classes are in English. In fact, they are generally university-level textbooks geared for native English speakers. For less proficient second-language (L2) students, reading in English can be an ordeal, often because of the great amount of unknown vocabulary that makes it difficult or even impossible to get the main idea or specific details of a text. Over the past decades researchers have investigated many aspects of reading including reading comprehension, reading interest, text difficulty and readability, reading strategies, and vocabulary acquisition and retention. One of the methods

appearing to make L2 reading more effective is the use of glosses either in L1 or L2 that are provided somewhere near the text. Researchers have examined various kinds of glosses¹ that facilitate reading comprehension or vocabulary acquisition by L1 speakers or L2 learners of Spanish (Jacobs, Dufon, & Hong, 1994), French (Joyce, 1997), Korean (Myong, 2005), Russian (Gettys, Imhof, & Kautz, 2001), German (Rott, Williams, & Cameron, 2002), and Chinese (Huang, 2003). Such glosses may consist of L1 translations or L2 definitions or both, with or without example sentences. Some benefits of glossing have been claimed to include (a) providing learners with essential target word knowledge for bottom-up processing (Gettys et al., 2001), (b) preventing learners from making wrong inferences (Paribakht & Wesche, 1997), and (c) replacing dictionary use during reading and making vocabulary access convenient for readers (Karp, 2002). Numerous empirical studies concerning the effects of L1 glosses have also been carried out. Hulstijn, Hollander, and Greidanus (1996) gave Dutch undergraduates studying advanced French an adapted story (1,306 words) and investigated 16 target words under two conditions during reading (L1 right-hand marginal glosses and dictionary use) to see if either would improve learners' incidental vocabulary learning. The results revealed that the effect of L1 marginal glosses was greater than that of dictionary use because readers seldom used the dictionary during their reading.

In order to probe whether marginal glosses would improve comprehension of a story, Davis (1989) recruited 71 American students in a French class and divided them into three conditions: (a) read-write-reread; (b) marginal glosses, questions, and comments prior to reading; and (c) marginal glosses, questions, and comments during reading. His findings indicated that subjects who received vocabulary help either before or during reading did significantly better than those who received no help. Some studies, however, have shown no significant effects with glossing. Using a recall protocol, Jacobs et al. (1994) asked 85 English-speaking subjects to (1) write down in their L1 everything they could remember after reading an L2 text, and (2) translate vocabulary items into English. Their study investigated three gloss conditions (L1 English glosses, L2 Spanish glosses, and no glosses) by giving subjects a Spanish text (613 words) with 32 glosses. Their overall findings suggested that although high-proficiency participants who had glosses recalled more of the text, and those who had glosses performed better in the vocabulary translation tasks, there was no significant difference among the three conditions on reading comprehension and vocabulary learning. They also pointed out that significant differences only appeared in the immediate vocabulary translation task; no difference was found on the delayed vocabulary translation task. The researchers attributed the slightly improved effect on the text recall to the subjects' high language proficiency and offered two possible explanations consistent with this view. One was that these subjects may have found the text too easy and did not rely on the glosses. The other was that the text might have been so difficult that the glosses were vital to processing the text. In the latter case, the researchers concluded that "perhaps only those students with higher than average proficiency possessed sufficient L2 competence to make effective use of the glosses provided" (p. 26). In either case, careful text selection is essential in assessing the value of glosses. We have attempted to address this issue in our study by following a strict protocol for selecting texts. As to the performance on vocabulary retention, the researchers indicated that the fact that the superior performance disappeared on the posttest was probably due to the lack of exposure during the four-week interval between the initial and follow-up tests.

Adopting a similar approach, Joyce (1997) also used recall protocols to test subjects'

comprehension after reading. She explored the effects of glossing one of the intermediate and advanced French textbooks which was being used at the University of Pennsylvania. An authentic text (485 words), in the field of journalism, was distributed to 90 undergraduates under two conditions (L1 English marginal glosses and no glosses). After the subjects read the text, they were instructed to write down whatever they could remember of the text in their L1 (English). The results from the recall protocol again showed that subjects receiving glosses did not recall significantly more than the control group. To account for this result, Joyce suggested that subjects might have understood much more than they recalled in a recall protocol task. She recommended that a variety of reading assessment methods should be employed to cross-check subjects' comprehension in future studies. Whether readers can acquire words incidentally and remember them for an extended period of time is another question that has been investigated using glosses. Huang (2003) looked at three kinds of gloss conditions (L2 English glosses, L1 Chinese glosses, L2 English glosses plus L2 English example sentences) for reading comprehension and vocabulary retention with 181 third-year junior high subjects in Taiwan and found a forgetting pattern. The text (*Fry Graph*, difficulty level = 7.0) was chosen from *Studio Classroom*, a popular local English learning magazine. It was pilot-tested, reduced in length to 313 words with 15 glosses placed before the beginning of the text. The 2-week long study (including a vocabulary pretest, reading comprehension test, immediate vocabulary recall test and two delayed vocabulary recall tests) took place in four different sessions. She gave each of the comparable, intact classes the text with just one of the three gloss conditions; she also had a control group with no glosses. Her findings showed that the groups reading with any of the three kinds of gloss conditions performed significantly better than the control group, which meant that glosses could indeed increase subjects' reading comprehension and vocabulary recall. Huang also investigated the language forgetting pattern exhibited in the subjects' three vocabulary recall tests. The trend observed in the study showed that vocabulary recall decreases over time; however, the decline is the sharpest between the immediate recall test and the first delayed recall test 1 week later. Similar results were also found in Watanabe's (1997) study. The decline between the first and second delayed tests was not nearly as steep in Huang's study, or in the case of Watanabe's study even rose slightly.

Both of these studies adopted Pimsleur's (1967) *graduated interval recall* hypothesis to account for their findings showing that after learners acquired vocabulary, these words would fade from their memory rapidly if there was no reviewing process. Pimsleur suggested, however, that if teachers reviewed part of the words periodically, students' memory would be reinforced when it began to fade. Hence, he suggested that teachers should review new vocabulary items very frequently right after they are first presented and continue reviewing them during the following days or weeks at increased intervals.

In sum, in many studies, there appears to be a positive effect of glosses in facilitating reading comprehension or promoting vocabulary acquisition. We also see significant glossing effects on subjects' short-term retention of vocabulary. In those studies where no such effect is observed, we might attribute this to other factors such as the level of language proficiency of subjects, the readability or appropriateness of the reading passages, the tasks employed to demonstrate comprehension or recall, or perhaps even gloss types.

We have observed that studies exploring the use of L1 glosses by Taiwan's university students,

particularly non-English majors, are lacking. Moreover, little is known about the effects of three specific combinations of vocabulary glosses. The current study investigated whether providing three kinds of L1 glosses in or near a reading text could assist Taiwan's technical university non-English majors at different English proficiency levels to read texts, acquire new vocabulary, and retain words over time. The three kinds of L1 glosses are (a) Chinese glosses plus English example sentences (presented prior to the text on a separate page), (b) Chinese in-text glosses (presented next to the target words inside the text), and (c) Chinese marginal glosses (presented below the text). In this paper, we report on the findings for the following research questions:

1. What are the effects of each of the individual gloss conditions (L1-gloss-L2-ex, L1-in-text-gloss, and L1-MG-gloss) on the subjects' performance on reading comprehension and vocabulary recall tests compared to a no-gloss control group?
2. Is language proficiency a factor in the effectiveness of gloss conditions?
3. What kind of forgetting pattern is seen on the subjects' immediate and delayed vocabulary recall tests, and what implications does it have for vocabulary retention?
4. What are the subjects' attitudes toward learning vocabulary through L1 or L2 vocabulary glosses or vocabulary glosses with example sentences?

Method

Participants

The study began with 265 participants who were non-English major undergraduates at a national university of science and technology in southern Taiwan. However, only those who completed all phases including the pretest, immediate, and two delayed tests were counted. The subjects of interest were business and engineering majors from eight intact classes (one engineering class and one business class from each of four levels) in the university's General English program. Foreign language majors (German and Japanese), who were placed in the same classes with business majors, were excluded in order to focus on students who did not have a particular interest in foreign languages. A total of 135 participants remained in the study as a result. Table 1 gives a summary of the subjects in each gloss condition. Appendix A provides more detailed descriptions of the gloss conditions.

The proficiency levels are designated as Levels 1 to 4 from the lowest to the highest level. Initial language proficiency was assessed using an internationally recognized standardized language test (English Placement Test developed by the University of Michigan). Subjects signed a consent form to indicate their willingness to participate in this study and to authorize the use of their responses on the tests and questionnaires for research purposes. It was made clear that whether they participated or not would not influence their course grades.

Table 1. *Subjects and gloss conditions*

Gloss Condition	Level								Subtotal
	1		2		3		4		
	B	E	B	E	B	E	B	E	
L1-gloss-L2-ex	2	6	5	5	4	5	5	2	34
L1-in-text-gloss	3	6	7	5	6	6	3	2	38
L1-MG-gloss	2	5	2	3	4	6	4	4	30
No-gloss	1	6	3	5	5	6	2	5	28
Subtotal	8	23	17	18	19	23	14	13	135
Total	31		35		42		27		135

Note. B = Business majors; E = Engineering majors

Gloss Conditions

In the study, each subject was given a reading text with or without gloss supports. Subjects in each class were randomly assigned to one of the four conditions that included three experimental groups and one control group: (a) L1-gloss-L2-example, (b) L1-in-text-gloss, (c) L1-MG-gloss, (d) no-gloss.

Reading Passages and Glossed Words

The results of three criteria were employed to attempt to ensure appropriate text selection: (a) the *Fry Graph* readability formula, (b) a text difficulty evaluation form, and (c) the findings of a pilot study. First, five reading texts were collected that were no longer than 350 words from a number of English reading series textbooks, *Weaving it Together*, *Active Skills for Reading*, and *The Active Reader—Reading for Meaning*, recommended as useful reading supplements for General English courses by the teachers at the university, but not likely to have been seen by the subjects before. The five texts were scaled at 6, 8, 8, 9, and 10 respectively using the *Fry Graph*. Next, four General English instructors who taught Levels 1 to 4 were invited to read the five texts, mark down 5 to 15 unknown words within each of the five texts, and complete a text difficulty evaluation form devised for this study. Last, 3 weeks before the main study, a pilot study using subjects with a background similar to the subjects of the main study was conducted in an effort to determine the most suitable reading texts and verify the appropriateness of all the tests designed for the study. In addition, it also served to detect any potential problems before the official study took place.

After examining the results of the pilot study, we selected one text that was deemed to be not too easy or too difficult for subjects at Levels 1 to 3. The text was “Unusual Marriage Ceremonies” (207 words, *Fry Graph* 6) from *Weaving it Together 2* (Broukal, 1993, pp. 57–58). However, none of the five pre-selected texts was deemed suitable for Level 4 because of a perceived ceiling effect: The pilot study subjects were able to get 11 correct out of 16 on the pretest, 13 correct out of 16 on immediate vocabulary recall test, and 4 or 5 correct out of 5 on the reading comprehension test (see *Procedure*). The solution to this was to find another text that exceeded the readability scale of 10. With the help of two General English instructors, another text, “Addicted to Chocolate” (343 words, *Fry Graph* 12) from *Active Skills for Reading 2* (Anderson, 2003, pp. 6–7) was selected for Level 4.

Each of the two texts had 16 vocabulary items that were targeted for glossing under the three experimental conditions because they were thought to be unfamiliar to the subjects. As noted above, the words to be glossed were recommended by teachers familiar with the subjects' level of proficiency. Appendix A shows the four gloss conditions. The words highlighted in the no-gloss condition show the targeted vocabulary items. Appendices B and C list the complete set of 16 target items plus distractors.

Procedure

The study consisted of the following three phases carried out in three sessions. For the purpose of assessing subjects' retention of learned target words and forgetting pattern over time, two unexpected delayed vocabulary recall tests were arranged after the first phase of the experiment. Thus, subjects were not informed ahead of time that there would be any kind of delayed test or task for them to perform.

- *First Phase* (approximately 50 minutes)
 - (1) Pretest: Vocabulary test (10 minutes, 16 items)
 - (2) Reading a text (15 minutes)
 - (3) Posttest (See the Appendices B & C):
 - a) Reading comprehension test (5 minutes, 5 items)
 - b) Immediate vocabulary recall test (10 minutes, 16 items)
 - (4) Questionnaire about vocabulary gloss use for all subjects (10 minutes, 24 items)
(See Results and Discussion hereafter.)
- *Second Phase* (15 minutes): One week after reading
First delayed vocabulary recall test (16 items)
- *Third Phase* (15 minutes): Two weeks after reading
Second delayed vocabulary recall test (16 items)

In the first phase, for the pretest, subjects were given a piece of paper with 24 Chinese terms on the top of the page for 16 English target words shown on the bottom half of the page. Subjects indicated the meaning of the English words by choosing the Chinese equivalent translation presented above. After subjects turned in the pretest, they then read the assigned passage, which they returned before commencing the posttest. During this time, they could read the passage as many times as they wanted. They were also told to circle any words that they might not know within the passage that were not already glossed. The purpose of directing subjects to do this was to distract them from focusing exclusively on the glosses. The words circled were not intended to be analyzed as data for this study. In the posttest, subjects were first required to complete a reading comprehension test with five multiple-choice questions, and then an immediate vocabulary recall test, which presented 16 English sentences, each with a blank in it that subjects had to fill in by selecting the best term from a list of 24 English items. The 24 items included the 16 target words from the reading text and 8 words that did not appear in the reading text. These 24 items were the same as on the pretest. The format of the pretest was different from the vocabulary recall test in order to minimize transfer from the pretest to the posttest. The vocabulary test was followed by the questionnaire, which elicited information about subjects'

previous exposure to and personal views concerning the use and perceived effectiveness of glosses.

The content of the three vocabulary recall tests was the same: 24 vocabulary items on the top of the page and 16 sentences to complete on the bottom of the page. However, the order of the 16 sentences was varied for each test. (See the Appendices B & C for the vocabulary pretest and the immediate recall test for both passages.) It is noteworthy that each vocabulary test served a different purpose. The purpose of the pretest was to assess how many words subjects knew prior to reading. The immediate vocabulary recall test was used to find out whether reading with or without glosses facilitated subjects' vocabulary knowledge. The delayed recall tests were designed to investigate subjects' vocabulary retention over time in relation to the different glossing conditions.

Scoring consisted of giving 1 point for each correct response on all tests; that is, there were 5 points possible for the reading comprehension posttest and 16 points for each of the vocabulary tests. No points were deducted for wrong answers.

Results

Effects of Glosses on Immediate Vocabulary Recall and Reading Comprehension

Table 2 indicates the scores for the posttest, including reading comprehension (ReadComp) and immediate vocabulary recall test (VocTest). As can be seen in Table 2, reading comprehension for all four levels combined shows no gains for any of the experimental conditions over the control. In contrast, the VocTest 1 scores are significantly higher than the no-gloss control group as indicated by the *t* values, though the effects of the L1-MG-gloss condition are not quite as strong as the other gloss conditions.

Table 2. *Posttest results for combined subjects*

Test	M_{test}	$M_{\text{no-gloss}}$	$M_{\text{difference}}$	<i>t</i>	<i>p</i>
L1-gloss-L2-ex (<i>n</i> = 34)					
ReadComp (5)	2.97	2.61	0.36	1.112	.270
VocTest 1 (16)	9.94	5.48	4.46	4.441	.000***
L1-in-text-gloss (<i>n</i> = 38)					
ReadComp (5)	3.16	2.61	0.55	1.769	.082
VocTest 1 (16)	9.58	5.48	4.09	3.917	.000***
L1-MG-gloss (<i>n</i> = 30)					
ReadComp (5)	2.73	2.61	0.13	0.381	.705
VocTest 1 (16)	8.07	5.48	2.58	2.656	.011*

Note. M_{test} = mean of the test condition; $M_{\text{no-gloss}}$ = mean of the no-gloss condition (*n* = 33); $M_{\text{difference}}$ = mean difference between the test condition and the control condition. ReadComp = reading comprehension; VocTest = vocabulary recall test.

p* < .05. *p* < .01. ****p* < .001.

To answer the first research question, for the subjects at the four levels viewed as a whole, all of

the three experimental gloss conditions facilitated the subjects' vocabulary retention, but they did not help the subjects on reading comprehension. Three possible explanations for the observed positive effects of vocabulary glosses on the vocabulary recall test are suggested as follows. Two of them, the first and third explanations, involve greater depth of processing.

First, the L1-gloss-L2-example items on a separate page may have served as a prereading activity activating readers' word knowledge before they began to read. This process may have provided subjects with a certain amount of background knowledge of the text and at the same time focused subjects' attention on vocabulary glosses. Moreover, the subjects in the L1-gloss-L2-example group may have retained the vocabulary more effectively because they had gone through a deeper processing of the example sentences than merely glancing at simple glosses would have required (Laufer & Hulstijn, 2001). Second, the subjects in the L1 in-text gloss group may have done significantly better in vocabulary recall than the control group because they could refer to the in-text Chinese equivalents immediately during reading without being distracted by looking back and forth between words in the text and their glosses somewhere else on the page. Each of these explanations suggests how glossing could contribute positively to the acquisition of vocabulary. Third, the subjects in the L1-MG-gloss group can be viewed as engaging in three recurring tasks during reading, which also encourage deeper processing. These were identified by Hulstijn et al. (1996) in their study: Subjects read the text and encountered unfamiliar words. Then, they referred to the glosses at the bottom of the same page and noted the meanings. Keeping the meanings in mind, they then had to return to the text and recall what they had just read to match the meanings with the textual information. Although this reinforced their knowledge of the words and led to significantly improved retention over the control group, the L1 marginal glosses might have been somewhat less effective than the other two glossing conditions because of the disruption in the flow of reading, which hampered subjects' natural reading processing. (See the Conclusion and Implications section below for a summary of the ranking of the effectiveness of the three gloss conditions and Cheng, 2005, for in-depth discussion.)

As shown in Table 2, even though glosses aided in vocabulary acquisition or retention, no significant difference was found for reading comprehension between the control group ($p < .05$) and each of the three experimental conditions (L1-gloss-L2-ex, $p = .270$, L1-in-text-gloss, $p = .082$, and L1-MG-gloss, $p = .705$). Subjects in the three gloss conditions correctly answered approximately three out of the five questions (with means of 2.73 to 3.16) while those in the control group scored slightly lower ($M = 2.61$). The presence of vocabulary glosses might have diverted subjects' attention somewhat from the text and made them focus more on processing the word meanings, rather than on reading for the main idea or remembering specific details of the text. Alternatively, 16 unknown or difficult words may have been too many: Subjects may have focused so much of their cognitive resources on deciphering vocabulary that they had few remaining mental resources to devote to comprehending, integrating, and remembering what they read. It is also possible that five questions about the reading passage were too few to discriminate effectively between subjects.

Language Proficiency and Glosses

In this section, we turn to the second research question, which addresses whether language

proficiency is a factor in the effectiveness of glosses. A one-way analyses of variance (ANOVA) was adopted to assess whether there were significant differences between the four levels of proficiency for the four different gloss conditions. Scheffé Post-Hoc Multiple Comparisons were further computed to determine where the difference among the subgroups was if significant results were found. We acknowledge that the size of each subgroup at the four levels is quite small, which can not only affect whether significant differences are found, but also the generalizability of our findings. Replication of this study with larger subgroup sizes should yield more robust results that could be generalized with more confidence.

To answer the second research question, none of the four kinds of gloss conditions significantly facilitated Levels 1 to 4 subjects' reading comprehension. ANOVA and Scheffé tests both indicate that for the reading comprehension test there was no statistically significant difference between gloss types and each individual level. Nevertheless, Level 1 subjects' reading performance with glosses was enhanced by a small amount. As can be seen by the means in Table 3, all three gloss types improved reading comprehension somewhat over the no-gloss condition. Nonetheless, the highest mean score was less than three correct answers out of the five questions. Perhaps the text was too difficult for them, and the Level 1 subjects were able to derive only limited benefit from the glosses in their reading comprehension. We note, however, that the results approach significance ($p = .089$).

Table 3. Summary of posttest means for subjects from Levels 1 to 4

Test	Condition							
	L1-gloss-L2-ex		L1-in-text-gloss		L1-MG-gloss		No-gloss	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Level 1	<i>(n = 8)</i>		<i>(n = 9)</i>		<i>(n = 7)</i>		<i>(n = 7)</i>	
ReadComp	2.50	1.309	2.67	1.323	1.71	0.488	1.43	0.787
VocTest 1	5.25	2.866	4.67	4.213	4.00	4.163	2.71	2.498
VocTest 2	3.25	2.915	4.22	3.270	4.43	4.962	2.43	1.272
VocTest 3	4.88	3.091	4.78	3.598	3.00	4.163	2.71	1.890
Level 2	<i>(n = 10)</i>		<i>(n = 12)</i>		<i>(n = 5)</i>		<i>(n = 8)</i>	
ReadComp	2.90	1.524	3.25	1.288	2.80	1.304	2.63	1.506
VocTest 1	11.90	3.604	11.17	4.914	8.40	2.881	5.50	3.964
VocTest 2	10.40	4.858	10.75	3.388	6.40	2.881	5.88	3.980
VocTest 3	11.10	3.872	10.83	3.407	7.60	3.507	6.00	4.440
Level 3	<i>(n = 9)</i>		<i>(n = 12)</i>		<i>(n = 10)</i>		<i>(n = 11)</i>	
ReadComp	3.89	.782	3.58	1.165	3.70	1.252	3.36	1.206
VocTest 1	12.56	3.609	12.00	2.923	9.90	3.604	7.27	3.036
VocTest 2	11.44	3.539	10.33	3.651	8.20	4.686	5.27	3.524
VocTest 3	11.67	3.536	10.67	4.438	9.90	4.630	6.82	5.741
Level 4	<i>(n = 7)</i>		<i>(n = 5)</i>		<i>(n = 8)</i>		<i>(n = 7)</i>	
ReadComp	2.43	1.134	2.80	1.095	2.38	1.061	2.57	1.272
VocTest 1	9.14	4.220	8.80	4.438	9.13	2.997	5.43	4.117
VocTest 2	8.14	3.761	7.80	3.271	8.25	3.576	6.00	3.266
VocTest 3	8.71	3.200	7.80	3.114	8.63	3.701	5.29	3.729

Note. ReadComp = reading comprehension; VocTest = vocabulary recall test.

This conclusion is similar to the one Lee and Good (2003) reached though they provided glosses for an engineering text rather than a more general EFL text. In order to decrease engineering students' processing load when reading scientific textbooks, they examined three conditions: a simplified English text, an original English text with Chinese in-text glosses, and the original unmodified English text as control. They found that experimental groups did not outperform the control group in the reading comprehension test; they suggested this might be "due to overall weakness in reading ability, which is not remediable merely through vocabulary manipulation" (p. 18). Watanabe (1997) also suggested that even if explanations were inserted for unfamiliar words, and the explanations were comprehensible, students with small vocabulary size would still not be able to make effective use of the glosses (p. 303).

As for vocabulary retention, in Table 3 it can be seen that the means for each of the gloss conditions are higher than for the control group. Table 4 shows that significant differences within levels were found only for Level 2 and Level 3 subjects. Cheng (2005) presented the statistical analyses showing that L1-gloss-L2-example and L1-in-text-gloss are the most effective glossing types for subjects at Levels 2 and 3; however, for Levels 1 and 4 subjects, no statistically superior ways of presenting the glosses were found. From these findings we might conclude that language proficiency is indeed a factor in gloss effects, but not all levels benefit equally. We attribute the Levels 2 and 3 subjects' improved performance to the vocabulary support. Several possible reasons for the lack of Levels 1 and 4 improvements are discussed below.

Table 4. *Statistical significance of the 4 levels by ANOVA*

Test	Level			
	1	2	3	4
ReadComp	.089	.793	.768	.921
VocTest 1	.561	.011*	.003***	.242
VocTest 2	.645	.024*	.004***	.596
VocTest 3	.448	.021*	.119	.241

Note. * $p < .05$. ** $p < .01$. *** $p < .001$.

First, although L1 (Chinese) glosses were presented, which were assumed to be direct and clear between words and meanings, those students may still have had a hard time remembering the meanings of the glossed words and applying them to a difficult reading text to aid them either in their comprehension of the text or to help them acquire the vocabulary. Hence, glosses may not be sufficient to ensure either comprehension or vocabulary acquisition for low proficiency technological university students. Students' performance may depend more on their overall English reading ability and text appropriateness for that ability level. This deserves further scrutiny.

Secondly, for Level 4 subjects, the reason that no significant effects of vocabulary glosses (see Table 4) were found may also be accounted for by text difficulty. Even though the Level 4 experimental groups did outperform the control group on each vocabulary test (see mean scores in Table 3), there were no significant differences in the vocabulary tests. The text (*Fry Graph* rating at Grade 12 level) might have been beyond subjects' comprehension capacity; hence, glosses did not efficiently facilitate subjects' understanding of the whole passage or allow them to acquire the target words. As described above, the text "Addicted to Chocolate" was selected

after carrying out the pilot study. The substitution was made because the reading comprehension scores shown in the pilot study were quite high, so the original pilot-study text seemed to be too easy for Level 4 students. However, contrary to expectation, the new text may have been too difficult and resulted in little positive effect of vocabulary glosses on reading comprehension or vocabulary recall in the main study. Jacobs et al. (1994), as discussed above, cited text difficulty (a text being either too difficult or too easy) as a possible explanation for why their high-proficiency subjects did not show treatment effects. Thus, as we can see in Table 3, the mean scores for the reading comprehension may imply that the text for Level 4 subjects was too difficult: None of the mean scores for any of the gloss conditions or the control group reaches 3 points out of a possible 5. The matching of the readability level of a text with the language ability of subjects is likely a key factor in the effectiveness of gloss support. Another plausible explanation for Level 4 subjects' poor performances might be that Level 4 had the smallest number of subjects ($n = 27$) among all groups (Level 1 = 31, Level 2 = 35, Level 3 = 42); more robust results might be obtained with a larger sample size.

Third, the positive effects found on vocabulary recall for the Level 2 and Level 3 groups reveal that although the text may have been too hard for Level 1 students, it appears to have been appropriate for students at the low-intermediate or intermediate levels: Their vocabulary performance scores improve significantly. Even though there was facilitation in vocabulary acquisition and retention, reasons for the failure in improving reading comprehension are elusive. Possible explanations may include the passage being too short and having only five questions on the reading comprehension was too few to effectively judge or to evaluate subjects' comprehension ability. Most of the Levels 2 and 3 subjects got nearly 3 to nearly 4 points out of 5 on the reading comprehension. Level 1 subjects got 1+ to 2+ points out of 5; on the more difficult reading passage, Level 4 subjects got between 2 and 3 points out of 5. In other words, about 40% to 60% of subjects got the right answers. It is suggested that in future studies more questions should be provided to assess students' comprehension ability, perhaps as many as eight to ten questions; this will also likely necessitate the use of a longer reading passage.

Language Forgetting Patterns and Vocabulary Retention

The means for each test are presented in Table 5.

Table 5. Subjects' performance in four gloss conditions on vocabulary tests

Condition	Pretest		Voctest 1		Voctest 2		Voctest 3	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
L1-gloss-L2-ex ($n = 34$)	7.44	3.42	9.94	4.49	8.53	4.90	9.29	4.28
L1-in-text-gloss ($n = 38$)	7.71	3.06	9.58	4.93	8.68	4.25	8.95	4.43
L1-MG-gloss ($n = 30$)	7.03	2.57	8.07	4.05	7.03	4.32	7.57	4.73
No-gloss ($n = 33$)	6.75	2.82	5.48	3.66	4.97	3.39	5.42	4.50

To investigate longer-term vocabulary retention, which is the focus of the third research question, we look at forgetting patterns. The data from the four vocabulary tests for the combined four proficiency levels were submitted to a simple-factorial ANOVA. The dependent variable in each ANOVA was the individual vocabulary test under investigation, and the independent variable was the type of gloss conditions. The ANOVA results, as can be seen in Table 6, indicate that significant differences were found on VocTest 1 ($p < .001$), VocTest 2 ($p = .001$), and VocTest 3

($p = .002$). However, no significance was found on the Pretest ($p = .560$), which is what we would expect if the students in each subgroup or gloss condition are comparable.

Table 6. ANOVA for vocabulary tests among subjects at four gloss conditions

Test	Subgroup	SS	df	MS	F	p
Pretest	Between	18.701	3	6.234	0.690	.560
	Within	1183.225	131			
VocTest 1	Between	416.745	3	138.915	7.358	.000***
	Within	2473.255	131			
VocTest 2	Between	304.864	3	101.621	5.611	.001***
	Within	2372.617	131			
VocTest 3	Between	313.219	3	104.406	5.196	.002**
	Within	2632.381	131			

Note. * $p < .05$. ** $p < .01$. *** $p < .001$.

To investigate the performance of the four subgroups on VocTests 1, 2 and 3, three Post-Hoc Multiple Comparisons (Scheffé Tests) were also computed. The results in Table 7 indicate that significant differences were found in both the L1-gloss-L2-example and the L1-in-text-gloss conditions, but not in the L1-MG-gloss and the no-gloss conditions on VocTests 1, 2, and 3 (see Cheng (2005) for in-depth discussion). The graph of the means for each test (based on Table 5) shown in Figure 1 gives us a clearer picture of the forgetting pattern in the current study.

Table 7. Post-hoc multiple comparisons of gloss conditions on vocabulary tests

Test	M	p	
VocTest 1	9.94 (L1-gloss-L2-ex)	5.48 (No-gloss)	.001***
	9.58 (L1-in-text-gloss)	5.48 (No-gloss)	.002**
VocTest 2	8.53 (L1-gloss-L2-ex)	4.97 (No-gloss)	.010*
	8.68 (L1-in-text-gloss)	4.97 (No-gloss)	.005*
VocTest 3	9.29 (L1-gloss-L2-ex)	5.42 (No-gloss)	.007*
	8.95 (L1-in-text-gloss)	5.42 (No-gloss)	.015*

Note. * $p < .05$. ** $p < .01$. *** $p < .001$.

As shown in Figure 1, the means for the subjects' vocabulary recall tests in the three experimental conditions rise from the pretest to the immediate posttest VocTest 1, decline on VocTest 2 and rise slightly on VocTest 3. For the no-gloss group, there is an apparent decrease between the pretest ($M = 6.75$) and VocTest 1 ($M = 5.48$) and another decrease in VocTest 2 ($M = 4.97$). We can see that their retention declined between the immediate and first delayed vocabulary recall tests by 0.51 points, which was also observed in past studies (Huang, 2003; Watanabe, 1997). However, between the first and second delayed recall tests, a slight increase in retention was found for all conditions. In other words, subjects' retention did not decline between the first and second delayed recalls, but rather showed a certain degree of growth including for the control group. To be more specific, the Scheffé Tests in Table 7 revealed that declines and rises were statistically significant only in L1-gloss-L2-ex and L1-text-gloss conditions; no significance was found in L1-MG-gloss and no-gloss conditions.

A possible explanation for this unexpected rise from the first to the second delayed recall tests is that the vocabulary recall tests serve as a review or reinforcement for subjects. In other words, subjects look at the same words in order to choose correct answers 1 week and 2 weeks after the initial exposure. Such periodic review is one way to enhance retention over time, as suggested by Pimsleur (1967) above. We also see a correlation between each kind of gloss and retention, though even the control group exhibited an increase between VocTest 2 and VocTest 3, which is what we would expect if repeated encounters with the vocabulary serve as reinforcement.

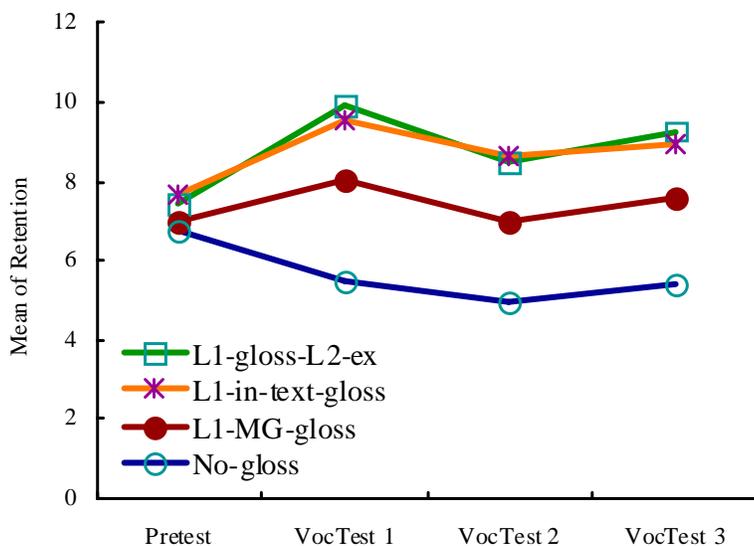


Figure 1. Vocabulary forgetting patterns. Each test has 16 items or points. Pretest was administered prior to reading (first phase); VocTest 1, following reading (first phase); VocTest 2, 1 week after reading (second phase); and VocTest 3, 2 weeks after reading (third phase).

Nevertheless, if retesting as reinforcement is the only factor, we might wonder why Huang's (2003) forgetting pattern does not show a similar growth. In addition to the benefit arising from repeated review, the slight increase between the first and the second delayed vocabulary recall in retention might be accounted for by two alternative explanations. First, we note that this forgetting pattern is similar to that in Watanabe's (1997, p. 302) study, but he found a much larger increase, which was statistically significant on the second delayed test. He speculated that the increase between the first and the second delayed tests might be because some students might have looked up the words they saw on the tests, or they might have encountered some of the target words in subsequent reading (e.g., at home, in other classes). If this explanation is valid, we might say that the increase is attributable to external factors, and may not be regarded as retention of what they learned from reading or through glosses. Alternatively, we might speculate that perhaps the three kinds of vocabulary glosses used in the present study were somehow more effective than the three in Huang's study (see the Introduction for a description) in promoting vocabulary retention. However, since the gloss types in the present study and Huang's are not the same, it is difficult to compare them to determine whether this proposal is a viable explanation.

Questionnaire: Previous Experience with Vocabulary Glosses

Descriptive statistics were computed for the means of each question item on the questionnaire addressing the fourth research question to understand learners' attitudes about various kinds of glosses. The questionnaire was written in Chinese and consisted of two parts. The eight questions in Part A ask about subjects' past experience with seeing vocabulary glosses at two different stages: pre-university stage (in a vocational high school or five-year junior college), and university stage (at their current university). Subjects chose frequency rates numbering from 1 (*never*) to 5 (*usually or always*) to answer the questions. In Part B, eight additional questions asked subjects about their opinions concerning the effectiveness of these vocabulary glosses for reading comprehension and vocabulary learning. Subjects were provided with clear definitions and examples of the three gloss types of interest in this study. In addition, three other types of glosses that often appear in EFL textbooks were also listed with examples on the first page of the questionnaire so subjects could be questioned about them also. Subjects were required to look at them before they proceeded to Parts A and B. As noted in the footnote in the introduction, the word for *gloss* and *glosses* in Chinese means *vocabulary explanations* and therefore is immediately understandable to all Chinese. The examples of the types were provided so that subjects would know how these could be presented in a text. This questionnaire was given to all subjects immediately after they completed the reading comprehension and vocabulary recall test in the first phase.

Questionnaire Part A.

Responses for items 1 to 8 are discussed one by one in this section.

(1) *Have you ever seen vocabulary glosses in your English textbooks?* As many as 40.9% of subjects indicated that they had "seldom" seen glosses in their university English textbooks. In addition, 24.5% responded that they had "sometimes" seen glosses before entering the university. The former result may seem rather surprising, but this might be explained by the classroom textbooks used. At the university surveyed, general English instructors who teach Levels 1 to 4 choose their own textbooks according to different teaching objectives. The use of glosses is mostly confined to reading textbooks. In courses emphasizing conversation, sentence pattern practice, or writing only, subjects might not see glosses in their textbooks. It is also highly unlikely that glosses appear in any of their content-area textbooks, which are written for native English speakers and are adopted by Taiwan instructors for their engineering or business course.

(2) *Have you ever seen vocabulary glosses in EFL magazines or English newspapers?* Subjects at both pre-university and university stages "sometimes" had seen glosses provided in EFL magazines or newspapers. The use of vocabulary glosses has become so common during recent years that the subjects could hardly miss it in language-learning media available on magazine racks all over Taiwan.

(3) *Have you ever seen L1 glosses plus L2 example sentences near the text?* Subjects had "seldom" seen these glosses at either stage. From our experience, these kinds of glosses and L1 marginal glosses appear in some senior high school English reference books (e.g., Cheng, 2003). Since most of the subjects in the study were vocational high school or 5-year junior college

graduates, they might have never used such books before, and so had few opportunities to see examples of L1-gloss-L2-ex. The English textbooks they use in the university now are usually provided with L2-gloss-L2-ex, which further explains why subjects “seldom” saw L1-gloss-L2-ex in the two stages. Of course, they may not read many English learning magazines, where L1-gloss-L2-ex is fairly common.

(4) *Have you ever seen L1 marginal glosses below the text?* The highest percentages of subjects seeing this kind of gloss in the past and at the present are 26.4% (sometimes) and 38.4% (seldom). L1 marginal glosses are popular in junior and senior high school textbooks and EFL learning magazines in Taiwan (e.g., *Time Express*). Subjects may have been exposed to texts with these kinds of glosses in the past but it appears they rarely saw them in their current university textbooks, where the glosses could be L2 marginal glosses instead of L1.

(5) *Have you ever seen L1 marginal glosses in the right margin?* Subjects replied that they had “seldom” seen L1 marginal glosses next to the text at the two stages. It is true that this kind of gloss does not often occur and most are L2 or L1 glosses with L2 example sentences. However, we have seen them provided in *Let’s Talk in English*, an English learning magazine, in a word box next to the text.

(6) *Have you ever seen L1 in-text glosses in the text?* The results reveal that subjects “seldom” saw these glosses during either stage. Compared with the glosses mentioned above, L1 in-text glosses are quite rare. A possible explanation for this situation is that not every learner needs to use all of the provided glosses. Writers may also be aware that if the glosses are embedded in a text, they might affect learners’ attention and natural reading process. In contrast, if glosses are displayed outside of the text (e.g., marginal glosses), learners can decide whether to consult them or not.

(7) *Have you ever seen L2 glosses plus L2 example sentences near the text?* Subjects replied that they had “seldom” seen this kind of gloss in the past or at the university. These glosses can be found in some university reading textbooks (e.g., *Reading Advantage* by Malarcher, 2004). Since general English instructors at the university that we surveyed have some freedom in choosing different textbooks for classroom use, subjects may not have been exposed to textbooks with these kinds of glosses.

(8) *Have you ever seen L2 marginal glosses below the text?* Some subjects reported having “seldom” seen these types of glosses in the past, while others “sometimes” have seen them at the university stage. L2 glosses are usually provided in EFL or ESL oriented textbooks. There are textbooks (e.g., *Active Skills for Reading* by Anderson, 2003) that have L2 marginal glosses right below the text. L2 glosses of any kind are most likely to be found in textbooks with international markets like Anderson’s; no language-specific adaptation is done, which is what would be necessary if L1 glosses were used.

To sum up, in Part A, a majority of subjects replied that they had “seldom” or only “sometimes” seen glosses at either of the two stages of their education. However, the reason why subjects had “seldom” encountered some types of glosses is perhaps because these subjects are non-English majors, and so they may not have had much exposure to many L2 learners’ magazines or

textbooks.

Questionnaire Part B.

Eight questions asked about subjects' perspectives concerning the use of vocabulary glosses during the experiment and beyond. Items 1 to 4 were designed to investigate how many glossed words subjects knew prior to reading. Items 5 to 8 were used to find out which kind of glosses subjects believed would help them most or least in reading or learning new words.

(1) *Among the 16 glossed words, how many of them did you know before you read the text?* Many students (48%) replied that they knew 4–7 of the glossed words before reading. In other words, most of the subjects did not know even half of the target words before reading, which suggests that the glossed words were selected appropriately. If subjects recognized none of the words, the text might have been too difficult for them; if they recognized too many of the words, the text could have been too easy. This result is consistent with the scores on the vocabulary pretest. On average, subjects got 7 out of 16 on the pretest, which accords with their self-assessment (see Figure 1).

(2) *What kind of glosses did you just see while you were reading the text?* Item 2 asked subjects to recall which kind of vocabulary glosses, if any, accompanied the texts they had just seen as they read the passage. The expected number for each of the first four answer options (no-glosses, L1-gloss-L2-ex, L1-in-text-gloss, L1-MG-gloss, I forgot/I am not sure what I had seen) should be roughly 25%, but subjects' answers varied a lot. It can be inferred that subjects did not pay close attention to or remember what kinds of glosses they had seen in the text.

(3) and (4) *Do you think vocabulary glosses were helpful for text comprehension / learning new words during reading?* Because the control group did not receive any treatments, they were instructed to skip these two questions. The results for both items showed that 75% of subjects in the experimental groups responded positively saying vocabulary glosses were helpful for text comprehension and vocabulary learning during the study. That means most of the subjects hold a positive attitude toward vocabulary glosses.

For items 5 through 8, we are exploring subjects' subjective evaluations about which of these glosses they thought would be most or least helpful to them in their reading or in learning new vocabulary. Both the control groups and experimental groups answered these questions since the solicited opinion was not restricted to what they had just seen on the posttest. Of course, our experimental design allowed each subject to see only one of the three gloss types, and the control group did not see any of them. Subjects' responses reflect either their previous experience with glosses, which as we saw in Part A was limited, or their speculations about which type might be most helpful to them. We are attempting to determine whether subjects' opinions correlate with their actual performance. If a subject failed to provide a response to any item in this section, it was coded as *no response*.

(5) and (6) *Which kind of the following glosses do you think could best help you read texts / learn new words?* For Item 5, we can see that 34% of the subjects responded that L1 in-text glosses would be helpful for text comprehension, which is consistent with the findings in the present

study indicating that the L1 in-text glosses ($M = 3.16$) were the most effective ones for text comprehension (see Table 2).

L1 glosses plus L2 example sentences	L1 in-text glosses	L1 marginal glosses	No Response
27%	34%	22%	17%

In response to Item 6, 42% of subjects thought that L1 glosses plus L2 example sentences could best help them learn new words. This finding might lead us to expect this gloss type would result in significantly better performance; however, we found that the mean scores on the immediate vocabulary recall test showed L1 in-text glosses ($M = 8.68$) were slightly higher than the L1-gloss-L2-example ($M = 8.53$), though the difference was not significant (see Table 5).

L1 glosses plus L2 example sentences	L1 in-text glosses	L1 marginal glosses	No Response
42%	24%	16%	18%

(7) and (8) *Which kind of the following glosses do you think could least help you read texts / learn new words?* These items, like the previous two, call on subjects to respond based on their experience or their speculation concerning which might be most helpful. For Item 7, as indicated below, 36% of subjects replied that they thought L1 marginal glosses would be the least helpful for text comprehension, which is consistent with our findings ($M = 2.73$) (see Table 2).

L1 glosses plus L2 example sentences	L1 in-text glosses	L1 marginal glosses	No Response
17%	30%	36%	17%

For Item 8, 40% believed L1 in-text glosses would be the least helpful gloss type for learning new words, as is shown below.

L1 glosses plus L2 example sentences	L1 in-text glosses	L1 marginal glosses	No Response
13%	40%	28%	19%

By comparing the means of the immediate vocabulary recall test, we discovered that despite the greatest number of subjects' opinion that L1 marginal glosses would be the most effective, this was not supported by our findings ($M = 8.07$; see Table 2), though the difference between gloss conditions was not significant. However, since some subjects selected each of the gloss types as potentially most helpful, and no clear majority selected a single one, we can only conclude that the subjects are not sensitive to which gloss type might yield the best results.

The original purpose for including Items 5 through 8 on the questionnaire was twofold: One was to serve as a cross-reference to allow the comparison of these answers of these four questions with the results of our vocabulary tests to see if subjects would be able to identify the gloss type that the results of the vocabulary tests showed to be most helpful. From the above findings, it does not appear that they clearly prefer one type over another. The other purpose was to give textbook writers and university instructors some ideas about learners' perceptions of vocabulary

glosses. Students have no choice about what kinds of glosses, if any, will be presented to them. If one kind of gloss is found to be or perceived by subjects to be more useful, writers and instructors could use that type more in the textbooks they write or select. In the present study, we observe that the mean scores on VocTests 1, 2, and 3 for L1-gloss-L2-example and the L1-in-text-gloss conditions were significantly higher than the no-gloss conditions. According to the means given in Table 2, it appears that the most effective glosses to the least effective ones were: the L1-gloss-L2-ex ($M = 9.94$), the L1-in-text-gloss ($M = 9.58$), and the L1-MG-gloss ($M = 8.07$). This seems somewhat surprising because according to the questionnaires, the most-frequently-seen type of gloss seen by subjects were the marginal glosses; however, these were not shown to be as effective as the other two gloss conditions. This finding could serve as an important reminder for L2 textbook writers to include other kinds of glosses besides the more popular marginal gloss.

To sum up, results in Part B regarding subjects' opinions about the use of vocabulary glosses indicate that subjects think they are useful. The answers for Items 1 to 4 serve as references for the present study. Responses for Items 5 to 8 ask about subjects' perception of relative benefit of gloss types. We have compared their reported perceptions with the test results to see if subjects are accurate in their perceptions of the relative merits of gloss types. The findings from the questionnaire show that their perceptions do not match their performance, but inasmuch as the differences are not significant we cannot conclude whether they are able to distinguish between the benefits of the three types of glosses. A larger sample size would make any findings more robust.

Subjects view the benefits of glosses positively as shown in Part B of the questionnaire, and they have had relatively few opportunities to encounter and use them according to Part A. This suggests that providing students with more texts that are glossed would be welcomed by them.

Conclusion and Implications

The findings for the four research questions investigated in this study are summarized as follows:

1. Combining all proficiency levels for the four conditions, it was found that subjects in the three experimental gloss conditions outperformed subjects in the no-gloss condition on all vocabulary recall tests. Glossing, however, did not show significant effects for reading comprehension for any of the levels of proficiency, even though there was a marginal effect ($p = .089$, see Table 4) for Level 1.
2. We found that language proficiency was indeed a factor in gloss effects, but not all levels received equal benefit. As noted in the previous conclusion, none of the four kinds of gloss conditions significantly facilitated Levels 1 to 4 subjects' reading comprehension. As for vocabulary recall, significant differences were found only for Level 2 and Level 3 subjects; however, for Levels 1 and 4 subjects, no statistically superior ways of presenting the glosses were found. The ranking of the effectiveness of gloss type on vocabulary retention appears to be L1-gloss-L2-ex > L1-in-text-gloss > L1-MG-gloss > no-gloss (see Cheng, 2005, for additional discussion).

3. Examining the longer-term vocabulary retention as represented by the vocabulary forgetting pattern, we saw that subjects obtained relatively higher scores on the immediate vocabulary recall test after their first encounter with the text and the accompanying glosses. We also saw a decrease on the first delayed vocabulary recall test, and then a somewhat surprising slight increase on the second delayed vocabulary recall test. We suggested that the presence of glosses in the passages and repeated review might account for this.

4. Subjects' opinions about the value of vocabulary glosses were positive: 75% of them believed that vocabulary glosses could increase English reading comprehension and vocabulary learning. This study shows that their impressions about the value of glosses for comprehension were not as accurate as those for vocabulary. In addition, subjects' encounters with vocabulary glosses ranged from "seldom" to "sometimes," which suggests that glossing may not be widely used, at least in the materials non-English majors are exposed to.

Additional research to further enhance our understanding of the effect of vocabulary glosses on reading comprehension and vocabulary learning and how gloss types relate to proficiency levels is recommended. Future research could include the following areas. First, this study could be replicated with a larger subject sample in order to generalize the findings to a larger population. Many university graduate and undergraduate courses in Taiwan require students to use English textbooks, which may, in fact, be too difficult for them to read effectively. If glossing is confirmed as a useful technique for enhancing reading ability, as has been suggested by some other studies (Davis, 1989; Rott et al., 2002; Watanabe, 1997), it may be possible to encourage publishers to produce EFL versions of their textbooks or to provide vocabulary support supplements or bilingual glossaries that would be more useful to students than the currently available textbooks with no reading support.

Second, we cited Cheng's (2005) findings on the correlation between individual gloss types and English proficiency level. Moreover, the reason why not all levels of students benefit from L1 glosses needs to be investigated further. Unlike the current study, Jacobs et al. (1994) found that participants at a higher proficiency level seemed to benefit more from glosses perhaps because they possessed sufficient L2 competence and knew how to use the glosses effectively. They also pointed out that a text should be close to learners' current L2 knowledge. A text that is too difficult might result in the need for a large number of glosses, which could impede the flow of reading. A large number of unknown words also could discourage students from trying to read a text at all. The importance of matching text difficulty to students' language proficiency is an important factor in assessing the value of glosses in empirical studies, but it is even more important when instructors select English textbooks for content courses. Investigating text difficulty and other factors that influence how teachers choose content area textbooks could help teachers select textbooks at a level engineering and business students can make most effective use of.

It should be pointed out that for empirical studies to be meaningful, texts should also not be too easy, thus making the provision of glosses unnecessary. For this reason, careful text selection is

crucial for subjects who are at different proficiency levels. Even though the current study employed teacher surveys and a carefully designed pilot study as well as a readability formula to choose our texts, we subsequently had to select an additional, more difficult text for the highest proficiency level. Protocols for identifying suitable texts and matching text difficulty and subjects' ability at all levels should be investigated further.

Reading in English is perhaps the most important of the four language skills needed by students of any major and by future professionals, especially in business and engineering. They will probably use English to keep up with the latest trends and technology via the Internet or by reading professional journals and magazines. All of this reading presents a significant challenge to students with limited proficiency. Glossing has been shown to be effective in enhancing vocabulary acquisition and retention in this study and reading comprehension in other studies. The more widespread use of glosses in texts could provide L2 readers with the support they need to improve their vocabulary and reading skills so that they may become more independent and successful readers.

Acknowledgments

We would like to thank the editors of the special issue on reading and vocabulary (20, 2) for insightful comments on an earlier draft of the manuscript and two anonymous reviewers for their valuable advice about the content and organization of this article. We are responsible for any deficiencies remaining after their kind input.

Note

1. Glosses are called *zihuijieshi* (字彙解釋) “vocabulary explanations” in Mandarin Chinese and are often used to make new vocabulary in reading passages more accessible to readers in English learning magazines in Taiwan. The Chinese term is completely transparent, which cannot be said for the English word *gloss*, which may not be familiar to very many native English speakers at all.

References

- Anderson, N. J. (2003). *Active skills for reading 2*. Boston, MA: Thomson & Heinle.
- Broukal, M. (1993). *Weaving it together 2*. Boston, MA: Heinle & Heinle.
- Cheng, Y. H. (2003). *Lungteng senior high school English III (Book A)*. Taipei: Lungteng Cultural Co.
- Cheng, Y. H. (2005). *Effectiveness of using vocabulary glosses to enhance technological university business and engineering majors' EFL reading comprehension and vocabulary learning*. Unpublished master's thesis, National Kaohsiung First University of Science and Technology, Taiwan.
- Davis, J. N. (1989). Facilitating effects of marginal glosses on foreign language reading. *The Modern Language Journal*, 73(1), 41–49.

- Gettys, S., Imhof, L. A., & Kautz, J. D. (2001). Computer-assisted reading: The effect of glossing format on comprehension and vocabulary retention. *Foreign Language Annals*, 34(2), 91–106.
- Huang, Y. C. (2003). *The effects of vocabulary glosses and example sentences on junior high school EFL students' reading comprehension and vocabulary learning*. Unpublished master's thesis, National Cheng-Kung University, Taiwan.
- 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., & Hong, F. C. (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.
- Joyce, E. E. (1997). Which words should be glossed in L2 reading materials? A study of first, second, and third semester French students' recall. *Pennsylvania Language Forum*, 69, 58–64.
- Karp, A. S. (2002). *Modification of glosses and its effect on incidental L2 vocabulary learning in Spanish*. Unpublished doctoral dissertation, University of California, Davis.
- Laufer, B., & Hulstijn, J. (2001). Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. *Applied Linguistics*, 22(1), 1–26.
- Lee, M., & Good, R. L. (2003). Making English reading easier for engineering students: Exploring the effects of reducing vocabulary processing load on EFL learners' comprehension of scientific texts. *Proceedings of the 2003 Conference on Foreign Language Application and Research*, 1–23.
- Malarcher, C. (2004). *Reading advantage* (2nd ed.). Singapore: Thomson & Heinle.
- Myong, H. K. (2005). Glosses, comprehension, and strategy use. *Reading in a Foreign Language*, 17(2), 125–143.
- Paribakht, T. S., & Wesche, M. (1997). Vocabulary enhancement activities and reading for meaning in second language vocabulary acquisition. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition: A rationale for pedagogy* (pp. 174–200). Cambridge, England: Cambridge University Press.
- Pimsleur, P. (1967). A memory schedule. *The Modern Language Journal*, 51(2), 73–75.
- Rott, S., Williams, J., & Cameron, R. (2002). The effect of multiple-choice L1 glosses and input-output cycles on lexical acquisition and retention. *Language Teaching Research*, 6(3), 183–222.
- Watanabe, Y. (1997). Input, intake, and retention: Effects of increased processing on incidental learning of foreign language vocabulary. *Studies in Second Language Acquisition*, 19(3), 287–307.

Appendix A

Examples of the Four Gloss Conditions

(1) L1-gloss-L2-example

paradise	天堂	Taiwan is a fantastic sightseeing <i>paradise</i> .
specialty	特產	You can try many local <i>specialties</i> , experience exotic cultures, and see beautiful scenery.

(2) L1-in-text-gloss

Taiwan is a fantastic sightseeing <i>paradise</i> (天堂). You can try many local <i>specialties</i> (特產), experience <i>exotic</i> (異國的) cultures, and see beautiful <i>scenery</i> (風景).

(3) L1-MG-gloss (below the text)

paradise	天堂	exotic	異國的	skyscraper	摩天大樓
specialty	特產	scenery	風景	spot	景點

(4) No-gloss

Taiwan is a fantastic sightseeing <i>paradise</i> . You can try many local <i>specialties</i> , experience <i>exotic</i> cultures, and see beautiful <i>scenery</i> .

Appendix B

Vocabulary Pretest for Levels 1 to 3 Subjects ("Unusual Marriage Ceremonies")

1 島	7 複雜的	13 新娘	19 刺青
2 典禮	8 結婚	14 罕見的	20 新婚夫婦
3 宗教的	9 教堂	15 舉行; 發生	21 燕尾服
4 西穀米	10 敵人	16 單獨的/地	22 吐
5 牽(手); 結合	11 新郎	17 招待	23 不幸的
6 蜜月	12 公民	18 一口	24 靈魂

Select the best item and write down a number (1–24) on the blank below.

- | | | | |
|----------------|-------|---------------|-------|
| 1. marriage | _____ | 9. join | _____ |
| 2. ceremony | _____ | 10. spit | _____ |
| 3. complicated | _____ | 11. mouthful | _____ |
| 4. bride | _____ | 12. enemy | _____ |
| 5. tattoo | _____ | 13. spirit | _____ |
| 6. unusual | _____ | 14. sago | _____ |
| 7. take place | _____ | 15. newlyweds | _____ |
| 8. bridegroom | _____ | 16. island | _____ |

Immediate Vocabulary Recall Test for Levels 1 to 3 Subjects

1 marriage	7 complicated	13 bride	19 tattoo
2 ceremony	8 island	14 unusual	20 newlyweds
3 religious	9 church	15 take place	21 tuxedo
4 sago	10 enemy	16 alone	22 spit
5 join	11 bridegroom	17 reception	23 unlucky
6 honeymoon	12 civil	18 mouthful	24 spirit

Select the best item and write down a number (1–24) on the left blank below.

1. ___ Some young people like to have their arms or bodies _____ed.
2. ___ I invited my parents to come to my graduation _____ in June.
3. ___ The _____ will go to Hawaii for their honeymoon.
4. ___ Do not _____ food or other items out of your mouth.
5. ___ They are going to get married. They will have a church _____.
6. ___ Singapore and Taiwan are _____ countries.
7. ___ His name is so _____. I have never heard of it.
8. ___ The question is so _____. It's not simple at all.
9. ___ The _____ kissed the bride. He looked so happy.
10. ___ The old man _____ed their hands together.
11. ___ My aunt just got married. She is the most beautiful ___ I've ever seen.
12. ___ I just ate a _____of food and left for school in a hurry.
13. ___ _____ rice soup, is a kind of thick, white soup.
14. ___ Don't go into the haunted house. There are bad _____s inside.
15. ___ Do not make friends with your _____. Be cautious!
16. ___ Grandmother's birthday party will _____ in Grand Hotel.

Appendix C

Vocabulary Pretest for Level 4 Subjects (“Addicted to Chocolate”)

1 獨特的	7 有害的	13 證據	19 研究
2 受試者	8 外層	14 刺激	20 危險
3 上癮的	9 喜悅	15 營養的	21 疾病
4 組合; 搭配	10 膽固醇	16 情況	22 蛀牙
5 意見	11 同等的	17 保護	23 引發; 促成
6 正面積極的	12 咖啡因	18 吃完; 喝完	24 化學成份

Select the best item and write down a number (1–24) on the blank below.

- | | | | |
|------------------|-------|--------------------|-------|
| 1. distinctive | _____ | 9. cavities | _____ |
| 2. stimulate | _____ | 10. research | _____ |
| 3. evidence | _____ | 11. caffeine | _____ |
| 4. nutritional | _____ | 12. positive | _____ |
| 5. coating | _____ | 13. consume | _____ |
| 6. contribute to | _____ | 14. participant(s) | _____ |
| 7. cholesterol | _____ | 15. damaging | _____ |
| 8. pleasure | _____ | 16. addicted to | _____ |

Immediate Vocabulary Recall Test for Level 4 Subjects

1 distinctive	7 contribute to	13 positive	19 consume
2 damaging	8 research	14 chemicals	20 risks
3 evidence	9 pleasure	15 participants	21 disease
4 nutritional	10 cavities	16 stimulate	22 combination
5 equivalent	11 condition	17 protect	23 addicted to
6 coating	12 caffeine	18 cholesterol	24 opinion

Select the best item and write down a number (1–24) on the left blank below.

1. ___ My brother is ___ playing on-line games. He is so crazy about them.
2. ___ The most common things containing ___ are: coffee, tea, and cola drinks.
3. ___ Dad wants to lose weight, so he goes to a doctor and ask for ___ advise.
4. ___ Men _____ more meat than women worldwide.
5. ___ John likes to eat sweets. He has many _____.
6. ___ Herbal tea may reduce ____, a kind of fat that is found in the bloodstream.
7. ___ Keeping a _____ attitude can make you energetic every day.
8. ___ We had great _____ in chatting with each other that night.
9. ___ It is said that Mozart's music can be used to ___ babies' brain development.
10. ___ A group of doctors had conducted ___ to find out the cause of the disease.
11. ___ Their disagreements _____ the failure of the marriage.
12. ___ Do you have any _____ that Mark took the money?
13. ___ The study needs 100 _____ to volunteer for the tests.
14. ___ Violent TV programs can be _____ to our children's mind.
15. ___ The restaurant is famous for its _____, home-made Italian food.
16. ___ We supply a range of ___ services including ceramic and stainless steel.

About the Authors

Ying-Hsueh Cheng is a doctoral student of Foreign, Second, and Multilingual Language Education at the Ohio State University. Her current research interests include reading comprehension, such as L1 and L2 glosses, and vocabulary acquisition; project-based learning; teachers' and students' expectations in a foreign language classroom from sociocultural perspectives. E-mail: sherrycheng85@gmail.com

Robert L. Good is an assistant professor in the Department of English at the National Kaohsiung First University of Science and Technology, Taiwan. His current research interests include vocabulary issues such as size, acquisition, and use; lexicography as an account of current language; the cultural vocabulary of cultural literacy; and international communication as informed by cross-linguistic sociolinguistics and cultural literacy belonging to national and international communities. E-mail: rgood@ccms.nkfust.edu.tw