GLOSSING IN INCIDENTAL AND INTENTIONAL LEARNING OF FOREIGN LANGUAGE VOCABULARY AND READING

MYONG HUI KO
University of Hawai‘i

INTRODUCTION

Numerous studies have been conducted on incidental vocabulary learning in second/foreign language learning (Pitts, White, & Krashen, 1989; Dupuy & Krashen, 1993; Day, Omura, & Hiramatsu, 1991; Omura, 1992; Ha, 1992; Knight, 1993). The results of these studies indicate that a second/foreign language learner can learn new words through reading. However, learning words from contexts while focusing on reading is considered inefficient due to the lack of contextual cues provided in the text (Hulstijn, 1992). There have not been enough studies on input modification to help students cope with insufficient contextual cues in learning new words through reading. So far, only a few studies (Holly & King, 1971; Jacobs & Dufon, 1990; Hulstijn, 1992; Watanabe, 1992) have tried input modification on incidental vocabulary learning through reading. Among input modifications, providing a gloss is one type of vocabulary learning aid and reading comprehension aid. Glossing in previous studies (e.g., Holly and King, 1971; Jacobs & Dufon, 1990; Hulstijn, 1992; Pak, 1986; Johnson, 1982; Davis, 1989; Jacobs, 1991) is used to provide intentional stimuli for either learning new words or enhancing reading comprehension. No studies have investigated how glossing affects different learning conditions. In general, there is insufficient research on the effect of glossing in a second/foreign language vocabulary learning and reading comprehension. Hence further investigation in this area is warranted.

This study attempts to address how glossing affects learning vocabulary and reading comprehension in a foreign language. It compares a gloss condition with a no gloss condition, as well as a L1 gloss versus L2 gloss. It also investigates how different learning conditions (intentional vs. incidental) affect vocabulary learning and reading comprehension in a foreign language. Moreover, it investigates retention of learned words over the long term. Here, the long term is described as any duration from at least 24 hours to one or more weeks after treatment (Nation, 1982).
The Role of Vocabulary in Second/Foreign Language Learning

The importance of vocabulary in second/foreign language learning has been relatively neglected in ESL/EFL research. Generally, vocabulary has taken on a secondary role in favor of grammar. Vocabulary is generally not taught as a skill in itself. Rather, it is taught as a tool to improve reading and listening comprehension in both second/foreign language reading and listening courses (Davis, 1989; Judd, 1978; Coady, 1993; Gass, 1988; Taylor, 1990).

Some researchers (e.g., Stein, 1993; Davis, 1989; Judd, 1978; Omura, 1992) argue that vocabulary is one of the major aspects in learning second/foreign language since, without access to a wide range of words, meaningful communication cannot take place. Gass (1988) states that inappropriate lexical use causes relatively crucial consequences in communication because a lexical item carries a speaker's or writer's intended message, whereas grammatical and phonological errors can be overcome from other information provided. In other words, limited word knowledge can handicap learners from expressing their thoughts clearly through productive skills such as speaking and writing. Nation (1982) adds that the lack of vocabulary can also bring about a failure to receive information through receptive skills such as listening and reading. Kruse (1979, p. 208) states that "of prime importance in reading is vocabulary skill. Readers must know the meanings of enough of the words in a sentence for it to make sense. They must also know how to combine individual word meanings within a sentence. Once students are passed the initial stages of reading, they spend a large percentage of their time encountering new vocabulary." Saville-Troike (1984, cited in Stein, 1993) also indicates that knowledge of vocabulary is the most crucial factor in achieving academic success for second language learners. Because second language students have to read a great amount of material in their everyday academic life, their knowledge of vocabulary plays a very important role in the reading comprehension needed to carry out their academic courses.

Though each has unique characteristics which separate one from the other, reading and vocabulary seem to provide mutual benefits in second language learning (Dubin, 1989). Laufer (1992) points out that both the knowledge of words and the content of the given topic are the most important variables in second language reading. And because learners tend to depend on the meaning of words first, and content and syntax second and third, we can see that "the nature of the language threshold for reading purpose is largely lexical" (Laufer & Sim, 1985a, 1985b, cited in Laufer, 1992). Background knowledge in the given topic is another important
GLOSSING IN INCIDENTAL AND INTENTIONAL LEARNING

But when readers are confronted with a reading passage which contains many low frequency vocabulary items, their content schema cannot help them because they have to give more attention to decoding the unknown words (Davis, 1989).

Among several ways of measuring vocabulary difficulty, frequency and word length are common categories. In other words, if sentences contain words that are more frequent in occurrence and shorter in length, then the text is considered more readable. In fact, vocabulary difficulty may be one of the strong predictors of comprehension difficulty (Nation & Coady, 1988).

Depending on how one learns a new word, two types of classification of vocabulary knowledge are possible. Vocabulary knowledge can be seen as a continuum "with the initial stage being recognition and the final stage being production" (Gass, 1988, p. 94). On one end of the continuum, we find potential vocabulary for which learners understand the meanings though they have not previously come across the words. At the other end of the continuum, we find real vocabulary which includes words that learners have already encountered and "can either only understand (passive real vocabulary) or both understand and use (active real vocabulary)" (Berman, Buchbinder, & Beznedeznych, 1968, cited in Palmberg, 1990, p. 1).

The latter distinction between active and passive is also referred to as productive versus receptive. Productive knowledge of vocabulary means the ability that a learner has to use the word through the productive channels of speaking and writing. Knowing a word in this sense involves several features, including "phonological, orthographic, morphological, syntactic and semantic" (Laufer, 1990, p. 148). So productive knowing of a word means that a learner is able to pronounce the word, to write it, to use in a grammatically correct way, to produce some derivations from the basic morphemes, and to use it appropriately in various contexts. Receptive knowledge of vocabulary refers to the ability of a learner to understand the word through the receptive channels of reading and listening. Compared to productive knowledge of vocabulary, receptive knowledge of vocabulary requires less features. It involves abilities to recognize and recall the word's meaning when it is seen or heard, to know its collocations, and to know grammatical patterns of the word (Crow, 1986; Nation, 1990; Laufer, 1990).

The requirements for productive knowledge of a word include the receptive knowledge of it. Obviously, much time and effort are required to control productive vocabulary. However, the amount of receptive knowledge is much larger than
productive knowledge because it includes low frequency words. Thus even native speakers take a long period of time to acquire receptive vocabulary (Nation, 1990). Almost 80 percent of words in an average two-page text are contained in a list of about 2,000 high frequency words (Coady, 1993). These, 2,000 high frequency words may be an adequate productive vocabulary for a second language learner. The 2,000 words have high probabilities of occurrence in a text, but in general, they are low in the specific information which helps to decode the content of the message. Thus it is the low frequency words which carry the content information. While reading an average unmodified text, readers may not be able to understand about 20 percent of the words, which are high in information content. Hence they may have difficulty in comprehending the reading passage. Accordingly, second language learners may need many more than 2,000 words of receptive vocabulary for reading comprehension (Fox, 1987; Richards, 1974).

**Learning Words through Context**

Considering the fact that native children learn as many as 3,000 words per year between grades three and twelve, mostly through other than direct instruction, we can see that indirect learning plays an important role in acquiring new words. This fact leads to the conclusion that second language learners need exposure to a great amount of reading material and many listening tasks (Nation, 1982; Hulstijn, 1993). Krashen (1989) states that massive reading alone can serve as sufficient comprehensible input for acquiring new words. Based on his input hypothesis, he claims that vocabulary can be learned as a by-product of reading.

There are several reasons for learning words through reading. First of all, high-frequency words are certainly justified in direct instruction due to their frequent occurrences in an average text, whereas low-frequency words are not. Moreover, because these high-frequency words play a crucial role in first or second language reading comprehension, they should be taught. However, low-frequency words are very poor candidates for instruction and learning in this sense. Thus these less frequent words may need to be learned through contact in context in extensive reading. However, learning words through extensive reading is effective only after reaching a critical level of word-recognition knowledge on core vocabulary or high-frequency words (Coady, 1993). The second reason for learning words through reading is the tremendous number of words to be encountered in school or in everyday reading materials; direct instruction alone will never be able to meet the
demand, considering the fact that "even extremely ambitious vocabulary programs do not cover more than a few hundred words per year" (Nation & Coady, 1988; Nagy & Herman, 1987, p. 23). Third, the level of word knowledge required to improve reading comprehension will not be fulfilled by receiving one-time instruction on word definitions because this approach does not provide very deep knowledge of the words. Thus multiple exposures in different contexts are necessary, and extensive reading can provide such multiple exposures (Nagy & Herman, 1987).

Among several techniques for teaching vocabulary is learning words from context. Using intelligent guesses, learners derive the meaning of a word from the grammatical and pragmatic contexts (Stein, 1993). This method is based on a positive assumption. If learners have to infer the meaning of unknown words from the context, they have to exert more mental effort than if they were given with the meaning initially. Thus they will retrieve or recall the meanings of words better than when they put less mental effort into learning their meanings.

Nagy, Herman, and Anderson's (1985) study reports that students can learn words from a natural context. A total of 57 eighth-grade students read either an expository or a narrative text. After reading, to assess their knowledge of 15 target words from each text, students were asked to do two tasks: an individual interview and a multiple-choice test. Those who read either type of passage showed small but statistically significant gains in target words. However, subjects in this study were native speakers of English.

Second/foreign language studies have demonstrated that, in the case of second/foreign language learners, learning words from context can be risky and not effective at all. This is because the context can supply limited information about the unfamiliar word and is sometimes misleading. In addition, because of their lack of ability in their target language, second language learners do not have good intuitions that can help them make sophisticated guesses. Hulstijn (1992) demonstrated that second language learners tended to guess the meanings of words incorrectly from the context when no cue was given. Watanabe (1992) also reported that providing cues in a text led to significantly better performance than the original text, where subjects were given a situation to derive word meanings from the context.

Kruse (1979) indicated that L2 learners should not be pushed to make intelligent guesses like native speakers because "they do not have the ability to abstract semantically in the language that a native speaker does, and so, especially in early stages, must be aided in using all the concrete clues available" (p. 210). Researchers
such as Bensoussan and Laufer (1984) and Hulstijn (1992) have also argued against context-based vocabulary learning because of problematic issues in this area. Hulstijn (1992, p. 114) sums up these problems in this way:

[First], context seldom offers enough information for the inferring method to be successfully applied; [second], the inferring method has the inherent risk that learners will make wrong inferences, and hence learn the wrong meaning (which they have to 'unlearn' subsequently); [third], the inferring method only works well with learners who have good problem-solving skills.

Therefore, there is good reason for skepticism about learning words from context. Stein (1993) states that although many second language teachers present reading materials which have clear linguistic clues when they teach a context-based inferring method to their students as a way of learning vocabulary, generally everyday authentic reading materials do not include such clear-cut clues. Therefore, second/foreign language teachers have a responsibility to train their students to cope with this deficiency. Encouraging students to use a dictionary to confirm their guesses, and selecting and providing reading materials which have glosses, may be some good alternatives to consider.

Two Ways of Learning Vocabulary: Intentional versus Incidental

Vocabulary can be learned in two ways: intentional learning and incidental learning. Intentional and incidental learning are widely used terms in psychology and applied linguistics. These terms are operationalized depending on the conditions of task instructions. In psychology, intentional learning is used when the objective of a task is explained explicitly, that is, a learner needs to know or is responsible for a certain information while performing a given task. In intentional learning, the learner has a clear idea about the intentional stimuli before getting on with the task. In contrast, incidental learning, which is also referred to as unintentional learning, is used when the objective of a task is not overtly taught but implied within the task. Thus the learner is not responsible for knowing the information which will be evaluated later; but the learner may notice through cues provided in the task because these cues are often associated with the intentional stimuli. For example, students were given tests on both content and slide recognition after showing unannounced slides during a lecture. Students knew that they were responsible for the content of the lecture (intentional learning), but they were not responsible for knowing the slide, although it was relevant to what they were learning during class (incidental learning).
GLOSSING IN INCIDENTAL AND INTENTIONAL LEARNING


The terms, intentional and incidental learning are also common in the second/foreign language learning field. However, their definitions are not consistently stated. In fact, there is some discrepancy among researchers, especially in describing the intentional learning condition.

Hulstijn (1992) describes two learning conditions depending on whether students were told in advance what they are responsible for. Those who were told in advance that they would have a test on target words were treated as the intentional vocabulary learning group, whereas those who were not told were treated as the incidental vocabulary learning group.

Konopak, Sheard, Longman, Slaton, and Atkinson (1987) refer to an intentional vocabulary learning group in which students read a text with target words underlined, and write the meaning of underlined words, guessing from the context. An incidental learning group was described as occurring when students read a text in which the target words were not underlined and the students were not told that they would have a vocabulary test later. Here, intentional and incidental learning depended on three factors: first, whether the attention on the target words was incorporated within the task or not; second, whether the students performed the relevant work which they would be tested on later; and third, whether the students were told in advance what they would be responsible for.

Nation (1982) uses direct and indirect vocabulary learning instead of intentional and incidental learning. He describes direct vocabulary learning to be when "a conscious effort is made to learn vocabulary either in context or in isolation, for example, by learning lists of word forms and their meanings, by doing vocabulary learning exercises, or by studying affixes and roots." Indirect learning is described as occurring where "new words are learned incidentally while reading or listening, usually as the result of information provided by the context" (p. 15). According to him, the decontextualized nature of vocabulary study is considered intentional learning, whereas incidental learning is similar to the definitions used in psychology and other second/foreign language learning fields.

Gee (1988) uses acquisition interchangeably with incidental learning, and overt learning with intentional learning. He discusses two modes of learning: "acquisition, or incidental learning, is development by exposure to models, without conscious learning and without overt explanation through direct teaching. It is the process by
which everyone acquires their first language. The other process is overt learning, or intentional learning, in which the teacher often explains overtly and the learner applies various intentional strategies (memorization, searching for connections, asking for rules, rehearsal, study)" (p. 217).

Schmidt (1990) states that consciousness, drawing learners attention to the formal properties of language, is a necessary condition in second language learning. He has classified different types of consciousness. With respect to consciousness as intention, he explains that it is "conscious efforts, attempts, and strategies, referring to the volitional, deliberate nature of the action" (p. 133). When a learner does something consciously, it means he or she is making a deliberate plan. Incidental learning is the condition where an individual learns "without the intent to learn or the learning of one thing (e.g., grammar) when the learner's primary objective is to do something else (e.g., communicate)" (Schmidt, 1994, p. 9). In incidental learning, attention is not focused. Schmidt believes that incidental learning may require less attention whereas the other requires more attention because attention is a necessary condition in order to learn something; he rejects the possibility of learning without attention (Schmidt, 1990; 1993; 1994).

According to Krashen (1989), comprehensible input can provide all that is necessary for vocabulary acquisition. He claims that even though some studies on intentional learning indicate that it is superior to incidental learning, because we do not know how much time and effort were put into skill-building exercises, we cannot tell which is more effective. Thus he argues that extensive reading is as effective as focus-driven learning and provides more interesting tasks to perform. However, it should be noted that controlled studies have consistently demonstrated the superiority of intentional learning over incidental learning for vocabulary.

A study done by Konopak et al. (1987) demonstrates that intentional vocabulary learning results in better performance than incidental vocabulary learning. An intentional learning group was asked to read the first form of a passage in which 10 target words were underlined and to provide the meaning of each underlined target word based on contextual clues. An incidental learning group read the second form of the passage in which the target words were neither underlined nor emphasized. A control group read a newspaper article which did not contain the target words. Then, all subjects took a posttest, writing down word definitions along with providing information about their knowledge of each word. The results indicated that there was a significant difference for both treatment conditions (intentional and incidental
groups) from the control group. Although the intentional vocabulary learning group performed much better than the incidental vocabulary group, Konopak et al. did not test for any significant effect between these two groups since they were interested in two issues: the ability of students to learn new words through contextual information and the effect of the number of exposures to target words on students’ vocabulary learning.

Hulstijn (1993) points out the weakness of incidental learning in second/foreign language vocabulary. Although readers can pick up some new vocabulary incidentally while reading, their intention is to get the meaning of the text. Hence in most cases, they forget unknown words once their meanings are established. The learning of words depends on how much mental effort the reader devotes to a particular word. In order to acquire productive vocabulary, a reader has to pay attention to the link between a word’s meaning and its form, but he/she can hardly do this while reading for meaning. Consequently, it is difficult to gain productive vocabulary knowledge while focusing on a reading task.

According to psychological models of memory, information is encoded in the short-term memory with the help of attention, and a certain amount of rehearsal is necessary in order to encode a stimulus from the short-term memory into the long-term memory. Considering the fact that humans have limited capacity for incoming information, the degree of efficient retrieval from memory later depends on both the quantity and quality of attention given at the time of encoding the stimulus. Processing in short-term memory is the first step and a necessary condition for permanent storage so failure to process into short-term memory means the loss of the stimulus. Information stored in short-term memory should be encoded into long-term memory, otherwise the stimulus will be lost again (Schmidt, 1990; 1993).

Incidental learning, which retains information as a by-product of another task, can be held in short-term memory for a while but cannot be processed to the long-term memory due to the lack of attention and rehearsal. However, intentional learning, in which the intention was generated by a predetermined goal, has the power to decide which information should be processed so it motivates the learner’s attention to the relevant information (Klauer, 1984). Therefore, because of the nature of human cognition, incidental vocabulary learning, though possible and commonplace, is not predicted to be superior to intentional vocabulary learning (McLaughlin, 1990; Schmidt, 1993).
Studies on Incidental Vocabulary Learning in Second/Foreign Language

Learning through context. There have been some studies on reading and incidental vocabulary learning in both ESL and EFL contexts. Pitts, White, and Krashen's (1989) study, which was a partial replication of Saragi, Nation, and Meister (1978), and Ferris' study (cited in Pitts, White & Krashen, 1989) showed that reading could be a good source for acquiring second language vocabulary for adult ESL learners. Dupuy and Krashen (1993), Day, Omura, and Hiramatsu (1991), and Omura (1992), demonstrated that incidental vocabulary learning was also possible in EFL situations. The primary implication of these studies is that EFL programs should encourage extensive reading to increase the foreign language vocabulary knowledge of EFL learners.

Knight (1993), Ha (1992), and Luppescu and Day (1993) were interested in whether access to a dictionary while reading for meaning can facilitate incidental vocabulary learning for foreign language learners. Knight (1993) divided his subjects into four groups according to the subjects' verbal ability (low or high) and reading conditions (dictionary access or no dictionary access). The results showed that those with dictionary access and those of higher verbal ability performed significantly better on vocabulary measures (incidental learning). These subjects also had better recall scores for the reading passage (intentional learning). Ha (1992) had three different groups: a bilingual dictionary group, a no-dictionary group, and a no-reading group. Luppescu and Day (1993) tested two groups, a bilingual dictionary group and a no-dictionary group. The results of these two studies showed that those who used a bilingual dictionary performed significantly better on the vocabulary test than those who did not use a dictionary.

Learning through input modification: Glossing. According to a bottom-up view of second language learning, glossing can aid learners' reading comprehension of difficult texts because the basic construct of a sentence and later a paragraph is the individual word. However, a top-down process of language learning would argue that glossing can disrupt the flow of the reading process due to the attention paid to the words while reading for meaning. At any rate, the use of vocabulary glosses in second language reading materials is a common practice (Jacobs & Dufon, 1990; Holley & King, 1971).

Marginal glossing is one way to help a learner comprehend the reading materials. By writing additional notes or information beyond the text in the margin of the same page or another page, marginal glosses guide the learner's attention and assist as a
GLOSSING IN INCIDENTAL AND INTENTIONAL LEARNING

mediator between the text and the learner. Marginal glosses have various functions in helping to decode the text by providing additional knowledge in specific content, skills, strategies, and definitions of difficult words. In general, the gloss in second/foreign language learning refers to one type among the many listed above, that is, providing information on the important words such as clues, definitions, or synonyms of vocabulary (Stewart & Cross, 1991; 1993; Richgels & Mateja, 1984). There are several reasons to use glosses in aiding vocabulary learning. First, they help readers understand new words more accurately, considering the fact that deriving meaning from context is difficult and risky in some aspects (e.g., Stein, 1993; Hulstijn, 1992). Second, frequent input, looking at the words in the glosses and in the context, can help to retain the meaning in the memory longer (Watanabe, 1992). Third, students prefer to have glosses in their second/foreign language reading materials (e.g., Jacobs & Dufon, 1990).

Only a few studies have been conducted on the effect of glossing on second/foreign language vocabulary learning. Holly and King (1971) compared different types of glosses placed at different positions in the text: side-of-page, bottom-of-page, or a gloss on an attached sheet. The results indicated no significant difference on either vocabulary or reading comprehension tests.

Jacobs and Dufon (1990) tested three conditions: no glosses, English (L1) glosses, and Spanish (L2) glosses. Both L1 and L2 gloss groups significantly outperformed those in the no-gloss group on the immediate vocabulary test. However, no significant difference was found on the delayed vocabulary test, which was given for weeks later. The researchers found no significant difference between L1 and L2 glosses.

Hulstijn (1992) introduced multiple-choice (MC) type glosses. He proposed that, since MC glosses give readers a cue, this can compensate for the limitations of context-based inferring. Because it requires more mental effort in selecting the right meaning among many, the MC analogy may produce better recall of the meanings of words. His findings showed that: (a) in a within-subject design, second language learners were more likely to pick up words when they inferred meaning from the text (high mental effort) than when the meaning of words was given to them (low mental effort), (b) second language learners tended to guess the meanings of words incorrectly when no cues were given, (c) the MC condition had a higher retention of vocabulary than synonym condition, but second language learners made substantial errors due to the appealing distracters in MC glosses.
Watanabe (1992) was interested in the comparison of three cue types: appositive, single glosses, and MC glosses. The results showed that those who read a passage with either single gloss format or MC gloss format performed significantly better than those who read the unmodified original passage in both immediate and delayed vocabulary posttests.

Studies on Glossing in Second/Foreign Language Reading Comprehension

Nation (1990) discusses some of the advantages of glossing to help second/foreign language reading comprehension. First, glossing can minimize interruption while reading is going on. Since glossing provides the definitions of low frequency words, second language readers do not have to constantly look up the words. Second, students can be more independent from their teachers since they can read on their own with the help of the glosses. Third, glossing allows for individualization. Considering the fact that not all students have problems with the same words, they can look up just the words which they do not know.

There have also been a few studies done on the effect of glossing on enhancing second/foreign language reading comprehension. Along with her primary interest, the role of cultural schema, Johnson (1982) tested the effect of vocabulary in four different conditions: no help in vocabulary, studying the definitions of words before reading, reading a passage with the target words glossed, and studying the target words before reading and reading the text with gloss. The results indicated that the cultural schema facilitated reading comprehension, but exposure to different types of vocabulary aids did not significantly affect comprehension. Pak (1986) categorized her subjects into either above-average or below-average groups based on their reading ability, and also tested three different conditions of vocabulary aids: no vocabulary help, with marginal gloss, and with the words defined parenthetically within the passage. The subjects in the above-average group in reading ability performed significantly better than those in low-average group. However, the different presentations of vocabulary aids did not appear to affect their reading comprehension. Jacobs and Dufon (1990) investigated the effects of L1, L2, and no gloss on foreign language reading comprehension, along with foreign language vocabulary learning. Again, there was no significant difference among the different treatments.

Davis (1989) and Jacobs (1991) both found a positive effect for glossing and foreign language reading comprehension. Davis (1989) compared three different
GLOSSING IN INCIDENTAL AND INTENTIONAL LEARNING

conditions. The first condition was the read-write-reread group, who read the passage for fifteen minutes and wrote what they could remember for ten minutes. The second condition was the vocabulary/guide before reading group, who studied the guide for ten minutes and read for fifteen minutes. The third condition was the vocabulary/guide during reading (gloss) group, who read the text for twenty-five minutes. Subjects who received vocabulary help either before or during reading did significantly better than those who received no help. Jacobs (1991) compared unglossed and English (L1) gloss conditions and found the glossed test group performed significantly better than the other group.

Purpose of the Study

Many researchers (e.g., Judd, 1978; Stein, 1993; Davis, 1989; Omura, 1992) have agreed that a sufficient amount of vocabulary is required to achieve a level of communicative competence in a second/foreign language. Even though some work has been done in incidental vocabulary learning in a second/foreign language, only a few studies (e.g., Holly and King, 1971; Jacobs & Dufon, 1990; Hulstijn, 1992; Watanabe, 1992) have been conducted on the effects of glossing on incidental vocabulary learning through reading in an ESL/EFL context. Among them, only the Jacobs and Dufon (1990) study compared the effect of L1 gloss and L2 gloss, and it did not find any significant difference between them. Further research is needed to better understand the role of different types of glosses in second/foreign language learning. The effect of glossing has also been tested in facilitating second/foreign language reading comprehension in some studies (e.g., Johnson, 1982; Pak, 1986; Jacobs and Dufon, 1990; Davis, 1989; Jacobs, 1991). However, the contradictory findings call for further research to arrive at more conclusive answers. So far, no studies have compared the effect of glossing on two different types of learning conditions: intentional learning versus incidental learning. In this study, incidental learning and intentional learning are compared in both foreign language vocabulary learning and reading comprehension. Subjects reading the same text with different three gloss conditions (no gloss, L1 gloss, and L2 gloss) are also compared in terms of foreign language vocabulary learning and reading comprehension.

Knowing a word involves many factors, such as phonology, orthography, morphology, syntax, and semantics (Laufer, 1990). Considering the fact that the subjects in my study are not taught the features of target words explicitly, as listed above, but only learn new words through reading, they fall into the category of
learning words receptively, which in general involves "being able to recall the translation of the foreign word when the foreign word has been seen or heard" (Nation, 1982, p. 19). Thus knowing a word in my study means being able to recall its definition when seen.

The definitions of incidental and intentional learning in second/foreign language learning vary due to inconsistencies in researchers' definitions (e.g., Hulstijn, 1992; Konopak et al., 1987; Nation, 1982; Gee, 1988; Schmidt, 1990). The learning conditions I adopted are similar to Schmidt (1990; 1994), in which incidental learning is a process of learning something unintentionally while focusing on something else, and intentional learning is a process where students have a deliberate plan guiding their learning, where the learning conditions are determined by predetermined objectives explained overtly before the task. Hulstijn (1992) provides an example. According to him, the different types of learning task were operationalized depending on whether the subjects were given any advance explanation about the objectives of the task. Those who were told in advance that they would have a test on the target words were treated as an intentional vocabulary learning group, whereas those who were not told in advance were treated as an incidental vocabulary learning group.

To address these issues, the following research questions were posed:

1. Under which condition (incidental vs. intentional), are EFL learners better able to learn new vocabulary?

2. Is there a significant difference for EFL readers in comprehending the reading passages under the two different learning conditions (incidental vs. intentional)?

3. What are the differential effects of text with no gloss, L1 gloss, and L2 gloss in facilitating L2 vocabulary learning for EFL learners?

4. What are the differential effects of text with no gloss, L1 gloss, and L2 gloss in facilitating L2 reading comprehension for EFL readers?

5. Are students able to retain the learned words after one week?

6. What are the effects of three different types of texts on a delayed vocabulary test given a week later? Are the results the same as on an immediate vocabulary test?

The following five hypotheses will be investigated in order to answer the above research questions:

1. Students in the intentional vocabulary learning condition will perform significantly better on the immediate vocabulary posttest than students in the incidental vocabulary learning condition.
2. Students in the incidental vocabulary learning condition will perform significantly better on the reading comprehension test than students in the intentional vocabulary learning condition.

3a. Students in two gloss conditions will perform significantly better on the immediate vocabulary posttest than students in no gloss condition.

3b. However, there will be no significant difference between L1 gloss and L2 gloss conditions.

4. There will be no significant difference among the gloss conditions on the reading comprehension posttest.

5. Students will be able to retain the learned words after one week.

6a. Students in two gloss conditions (L1 and L2 gloss) will perform significantly better on the delayed vocabulary posttest than students in the no gloss condition.

6b. However, there will be no significant difference between subjects' scores in L1 gloss and L2 gloss conditions.

6c. The overall mean scores of the delayed vocabulary posttest will be lower than the mean scores of the immediate vocabulary posttest.

Hypothesis 1 is based on Konopak et al. (1987), who compared three conditions: intentional, incidental, and control groups. Their results showed that the intentional group performed best, followed by incidental and control groups, respectively. Although Konopak et al. indicated there was a significant difference between treatment groups (intentional and incidental) and the control group, they did not report whether there was any significant difference between the two treatment groups since they were interested in investigating whether students could learn new words through contextual information without overt instruction and whether the number of exposures to the target words could affect on their vocabulary learning.

Hypothesis 2 is derived from VanPatten's (1990) study which indicates that paying simultaneous attention to both meaning and form is difficult for learners since humans have a limited capacity to comprehend incoming information at any given moment. In his study, after a listening task, students who paid attention to content only produced the highest recalls than those who paid attention to both content and other forms such as key lexical items, definite articles, and verb morphemes. Hence it is predicted in my study that if the students are told that they will have a test on the content of the story, they will focus more attention on the meaning of the passage compared to students who will be told that they will have a vocabulary test. In other words, since students in the intentional vocabulary learning group will read while
thinking about the vocabulary test, their attention may be divided between content and vocabulary. This divided attention may cause the subjects to perform less well on the reading comprehension test than those in the incidental vocabulary learning group, whose attention will be focused solely on the content of the story.

Hypotheses 3a and 3b are derived from Watanabe (1992) and Jacobs and Dufon (1990). Watanabe showed that those who read a passage with either synonym or MC gloss performed significantly better than those who read the unmodified original passage. But there was no significant difference between synonym and MC gloss. Jacobs and Dufon (1990) also found that gloss conditions (L1 and L2 glosses) performed significantly better than the no gloss condition on immediate vocabulary posttests. They showed no statistical significance between L1 and L2 gloss.

Hypothesis 4 is based on Jacobs and Dufon’s (1990) study. However, in the literature, the results are contradictory. Johnson (1982) and Pak (1986) indicated no significant positive effect for glossed conditions in foreign/second language reading comprehension, whereas Davis (1989) and Jacobs (1991) indicated significant positive effects for the glossed condition. I derived my hypothesis from Jacobs and Dufon because their study was similar to mine with respect to comparing L1, L2, and no gloss conditions in foreign language reading comprehension. However, the method of accessing students' vocabulary knowledge was different, as they measured it through a production test, whereas I measured it with a multiple-choice test.

Hypothesis 5 is derived from Watanabe (1992) and Ha (1992). Like Watanabe and Ha, Jacobs and Dufon (1990) carried out delayed vocabulary posttest, but they showed different results from Watanabe and Ha. I assume that the reason for this difference is the variation in the time interval between the immediate and delayed vocabulary posttests. Jacobs and Dufon gave a delayed vocabulary posttest after four weeks, whereas Watanabe and Ha gave it after one week. Because the delayed vocabulary posttest in my study will be given after one week, a result similar to Watanabe and Ha may be expected, which will demonstrate that students can retain learned new vocabulary after one week.

Hypothesis 6a is derived from Watanabe (1992) which demonstrated that the glossed group outperformed the original text group on the delayed vocabulary posttest given after a week. Hypothesis 6b is derived from Jacobs and Dufon (1990) which showed no significant difference between L1 and L2 glosses. Jacobs and Dufon gave a delayed vocabulary posttest four weeks later, whereas I will give it after one week. Because of this interval difference between immediate and delayed vocabulary
posttests, there is some doubt about predicting the outcomes based on their study. However, since they did not find any significant difference between L1 and L2 glosses on the immediate vocabulary posttest, it is unlikely that there will be a significant difference between L1 and L2 glosses on the delayed vocabulary posttest in this study.

Hypothesis 6c is derived from Watanabe (1992) and Ha (1992). Watanabe had two types of delayed vocabulary posttests, writing down the meaning of words without context and with context. He reported lower mean scores on the delayed vocabulary posttest without context except for the control and original passage, whereas there were higher mean scores on the delayed vocabulary posttest with context. Since my study will be testing students' vocabulary knowledge without providing context, I may attain a result similar to his vocabulary test without context. Ha (1992) also showed that the mean scores of delayed vocabulary posttest were lower than the mean scores of immediate vocabulary posttest.

METHOD

Design

This study adopted a between groups design for pretests, main tests, and the delayed test. There are two dependent variables (vocabulary and reading comprehension) and two independent variables: learning (with two levels) and glossing (with three levels). In total, there are six conditions, as shown in Figure 1.

\[
\begin{array}{c|c|c}
\text{Gloss} & \text{Incidental} & \text{Intentional} \\
\hline
\text{No gloss} & & \\
\text{L1 gloss} & & \\
\text{L2 gloss} & & \\
\end{array}
\]

*Figure 1. The 2 x 3 factorial design.*

A total of 189 university students were assigned to one of the six conditions. Students read a text which included 15 target vocabulary items. The study was designed to test different learning conditions and input modifications (providing gloss) on text comprehension and learning new words. With respect to the learning
conditions, I used Hulstijn’s (1992) definition for types of learning. The type of task was determined to be either incidental or intentional based on the learners’ responsibility for the task, which was specified in advance. Consequently, students who were told in advance that they would have a reading comprehension test after reading a passage were treated as incidental vocabulary learning subjects. Those who were told in advance that they would have a vocabulary test were treated as intentional vocabulary learning subjects. The different learning conditions were included in order to test whether a predetermined objective could influence students’ performance in comprehending the text and learning unknown words. Different types of gloss were included to investigate whether different gloss conditions facilitate vocabulary learning and reading comprehension. This study was also designed to compare gloss or no gloss conditions, as well as L1 and L2 glosses. Regarding vocabulary learning, two versions of the vocabulary test (L1 and L2 versions) were given in a counterbalanced manner in order to be fair to the groups who had different versions of glosses. This study also investigated the retention of learned words after a week by giving a delayed test.

**Subjects**

The participants in my study were 189 college undergraduates enrolled in different types of English courses at Kyongsang National University in Chinju, Korea. I used regular English class periods of three intact classes to conduct my study. The classes met every other day for 50 minutes. Randomization of the students from the three classes was not allowed. Thus I had to assign different learning groups using the intact classes. One class, which had approximately 110 students, was chosen as the incidental vocabulary learning group, and two classes, with a total of approximately 60 students each, were chosen as the intentional vocabulary learning group. I started with more than a minimum number of students because I suspected that I might lose some students while going through three days of experiments involving five- and seven-day intervals. As anticipated, quite a few students missed one day of the three-day experiment. I could not include students who participated in only the pretests and immediate tests, in only the pretests and delayed tests, or in just one day of the experiment, that is, pretests, immediate tests, or delayed test. I decided to include those who participated in the main body of the study, the second and third days of the study. Because of this, I had fewer students in my pretest than in the main body of study. As noted above, because of attrition, the final number of
students was 158 for pretests, and the 189 for the immediate and delayed tests. Students who participated in the pretest also participated in the rest of the study. In other words, 158 students participated in the entire three-day experiment. The different number of students between pretests and main tests (immediate and delayed tests) included those who missed the pretests but participated in both the immediate and delayed tests.

The classes I used were general English courses, which students usually take either after finishing their freshman English courses or concurrently to enhance their English skills. Students are required to take freshman English courses for two semesters. However, many students take additional English courses because they know that English will be a very important subject throughout their lives. English, along with mathematics and Korean language, is one of the most important subjects to pass in the college entrance examination in Korea. Later, English is important when applying for jobs after college. Furthermore, once the students get jobs, they have to keep up their English skills because they are tested in English on their promotion examinations later. Many students also study English in preparation for TOEFL (Test of English as a Foreign Language), and TOEIC (Test of English for International Communication), and study test preparation books designed to prepare for job related examinations in addition to their English course work.

Most students from the general English courses had a variety of majors and class standings. However, they were a homogeneous group in terms of their age and English educational background. In general, their ages ranged from 19 to 26. Typically, they received grammar-based and sentence-by-sentence level translation of English instruction for six years, starting in junior high school, before entering university. Because the medium of instruction for their English classes was their native language, Korean, the students had almost no chance to speak English in the class. Therefore, it is assumed that the students’ grammatical or linguistic competence is advanced, but their speaking and listening skills are quite low.

Materials

I used a reading text adopted by Ha (1992) from Improving College Reading (Jacobus, 1972) to investigate the role of bilingual dictionaries in incidental vocabulary learning through reading. The title of the reading passage was "Where Ghosts Are Used as Servants." According to Ha (1992), the content of the text was selected because it was interesting. The length of the passage was 854 words.
Measured on the Flesch-Kincaid readability scale, it was 8.8th grade level. In other words, the appropriate level for the text was for native speakers who were in the eighth month of the eighth grade. Ha conducted a pilot study with a similar group of college students in Korea to determine whether the text would be appropriate for Korean subjects and to choose target vocabulary for them. Because my subjects were Korean college students, the level of the text should be appropriate. As shown in Table 1, I used the fifteen target words chosen by Ha based on the frequency of marks which her subjects made when they encountered unknown words when reading the text.

Table 1
Frequency of Marks for Target Words (N=42) (from Ha, 1992, p. 28)

<table>
<thead>
<tr>
<th>Word</th>
<th>Number of Marks</th>
<th>Word</th>
<th>Number of Marks</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. squirming</td>
<td>42</td>
<td>9. peep</td>
<td>30</td>
</tr>
<tr>
<td>2. jerking</td>
<td>40</td>
<td>10. bang</td>
<td>28</td>
</tr>
<tr>
<td>3. scoff</td>
<td>37</td>
<td>11. thumping</td>
<td>27</td>
</tr>
<tr>
<td>4. mantle</td>
<td>35</td>
<td>12. pierce</td>
<td>27</td>
</tr>
<tr>
<td>5. bead</td>
<td>32</td>
<td>13. corpse</td>
<td>23</td>
</tr>
<tr>
<td>6. agony</td>
<td>31</td>
<td>14. harrowing</td>
<td>21</td>
</tr>
<tr>
<td>7. manifest</td>
<td>30</td>
<td>15. parallel</td>
<td>21</td>
</tr>
<tr>
<td>8. creep</td>
<td>30</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

I adapted the reading text into three different forms: a text with no gloss, a text with L1 gloss, and a text with L2 gloss. The text with no gloss was the original passage which did not provide any information on the target words (see Ko, 1995, Appendix 1). The text with L1 gloss had clues or definitions for the fifteen target words in Korean at the bottom of each page (see Ko, 1995, Appendix 2). The text with L2 gloss had English clues or definitions (see Ko, 1995, Appendix 3).

There were two types of pretests: a cloze test and vocabulary pretest. I adopted a cloze test (which requires students to fill in the blanks in a prose passage) to check
my students' general English proficiency because cloze is considered a good general measure of a learner's second/foreign language proficiency (Oller, 1972). I used the cloze test which Watanabe (1992) used in his study because my subjects had similar background in terms of their English education and age (see Ko, 1995, Appendix 4). The cloze passage had a 9.0th grade level when measured on the Flesch-Kincaid readability scale. Watanabe adapted the passage from a book, Developing reading skills: Intermediate. The passage contained 277 words in total, and every seventh word deletion along with some rational deletions was made, leaving 25 blanks. Subjects were supposed to choose a correct word for each black from a list of 30 choices at the bottom of the cloze test. Watanabe explained that the reason he provided more words than necessary was to minimize the students getting correct answers by eliminating choices.

The vocabulary pretest was made by listing 15 target words, for which the students were expected to produce meanings. The purpose of giving this test was to learn whether students have previous knowledge of these words (see Ko, 1995, Appendix 5). I asked them to write out the meanings of each target word on paper in their native language, instead of making them select a correct answer from a multiple-choice test, to minimize correct selection by chance. I also avoided giving a multiple-choice vocabulary pretest to minimize the practice effect when they were given a multiple-choice vocabulary posttest later.

Two types of immediate posttests were given to the students after the reading. The reading comprehension test consisted of 20 multiple-choice items. The students were expected to choose a correct answer from among three choices. One distracter was more plausible and the other distracter was somewhat less plausible. I consulted with two native speakers in judging the plausibility of the distracters. The purpose of giving this test was to check students' overall understanding of the story. I did not give any type of production test because the reading time was limited to 18 minutes and the story was quite long; students might have had difficulty if they were expected to write down a summary of the story. I initially made 25 items, but eliminated five items after a pilot study because they were either too easy or too difficult. I created questions from all parts of the passage so that I could test their overall understanding of the story (see Ko, 1995, Appendix 6).

I used a multiple-choice vocabulary test that Ha (1992) had developed in her study to assess my subjects' immediate vocabulary knowledge after reading the text (see Ko, 1995, Appendix 7). Multiple-choice format was used because the target
words were learned through a receptive learning procedure, reading. Thus it was reasonable for the students to be able to recall the meaning of words when seen. Students were expected to choose one correct answer from among three distracters, of which one distracter was very plausible. Since I had two different types of gloss conditions (L1 and L2), I made an English version of the vocabulary test in order to prevent favoring a particular language gloss (see Ko, 1995, Appendix 8). Thus students took both versions one after the other. The versions had different orders in the items and distracters in order to minimize the practice effect. Combining the Korean and English versions of vocabulary tests, the total number of items was 30.

The delayed vocabulary test was the same as the immediate vocabulary test, except for rearranging the order of items and distracters (see Ko, 1995, Appendices 9 & 10). This was administered to compare the retention of vocabulary among different groups after a week.

**Procedures**

This study was administered to college students in Korea during regular English class periods from the end of May to the middle of June in 1994. Because many sets of materials were distributed and collected within a 50 minute time limit, two research helpers assisted in my study. I met my subjects three times to complete the study, when giving pretests, providing treatment and giving immediate posttests, and giving a delayed vocabulary test. The students were not told the purpose and the procedures of my study until they finished the whole process in order to insure unbiased data. Because I learned from the pilot test that most students were ashamed of their vocabulary pretest scores, I explained that the tests were for diagnostic purposes and that I was not interested in the individual scores so they should not be anxious about their individual scores on the test. However, there were still a number of students who did not put their names on the test paper because they were concerned about having a low score. The time limit for the each test and the reading passage was set after examining a few students whose backgrounds were similar to my subjects.

Since I had to use intact classes, I gave the pretests five days before the main part of the study. I chose five days because summer vacation was close, and it was the best timing considering the remaining time necessary to conduct the study. The cloze test took 25 minutes to test subjects’ general English skills. The vocabulary pretest was administered with a 10 minute time limit during which the students were expected to write the meaning of given target words. I had them produce the meaning of each
word in Korean to see whether they knew the word clearly. Because most of the students did not know the target words, they finished early, and I collected their test papers after four minutes. The whole procedure took about 30 minutes.

Five days later, the subjects were randomly divided into three different conditions: no gloss, L1 gloss, and L2 gloss groups. Subjects in incidental vocabulary learning groups were told that they would have a multiple-choice reading comprehension test, but they were not told that they would have a vocabulary test. After reading the passage for 18 minutes, they were given a 15 minute reading comprehension test, consisting of 20 multiple-choice items. Then two versions of immediate vocabulary posttest, consisting of 15 multiple-choice items each, were given for 18 minutes. Half of the students in each group took the Korean version first for nine minutes, and the English version second for nine minutes, while the other half of each group took the tests the other way around for counterbalancing purposes.

Subjects in the intentional vocabulary learning group were told that they would have a multiple-choice vocabulary test before they started reading the story, but they were not told that they would have a reading comprehension test. The same reading comprehension test and vocabulary test were given after the treatment. Subjects were not allowed to use any type of dictionaries while going through this procedure. The reading passages were collected before they took the tests so the subjects could not refer to the texts while answering the questions. The whole procedure took about 50 minutes.

A week later, the students in both the incidental and intentional groups were given an unexpected delayed vocabulary posttest so that a comparison could be made of the students’ retention in the different learning and gloss conditions. They were given 18 minutes to complete it. The students who took the Korean version first in the immediate vocabulary posttest took the English version first and Korean version later, while the others did it the other way around for counterbalancing purposes. I gave this delayed test a week later because the majority of incidental vocabulary learning studies had given their delayed test a week later (e.g., Watanabe, 1992; Ha, 1992), although Jacobs and Dufon (1990) gave it four weeks later. Since my study also investigated receptive vocabulary learning, a delayed vocabulary test after a week seemed reasonable. The whole procedure took only 18 minutes.

The cloze, vocabulary, and reading comprehension tests were scored such that a correct answer for each item received one point and an incorrect answer got zero
points. In the case of the vocabulary pretest, in which students were expected to write out the meanings of the target words, they also received points if they put down other meanings of the target words instead of the expected meaning because some target words had multiple meanings. The score on each test was the total number correct.

**Analyses**

The analyses were conducted using the SPSS PC+ Version 4.0 software package (Norusis, 1990). I analyzed the pretests (cloze and vocabulary pretests), main tests (reading comprehension and immediate vocabulary tests), and the delayed vocabulary test, respectively. I calculated descriptive statistics (the number of subjects, the mean, and the standard deviation) for each test. I also estimated the internal consistency reliability for these dependent variable measures using Cronbach's alpha.

In analyzing the pretests, since there were two dependent variables (cloze and vocabulary pretests) and one independent variable (six groups), a one-way multivariate analysis of variance (MANOVA) was performed to investigate the equivalence of six groups regarding their general English proficiency and prior knowledge of target words. The main tests were analyzed using two-way MANOVA since there were two dependent variables (vocabulary and reading comprehension tests) and two independent variables (learning conditions with two levels and glossing with three levels). After that, in order to investigate the retention of target words among the various conditions, a two-way analysis of variance (ANOVA) was used, since there was one dependent variable (delayed vocabulary test) and two independent variables (learning conditions with two levels and glosses with three levels). The alpha level for the whole study was set at .05. However, because three analysis of variance procedures were conducted, the alpha level was adjusted to .05/3 = .017. In other words, any probability which was equal to or less than this boundary was considered statistically significant for each comparison, so that the experiment-wise alpha could be held steady at approximately .05.

**RESULTS**

**Pretests Analyses**

First, the scores on the cloze and vocabulary pretest of the six conditions were compared to check the equivalence of the groups in terms of their general English
GLOSSING IN INCIDENTAL AND INTENTIONAL LEARNING

proficiency and their prior knowledge of the target words at the beginning of the study. Since three intact classes were divided into the different learning conditions, the investigation of group equivalence was necessary. A total of 158 students participated. The Cronbach’s alpha reliability of the cloze test was .73 and, for the vocabulary pretest, it was .57. The reason for the relatively low reliability on the vocabulary pretest may be the difficulty of the unknown words. Most students received very low scores on the test, ranging from zero to two points out of fifteen points, which means they did not know the target words, just as I had expected. Thus the distribution of scores was skewed positively, and this restriction of range might have depressed the reliability.

Descriptive statistics for the cloze test are presented in Table 2. The total possible score was 25 points. The mean for all the participants was 10.92, and the standard deviation was 4.06.

Table 2
Means and Standard Deviations of the Cloze Test

<table>
<thead>
<tr>
<th>Conditions *</th>
<th>n</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incidental w/ no gloss</td>
<td>28</td>
<td>10.39</td>
<td>2.77</td>
</tr>
<tr>
<td>Incidental w/ L1 gloss</td>
<td>26</td>
<td>11.19</td>
<td>4.54</td>
</tr>
<tr>
<td>Incidental w/ L2 gloss</td>
<td>28</td>
<td>12.39</td>
<td>4.89</td>
</tr>
<tr>
<td>Intentional w/ no gloss</td>
<td>22</td>
<td>9.86</td>
<td>4.20</td>
</tr>
<tr>
<td>Intentional w/ L1 gloss</td>
<td>28</td>
<td>10.96</td>
<td>3.56</td>
</tr>
<tr>
<td>Intentional w/ L2 gloss</td>
<td>26</td>
<td>10.50</td>
<td>4.06</td>
</tr>
</tbody>
</table>

* Incidental w/ no gloss means incidental vocabulary learning group with no gloss.
Intentional w/ no gloss means intentional vocabulary learning groups with no gloss.

The descriptive statistics for the vocabulary pretest are presented in Table 3. Again, the same 158 students took the vocabulary pretest. The total possible score was 15 points on the vocabulary pretest. The mean for all the participants was 1.13 and the standard deviation was 1.44.
Table 3
Means and Standard Deviations of the Vocabulary Pretest

<table>
<thead>
<tr>
<th>Conditions</th>
<th>n</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incidental w/ no gloss</td>
<td>28</td>
<td>1.07</td>
<td>1.09</td>
</tr>
<tr>
<td>Incidental w/ L1 gloss</td>
<td>26</td>
<td>1.38</td>
<td>1.55</td>
</tr>
<tr>
<td>Incidental w/ L2 gloss</td>
<td>28</td>
<td>1.50</td>
<td>2.15</td>
</tr>
<tr>
<td>Intentional w/ no gloss</td>
<td>22</td>
<td>.95</td>
<td>1.21</td>
</tr>
<tr>
<td>Intentional w/ L1 gloss</td>
<td>28</td>
<td>.75</td>
<td>1.17</td>
</tr>
<tr>
<td>Intentional w/ L2 gloss</td>
<td>26</td>
<td>1.08</td>
<td>1.09</td>
</tr>
</tbody>
</table>

A one-way MANOVA was carried out to determine whether there was any significant difference among the six conditions at the beginning of the study as shown in Table 4. Since the three tests showed alpha levels were well above .017, there was no significant difference among the six groups. The six groups were therefore considered approximately equivalent in terms of their level of English proficiency and previous knowledge of the target words, as measured by the cloze test and production of the meaning of the target words.

Table 4
Multivariate Tests for Overall Group Effect.

<table>
<thead>
<tr>
<th>Test Name</th>
<th>Value</th>
<th>Approx. F</th>
<th>Hypoth. DF</th>
<th>Error DF</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pillais</td>
<td>.06</td>
<td>.96</td>
<td>10.00</td>
<td>304.00</td>
<td>.481</td>
</tr>
<tr>
<td>Hotellings</td>
<td>.06</td>
<td>.95</td>
<td>10.00</td>
<td>300.00</td>
<td>.488</td>
</tr>
<tr>
<td>Wilks</td>
<td>.94</td>
<td>.95</td>
<td>10.00</td>
<td>302.00</td>
<td>.484</td>
</tr>
</tbody>
</table>
Main Tests Analyses

The reading comprehension and immediate vocabulary test scores were analyzed using a two-way MANOVA. The total possible score on the reading comprehension test was 20 points, and the total score on the immediate vocabulary posttest was 30 [15 items x 2 versions (L1 and L2 versions)]. As I mentioned earlier, two versions of the vocabulary test were used to minimize the effects of one language or the other. The reliability of the reading comprehension test was found to be .74, and the reliability of the immediate vocabulary test was .80 (using Cronbach alpha). A total of 189 students participated. In this 2 x 3 design (learning with two levels and gloss with three levels), three effects were analyzed: the interaction effect for learning by gloss, the main effect for learning, and the main effect for gloss.

Table 5 shows descriptive statistics for the reading comprehension test for various combinations of learning and gloss.

Table 5
Means and Standard Deviations for the Reading Comprehension Test by Interactions of Learning and Gloss.

<table>
<thead>
<tr>
<th>Conditions</th>
<th>( n )</th>
<th>( M )</th>
<th>( SD )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incidental Total</td>
<td>93</td>
<td>13.27</td>
<td>3.76</td>
</tr>
<tr>
<td>Incidental w/ no gloss</td>
<td>31</td>
<td>13.97</td>
<td>2.98</td>
</tr>
<tr>
<td>Incidental w/ L1 gloss</td>
<td>30</td>
<td>13.87</td>
<td>3.67</td>
</tr>
<tr>
<td>Incidental w/ L2 gloss</td>
<td>32</td>
<td>12.03</td>
<td>4.29</td>
</tr>
<tr>
<td>Intentional Total</td>
<td>96</td>
<td>12.00</td>
<td>3.74</td>
</tr>
<tr>
<td>Intentional w/ no gloss</td>
<td>32</td>
<td>11.78</td>
<td>3.09</td>
</tr>
<tr>
<td>Intentional w/ L1 gloss</td>
<td>34</td>
<td>12.32</td>
<td>4.10</td>
</tr>
<tr>
<td>Intentional w/ L2 gloss</td>
<td>30</td>
<td>11.87</td>
<td>4.03</td>
</tr>
</tbody>
</table>
Table 6 shows descriptive statistics for the immediate vocabulary test for various combinations of learning and gloss.

Table 6
Means and Standard Deviations for the Immediate Vocabulary Posttest by Interactions of Learning and Gloss

<table>
<thead>
<tr>
<th>Conditions</th>
<th>n</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incidental Total</td>
<td>93</td>
<td>18.09</td>
<td>4.97</td>
</tr>
<tr>
<td>Incidental w/ no gloss</td>
<td>31</td>
<td>14.10</td>
<td>2.59</td>
</tr>
<tr>
<td>Incidental w/ L1 gloss</td>
<td>30</td>
<td>22.13</td>
<td>3.18</td>
</tr>
<tr>
<td>Incidental w/ L2 gloss</td>
<td>32</td>
<td>18.16</td>
<td>5.03</td>
</tr>
<tr>
<td>Intentional Total</td>
<td>96</td>
<td>18.70</td>
<td>5.56</td>
</tr>
<tr>
<td>Intentional w/ no gloss</td>
<td>32</td>
<td>14.38</td>
<td>4.47</td>
</tr>
<tr>
<td>Intentional w/ L1 gloss</td>
<td>34</td>
<td>22.18</td>
<td>3.98</td>
</tr>
<tr>
<td>Intentional w/ L2 gloss</td>
<td>30</td>
<td>19.37</td>
<td>5.15</td>
</tr>
</tbody>
</table>

Multivariate analysis of variance indicated that there was no significant interaction effect for the two variables, learning and gloss, on the reading comprehension test or the immediate vocabulary posttest ($p > .017$) as shown in Table 7.

Second, multivariate analysis of variance was used to investigate whether there was any significant effect for learning conditions on either the reading comprehension test or the immediate vocabulary posttest as shown in Table 8. All three multivariate statistics had alpha levels below $.017$.

Thus univariate $F$-tests were carried out to determine which tests were significantly affected by the learning conditions as shown in Table 9. The result of these $F$-tests showed that the alpha level for the reading comprehension test was very close to the boundary of significance, but was in fact not statistically significant. The probability level for the immediate vocabulary test was high enough to indicate that the different learning conditions probably did not influence vocabulary learning ($p = .402$).
### Table 7
**Multivariate Tests for Effect of Learning by Gloss**

<table>
<thead>
<tr>
<th>Test Name</th>
<th>Value</th>
<th>Approx. F</th>
<th>Hypoth. DF</th>
<th>Error DF</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pillais</td>
<td>.01</td>
<td>.65</td>
<td>4.00</td>
<td>366.00</td>
<td>.626</td>
</tr>
<tr>
<td>Hotellings</td>
<td>.01</td>
<td>.65</td>
<td>4.00</td>
<td>362.00</td>
<td>.629</td>
</tr>
<tr>
<td>Wilks</td>
<td>.99</td>
<td>.65</td>
<td>4.00</td>
<td>364.00</td>
<td>.627</td>
</tr>
</tbody>
</table>

### Table 8
**Multivariate Tests for Main Effect for Learning**

<table>
<thead>
<tr>
<th>Test Name</th>
<th>Value</th>
<th>Approx. F</th>
<th>Hypoth. DF</th>
<th>Error DF</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pillais</td>
<td>.05</td>
<td>4.33</td>
<td>2.00</td>
<td>182.00</td>
<td>.015</td>
</tr>
<tr>
<td>Hotellings</td>
<td>.05</td>
<td>4.33</td>
<td>2.00</td>
<td>182.00</td>
<td>.015</td>
</tr>
<tr>
<td>Wilks</td>
<td>.96</td>
<td>4.33</td>
<td>2.00</td>
<td>182.00</td>
<td>.015</td>
</tr>
</tbody>
</table>

### Table 9
**Univariate F-tests with (1,183) DF for Main Effect for Learning**

<table>
<thead>
<tr>
<th>Variable*</th>
<th>Hyp. SS</th>
<th>Error SS</th>
<th>Hyp. MS</th>
<th>Error MS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>R.C. Test</td>
<td>79.47</td>
<td>2549.78</td>
<td>79.47</td>
<td>13.93</td>
<td>5.70</td>
<td>.018</td>
</tr>
<tr>
<td>Im. Voc. Test</td>
<td>12.30</td>
<td>3189.80</td>
<td>12.30</td>
<td>17.43</td>
<td>.71</td>
<td>.402</td>
</tr>
</tbody>
</table>

* R.C. Test = reading comprehension test; Im. Voc. Test = immediate vocabulary test.
Third, the effect for the gloss was investigated. Descriptive statistics are presented in Tables 10 and 11.

Table 10
*Means and Standard Deviations for the Reading Comprehension Test in Different Gloss Conditions*

<table>
<thead>
<tr>
<th>Conditions</th>
<th>$n$</th>
<th>$M$</th>
<th>$SD$</th>
</tr>
</thead>
<tbody>
<tr>
<td>No gloss</td>
<td>63</td>
<td>12.86</td>
<td>3.21</td>
</tr>
<tr>
<td>L1 gloss</td>
<td>64</td>
<td>13.05</td>
<td>3.95</td>
</tr>
<tr>
<td>L2 gloss</td>
<td>62</td>
<td>11.95</td>
<td>4.13</td>
</tr>
</tbody>
</table>

Table 11
*Means and Standard Deviations for the Immediate Vocabulary Posttest in Different Gloss Conditions*

<table>
<thead>
<tr>
<th>Conditions</th>
<th>$n$</th>
<th>$M$</th>
<th>$SD$</th>
</tr>
</thead>
<tbody>
<tr>
<td>No gloss</td>
<td>63</td>
<td>14.24</td>
<td>3.64</td>
</tr>
<tr>
<td>L1 gloss</td>
<td>64</td>
<td>22.16</td>
<td>3.60</td>
</tr>
<tr>
<td>L2 gloss</td>
<td>62</td>
<td>18.74</td>
<td>5.08</td>
</tr>
</tbody>
</table>
Multivariate analysis of variance was performed to investigate the significance of differences due to glossing conditions on either reading comprehension or immediate vocabulary posttest as shown in Table 12. The results of all three multivariate statistics indicated that there were significant differences \( p < .017 \) somewhere in the results.

### Table 12

**Multivariate Tests for Main Effect of Gloss**

<table>
<thead>
<tr>
<th>Test Name</th>
<th>Value</th>
<th>Approx. F</th>
<th>Hyp. DF</th>
<th>Error DF</th>
<th>( p )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pillais</td>
<td>.43</td>
<td>24.98</td>
<td>4.00</td>
<td>366.00</td>
<td>.000</td>
</tr>
<tr>
<td>Hotellings</td>
<td>.72</td>
<td>32.42</td>
<td>4.00</td>
<td>362.00</td>
<td>.000</td>
</tr>
<tr>
<td>Wilks</td>
<td>.58</td>
<td>28.66</td>
<td>4.00</td>
<td>364.00</td>
<td>.000</td>
</tr>
</tbody>
</table>

Hence, univariate F-tests were performed to locate which test was influenced by the glossing conditions as shown in Table 13. The result of univariate F-tests indicated glossing did not have a significant impact on reading comprehension. However, there was a significant effect on the immediate vocabulary posttest for glossing conditions. Scheffé post hoc analysis was performed in order to locate specific significant differences among the three different gloss conditions. The Scheffé analysis indicated significant differences between all three possible comparisons. In other words, there was a significant difference between glossed conditions (L1 gloss and L2 gloss) and the no gloss condition, as well as between L1 gloss and L2 gloss conditions.

### Delayed Test Analyses

A two-way ANOVA was conducted to analyze the delayed vocabulary test. The total possible points was 30 [(15 items x 2 versions (L1 and L2)]. The reliability of the delayed vocabulary test was .79. A total of 189 students, the same students who were in the immediate posttests, also participated in the delayed vocabulary test. The descriptive statistics are reported in Table 14.
Table 13
Univariate F-tests for Main Effect of Gloss (with 2,183 df)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Hyp. SS</th>
<th>Error SS</th>
<th>Hyp. MS</th>
<th>Error MS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>R.C. Test</td>
<td>46.24</td>
<td>2549.78</td>
<td>23.12</td>
<td>13.93</td>
<td>1.66</td>
<td>.193</td>
</tr>
<tr>
<td>Im. Voc. Test</td>
<td>1999.05</td>
<td>3189.80</td>
<td>999.53</td>
<td>17.43</td>
<td>57.34</td>
<td>.000*</td>
</tr>
</tbody>
</table>

* p < .017

Table 14
Means and Standard Deviations for the Delayed Vocabulary Posttest for Various Combinations of Learning and Gloss.

<table>
<thead>
<tr>
<th>Conditions</th>
<th>n</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incidental Total</td>
<td>93</td>
<td>17.58</td>
<td>5.09</td>
</tr>
<tr>
<td>Incidental w/ No gloss</td>
<td>31</td>
<td>14.42</td>
<td>3.02</td>
</tr>
<tr>
<td>Incidental w/ L1 gloss</td>
<td>30</td>
<td>21.07</td>
<td>4.56</td>
</tr>
<tr>
<td>Incidental w/ L2 gloss</td>
<td>32</td>
<td>17.38</td>
<td>5.17</td>
</tr>
<tr>
<td>Intentional Total</td>
<td>96</td>
<td>18.43</td>
<td>5.28</td>
</tr>
<tr>
<td>Intentional w/ No gloss</td>
<td>32</td>
<td>14.38</td>
<td>4.36</td>
</tr>
<tr>
<td>Intentional w/ L1 gloss</td>
<td>34</td>
<td>22.12</td>
<td>4.03</td>
</tr>
<tr>
<td>Intentional w/ L2 gloss</td>
<td>30</td>
<td>18.57</td>
<td>4.28</td>
</tr>
</tbody>
</table>

Univariate analysis of variance was conducted to investigate any significant differences on the delayed vocabulary test for either learning or glossing conditions as shown in Table 15. The results indicated that there was no significant effect for either the two-way interaction between learning and gloss or the main effect for learning. However, the delayed vocabulary test was again affected by glossing.
GLOSSING IN INCIDENTAL AND INTENTIONAL LEARNING

conditions ($p < .017$). Scheffé tests, performed to locate specific significant differences among the three different gloss conditions, indicated the same significant differences between all three possible comparisons as exhibited on the immediate vocabulary posttest. Therefore, the glosses do appear to affect vocabulary learning over time.

Table 15
Univariate Tests for Main Effects of Learning and Gloss and Their Interaction

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>SS</th>
<th>DF</th>
<th>MS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learning</td>
<td>25.24</td>
<td>1</td>
<td>25.24</td>
<td>1.37</td>
<td>.243</td>
</tr>
<tr>
<td>Gloss</td>
<td>1650.44</td>
<td>2</td>
<td>825.22</td>
<td>44.93</td>
<td>.000*</td>
</tr>
<tr>
<td>Learning x Gloss</td>
<td>14.38</td>
<td>2</td>
<td>7.19</td>
<td>.39</td>
<td>.677</td>
</tr>
<tr>
<td>Residual</td>
<td>3361.31</td>
<td>183</td>
<td>18.37</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>5059.98</td>
<td>188</td>
<td>26.92</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* $p < .017$

Additional Analyses

Additional analyses were conducted to investigate the relationship between immediate and delayed vocabulary posttests. Table 16 presents the descriptive statistics of vocabulary tests conducted at two different times.

Figure 2 helps to show more clearly the relationship between immediate and delayed vocabulary posttests. Figure 2 shows the change in mean scores of six conditions over time. The mean scores of the delayed vocabulary posttest are either lower than or the same as the mean scores of the immediate vocabulary posttest, except for the original text in the incidental vocabulary learning group. In general, the intentional vocabulary learning group had higher mean scores over time compared to the incidental vocabulary learning group.
Table 16  
Means and Standard Deviations for Vocabulary Test Scores Over Time  
(N=189)

<table>
<thead>
<tr>
<th>Conditions</th>
<th>Posttest n</th>
<th>Posttest M (SD)</th>
<th>Delayed M (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incidental w/ no gloss</td>
<td>31</td>
<td>14.10 (2.59)</td>
<td>14.42 (3.02)</td>
</tr>
<tr>
<td>Incidental w/ L1 gloss</td>
<td>30</td>
<td>22.13 (3.18)</td>
<td>21.07 (4.56)</td>
</tr>
<tr>
<td>Incidental w/ L2 gloss</td>
<td>32</td>
<td>18.16 (5.03)</td>
<td>17.38 (5.17)</td>
</tr>
<tr>
<td>Intentional w/ no gloss</td>
<td>32</td>
<td>14.38 (4.47)</td>
<td>14.38 (4.36)</td>
</tr>
<tr>
<td>Intentional w/ L1 gloss</td>
<td>34</td>
<td>22.18 (3.98)</td>
<td>22.12 (4.03)</td>
</tr>
<tr>
<td>Intentional w/ L2 gloss</td>
<td>30</td>
<td>19.37 (5.15)</td>
<td>18.57 (4.28)</td>
</tr>
</tbody>
</table>

Figure 2. Mean Scores of the Vocabulary Tests Over Time
In order to investigate how students dealt with the gloss when they were reading the text, I included an additional question at the end of the immediate vocabulary posttest (L1 version). For the students who were given the reading material with either the L1 or L2 gloss, I asked them to choose from among the following four options: (a) I never looked at the definitions or clues, (b) I saw some of the definitions or clues, (c) I saw all the definitions or clues, and (d) I tried hard to memorize the definitions or clues. Table 17 shows the results of their responses.

Table 17

<table>
<thead>
<tr>
<th>Occasions*</th>
<th>Incidental L1</th>
<th>L2</th>
<th>Total</th>
<th>Intentional L1</th>
<th>L2</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Never</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Some</td>
<td>15</td>
<td>15</td>
<td>30</td>
<td>18</td>
<td>16</td>
<td>34</td>
</tr>
<tr>
<td>All</td>
<td>12</td>
<td>11</td>
<td>23</td>
<td>11</td>
<td>7</td>
<td>18</td>
</tr>
<tr>
<td>All with effort</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>No response</td>
<td>1</td>
<td>3</td>
<td>4</td>
<td>0</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Total</td>
<td>30</td>
<td>32</td>
<td>62</td>
<td>34</td>
<td>30</td>
<td>64</td>
</tr>
</tbody>
</table>

*Never = never looked at the definitions or clues.
Some = looked at some of them.
All = looked at all of them.
All with effort = looked at all of them and tried hard to memorize them.
No response = students who did not indicate any choice listed above.

DISCUSSION

The first research question was about the relationship between different learning conditions and foreign language vocabulary learning. As shown in Table 6, the mean scores of the intentional vocabulary learning group were slightly higher than the mean scores for incidental vocabulary learning ($M = 18.70$ vs. $M = 18.09$); however, the difference was not significant. Thus hypothesis 1 was not supported.
The fact that there was no significant effect for different learning conditions on vocabulary learning can be explained by the following. There was no significant difference between the two different learning groups in terms of the attention that students paid to glosses. As indicated in Table 17, the students’ responses to the question about their reaction to the glosses showed that there was not a wide gap between the two groups in terms of the amount of effort they put into the glosses. I had expected that most students in the intentional vocabulary learning group would pay more attention to the glosses and study hard to memorize the definitions of target words. It seemed that some students in the intentional vocabulary learning group might have thought that they could find some clues or definitions while reading. Hence they might have read the passages wondering what kinds of vocabulary would appear on the vocabulary test later. The students who thought the listed words in the glosses were important vocabulary might have studied harder. However, those who did not think that way might have wondered about the vocabulary while reading the passage.

The reason that the intentional vocabulary learning group in Konopak et al. (1987) did much better than their incidental vocabulary learning group, as opposed to a very slight difference between the two groups in my study, could have been the result of different task instructions. In their study, the task conditions between intentional and incidental vocabulary learning groups were very different from mine. The intentional vocabulary learning group read reading material which drew attention to the target words (underlined target words) and performed tasks relevant to the evaluation (writing down the meaning of target words, guessing from the context), whereas the incidental learning group just read the reading material without having the target words emphasized. In other words, the nature of their intentional vocabulary learning condition was designed to favor the students’ performance over the incidental vocabulary learning condition. However, the intentional and incidental conditions in my study were designed solely to provide different objectives for each group in advance. This difference in task instruction could explain the different outcomes between their study and my study.

The second research question was whether the different learning conditions were related to foreign language reading comprehension. The results indicated that the performance difference between the incidental vocabulary learning group and the intentional vocabulary learning group on the reading comprehension posttest was not significant. Thus hypothesis 2 was not supported. However, because Table 9 showed
that the probability for the reading comprehension test was very close to the boundary of a significant alpha level, it seems that the learning condition might be related in some way to the reading comprehension test.

From another point of view, the incidental vocabulary learning group was an intentional reading comprehension group. Since students in the incidental vocabulary learning group were told that they would have a reading comprehension test, they might have focused on the content only, whereas those in the intentional vocabulary learning group were told that they would have a vocabulary test and were given the same amount of time, so they might have focused both on the content and vocabulary simultaneously. Although the difference was not significant, the finding that the incidental vocabulary learning group performed better on the reading comprehension test than the intentional vocabulary learning group indicates that focused attention (attention on the content only) may result in better performance than divided attention (attention on both the content and vocabulary).

The third research question explored the relationship between the different gloss conditions and foreign language vocabulary learning. As shown in Table 13, there was a significant effect for gloss on the immediate vocabulary posttest. As shown in Table 11, the mean scores of the L1 gloss condition \((M = 22.16)\) were the highest, while the mean scores of the L2 gloss condition \((M = 18.74)\), and the no gloss condition \((M = 14.24)\) followed accordingly. There was a significant effect between glossed conditions (L1 and L2 glosses) and the no gloss condition, as well as between L1 gloss and L2 gloss conditions. Hence hypothesis 3a was supported, but hypothesis 3b was not.

With respect to the glossed conditions and no gloss condition, the result was similar to the results of studies done by Watanabe (1992) and Jacobs and Dufon (1990), in which the students in the gloss conditions outperformed the students in the original text. However, in term of the L1 gloss and L2 gloss, there was a difference between Jacobs and Dufon's study and my study. In their study, there was no significant difference between the L1 gloss and L2 gloss conditions, whereas the results of my study indicated a statistical significance between L1 gloss and L2 gloss conditions.

The fourth research question was whether the different gloss conditions would also affect the foreign language reading. Although my study differed from Jacobs and Dufon in how it accessed the students' reading comprehension, using writing recalls (in their study) versus multiple-choice tests (in mine), my findings were similar to
As shown in Table 13, there was no significant effect for the different gloss conditions on the reading comprehension test. Thus hypothesis 4 was supported. In the Jacobs and Dufon study, the sequence of performance from high to low was L2, L1, and no gloss conditions. My findings, presented in Table 11, showed a different order, that is, L1 gloss, no gloss, and L2 gloss. The gap between L1 and no gloss conditions was very slight, and comparatively, the L2 gloss condition was much lower.

Apparently, compared with the no gloss condition, if the glosses are written in L1, this effect on reading comprehension was almost the same as the no gloss condition, whereas the L2 glosses worsened the students' reading comprehension. In the case of L1 glosses, the duration of time spent glimpsing the definitions might have been shorter, compared to the L2 glosses, because the students could more easily understand the definitions. However, in the case of the L2 glosses, the students might find the meanings in the definitions rather obscure and difficult, which could cause relatively long interruptions in reading. The relatively long disruptions in the reading flow might in turn have made the students less efficient readers.

The fifth research question addressed whether the vocabulary learned under the different gloss conditions could be retained after a week-long interval. As shown in Table 16, in general, there was only a slight decrease in terms of the mean scores of the vocabulary test taken after a week compared to the mean scores of the immediate vocabulary test. This demonstrated that the retention of the learned words was possible after a week. Thus hypothesis 5 was supported, which is what earlier studies (e.g., Watanabe, 1992; Ha, 1992) have shown. One thing I would like to point out is that, although I rearranged the order of items and distracters to minimize any practice effect, that effect may still exist. One solution to this problem would be to use new distracters, at least two among three, as well as to rearrange the order of items and the existing distracters when making a delayed vocabulary test using the same target words.

The sixth research question asked whether the results for the delayed vocabulary test was the same as for the immediate vocabulary test. Table 15 indicated that there was a significant effect for the different gloss conditions on the delayed vocabulary posttest \((p < .017)\). This finding confirmed that the gloss conditions were related to the vocabulary learning. Scheffé post-hoc analysis showed that there was statistical significance between both glosed conditions (L1 and L2 glosses) and the no gloss condition, as well as between the L1 gloss and L2 gloss conditions. Thus hypothesis
6a was supported, but hypothesis 6b was not. In the comparing the mean scores for the six different conditions between immediate and delayed vocabulary posttests, the students received either the same or lower mean scores on the delayed vocabulary posttest than they did on the immediate vocabulary posttest except for one case in which the incidental vocabulary learning group with no gloss had slightly higher mean scores than the intentional vocabulary learning group with no gloss. In short, the overall mean scores on the delayed vocabulary test were lower than the overall mean scores on the immediate vocabulary test. Thus hypothesis 6c was supported.

CONCLUSION AND IMPLICATIONS

This study found that there was no significant effect for different learning conditions on either vocabulary learning or reading comprehension, and revealed a significant effect between different gloss forms and vocabulary learning in an EFL situation. Again, the different forms of gloss were significantly related to the vocabulary test given after one week. The study also confirmed previous findings that the retention of learned vocabulary was possible for at least one week.

The first issue that I would like to address is the relationship of the EFL students’ focused attention and divided attention on vocabulary learning and reading comprehension. Those subjects who were informed about the reading comprehension test—the focused attention group—did better on that test than the subjects who were told that there would be a vocabulary test—the divided attention group. The results, while not statistically significant, were very close to the boundary of significance. Thus, if a more powerful design had been used, the students’ distribution of allocated attention during the task might have significantly affected their performance later.

Although two different learning conditions stimulated different aspects of students’ attention depending on what was regarded as important, there was no statistical significance between different learning conditions in either the vocabulary learning or reading comprehension. With respect to the vocabulary learning, the additional survey in Table 17 showed that the glosses seemed to attract students’ attention regardless of intentional or incidental learning conditions. It seemed that the extra notes in the margin might have attracted the students’ attention and made them look at those while reading. Thus it is recommended that future research should involve designing different learning tasks in which students can truly be involved. In other words, students in the incidental group are supposed to learn the vocabulary as a by-
product of doing something else, whereas those in the intentional group are supposed to learn words with a deliberate plan.

It is the same case for the reading comprehension. Future research should involve designing a task in which students not only have different purposes but also focus on that purpose based on the different learning conditions. In short, the learning conditions should be carefully designed to involve students in what they are supposed to be doing.

Along with carefully designed task conditions, I would like to provide two suggestions for a better study in the future. One is the use of randomized classes. Even though I did not find any significant differences among the three intact classes at the beginning of the study, it would be worthwhile to investigate the same research questions with truly randomized groups. The second suggestion concerns different ways of assessing students’ prior knowledge of the target words. As I mentioned in the results section, the vocabulary pretest turned out to have fairly low reliability ($r = .57$) perhaps due to its level of difficulty. Hence in future research, the level of difficulty of the test should be better matched to the abilities of the students in order to spread the scores out fairly, and thus to have a more reliable test. As I experienced in my study, either too difficult or too easy a test may cause low reliability. Hence it may be worthwhile to think in advance of ways to measure with better reliability.

Even if I did not find any significant difference for different learning conditions, on the whole, there was a positive direction for intentional learning over incidental learning. As I discussed earlier in the literature review, the merits of intentional learning over incidental learning seem to be based on the greater attention given to the reading passage. The students in the incidental vocabulary learning group (the intentional reading comprehension group) might have paid more and better attention, in terms of quantity and quality, at the time of reading the passage. This might have been true also for the vocabulary learning. The overall mean scores of the intentional vocabulary learning condition were better on both the immediate and delayed vocabulary tests.

These findings imply that drawing students’ attention when they begin a task can have a positive impact on learning outcomes. If the students are old enough to understand, it may be desirable to state a clear objective or purpose for the activity. By doing this, the students can have a better understanding of why they are given the particular activity. Then this understanding may result in positive outcomes. If the
GLOSSING IN INCIDENTAL AND INTENTIONAL LEARNING

activity is routinized and they know its purpose, then it may not be necessary to tell them every time. Otherwise, it may be worthwhile to make the students aware of why they are doing the activity.

The second issue that I would like to discuss is the relationship between different gloss conditions and vocabulary learning for EFL learners. Besides students' preference for glosses in the foreign language reading materials as reported in Jacobs and Dufon's study (1990), the results of my study and previous studies (Watanabe, 1992; Jacobs & Dufon, 1990) have shown that providing glosses in the reading materials facilitates foreign language vocabulary learning.

With respect to the L1 and L2 glosses, L1 glosses were more efficient for vocabulary learning and at the same time they did not impede reading comprehension. One reason that the students did not do as well on the L2 gloss in my study could be attributed to their habitual use of a bilingual dictionary (English and Korean) instead of a monolingual dictionary (English only). Because they were not accustomed to reading L2 definitions of words, they might have understood the L2 definitions less clearly than the L1 definitions. If the students could understand the definitions clearly, then the L2 glosses might be just as (or even more) effective than the L1 glosses. The L2 glosses might be more effective because, if a foreign word is a culture specific word which is difficult to explain in the L1, then it might be more efficient for the definitions to be explained in the target language. Since the semantic boundary of a foreign word may not always be the same as the corresponding native language word, explanation in the target language may be clearer and more efficient. For beginners or intermediate learners of a foreign language, L1 glosses may be efficient because, if they do not understand the written definitions, they may not learn from them. This interpretation is based on the theory that, if the input is not comprehensible, then the students will not able to learn from it. However, in the case of the advanced learners, L2 glosses, if understandable, may be efficient. Actually, the majority of college students in Jacobs and Dufon's study indicated their preference for L2 glosses over L1 glosses if they could understand the explanation of the words in the L2.

Several studies have demonstrated that incidental vocabulary learning is possible while reading. Due to the fact that the semantic boundary and collocation of the foreign words can be different from the corresponding native language words, learning words in context may be helpful in order to grasp the full range of meaning, including connotations. However, just deriving the meaning of unknown words might
be misleading and risky because the contexts sometimes do not provide enough contextual cues, as I stated earlier in the literature review. Hence learning words through reading with glosses may be the most efficient way to learn new words. As my study has demonstrated, students can certainly learn new vocabulary while reading if glosses are presented. Therefore, people involved in the L2 teaching field should be aware of this benefit and consider it in providing reading materials with glosses when encouraging students to learn vocabulary.

One problematic issue regarding glosses may be their limited usage in the real world. On the whole, providing glosses is limited to the L2 teaching materials; the students may not receive help in their voluntary for reading from various sources outside of the classroom. Besides, there can be drawbacks in selecting target words for glosses. There may exist a difference between how the words are selected and the students’ corresponding needs. People involved in materials design in L2 teaching have to consider this fact and need to conduct research before selecting words for glosses.

REFERENCES


GLOSSING IN INCIDENTAL AND INTENTIONAL LEARNING


Myong Hui Ko
403-024
Inchon Kwong yuk-Si Bupyong-Gu
Sangok-Dong 294
Woosung Apt 108-1305
South Korea

e-mail kos@hawaii.edu