Effects of Text Modification and Learning Condition on EFL Vocabulary Learning

Youngkyu Kim
(University of Hawaii at Mānoa)


This study investigated whether reading elaborated text would facilitate intentional and incidental learning of foreign language vocabulary. 211 Korean learners of English as a foreign language at university level were asked to read unmodified or elaborated text under an intentional or an incidental vocabulary learning condition and to take a reading comprehension test. Right after the treatment, an immediate vocabulary posttest, and one week later, both a delayed vocabulary posttest and a cloze test, were administered. The results indicated that: a) incidental vocabulary learning through reading was possible with these learners although the amount of such learning was small, but significant; b) elaborated text better facilitated vocabulary learning from reading than unmodified text although the difference between the two text types was not statistically significant; and c) intentional vocabulary learning from reading was better than incidental vocabulary learning, regardless of the text modification type, on both immediate and delayed vocabulary posttests. This study also suggested pedagogic implications of the research findings to foreign language reading and vocabulary teaching.

I. INTRODUCTION

Vocabulary has been considered a stumbling-block to, or a springboard for, successful foreign language (FL) learning. Improving proficiency in foreign language vocabulary has thus long been one of the key objectives in foreign language instruction. Reflecting this pedagogical interest in vocabulary, issues of how to teach and learn vocabulary effectively and efficiently

* I would like to express my gratitude to J. D. Brown, T. Hudson, M. Long, and three anonymous reviewers for their constructive feedback to an earlier version of this article.
have been much discussed in the native and foreign language vocabulary instruction literature (e.g., McKeown & Curtis, 1987; Nation, 2001).

As a default explanation of the fact that students learn words not explicitly taught to them, vocabulary learning researches have proposed an incidental vocabulary learning hypothesis, which has been empirically tested in both first language (L1) and second language (L2) acquisition (e.g., Herman, Anderson, Pearson, & Nagy, 1987; Huckin & Coady, 1999; Hulstijn, 2001, 2003). This incidental learning hypothesis is closely related to a vocabulary-learning-from-context hypothesis (e.g., Krashen, 1989; Sternberg, 1987), whereby the amount of incidental vocabulary learning can be expected to depend crucially on the quality and quantity of supportive contextual clues surrounding unfamiliar lexical items embedded in written or spoken text. This has also led L1 and L2 researchers and practitioners alike to focus on context modification and its effects on reading comprehension and vocabulary learning.

Reading is still widely believed to be one of the best ways to learn syntactic, semantic, discoursal, or lexical aspects of English in English as a foreign language (EFL) situations where written input via reading materials is more readily available to EFL learners than oral input. Hence, the use of reading materials originally written with native speakers (NSs) of English in mind can be expected to play a pedagogical role by providing a wealth of helpful input. However, in fact, a recent review of L2 research into text modification suggests that learning English by reading text originally written for NSs of English may not always result in full understanding of the information contained in the text, let alone learning of new lexical and grammatical items because the input contained in such text is often not comprehensible to EFL learners (e.g., Yano, Long, & Ross, 1994). This problem has led to the development and widespread use of modified graded readers in EFL reading instruction. The traditional practice in text modification has been to simplify reading materials (see e.g., Simensen, 1984, 1987a, 1987b, for the principles of text modification adopted by several British English Language Teaching (ELT) publishers of graded readers). However, text simplification often results in the distorted discourse patterns (e.g., see discussion in Yano, et al., 1994; Long & Ross, 1992) and the elimination of linguistic items which, while difficult, are nonetheless necessary for foreign language development (e.g., Simensen, 1990, 1992; Yano, et al., 1994). As a way of overcoming disadvantages of such text simplification, text elaboration has been suggested on the basis of research on the interactional modification of spoken discourse by NSs conversing with NNSs (nonnative speakers) (e.g., Larsen-Freeman & Long, 1991, chap. 5).

Research on text elaboration has great pedagogical potential for suggesting ways to make input comprehensible and at the same time keep nutrients for language acquisition intact, as in the case of interactional modification, and when combined with research on vocabulary learning, can provide answers to whether an increase in comprehension of input leads to learning of that
input. As a continuation of this line of research, the present study investigates the effects of text elaboration on EFL vocabulary learning through reading.

II. REVIEW OF THE LITERATURE

1. Incidental L2 Vocabulary Learning

Incidental learning, as in the form of an incidental vocabulary learning hypothesis, is defined as "learning 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. 16). It has been proposed as one of the explanations to a question of how L1 and L2 learners learn vocabulary. In support of the hypothesis, well-documented evidence from research on child L1 vocabulary learning (cf. Beck & McKeown, 1991; McKeown & Curtis, 1987) reveals that only a fraction of the large amount of vocabulary that children learn during the school years can be explained causally by explicit vocabulary instruction. The majority of the vocabulary that children learn is thus believed to be incidentally learned, while they are primarily engaged in reading or listening to a variety of oral or written input to comprehend what is going on. Thus, research on incidental learning of vocabulary from reading or listening as one of the major sources for vocabulary development has been of considerable interest to L1 reading and vocabulary researchers (see e.g., Elley, 1989, for incidental vocabulary learning from listening and Herman, et al., 1987; Jenkins & Dixon, 1983; Jenkins, Stein, & Wysocki, 1984; Konopak, Sheard, Longman, Lyman, Slaton, & Atkinson, 1987; Nagy & Anderson, 1984; Nagy, Herman, & Anderson, 1985; Nagy, Anderson, & Herman, 1987; Shu, Anderson, & Zhang, 1995, for incidental vocabulary learning from reading).

The incidental learning hypothesis (Herman, et al., 1987, p. 264) has also been empirically tested in SLA research in general (cf. Hulstijn, 1989a, 1989b) and in L2 vocabulary learning research in particular (see e.g., Neuman & Koskinen, 1992, for incidental vocabulary learning from captioned television, Ellis, Tanaka, & Yamazaki, 1994, for incidental vocabulary learning through oral interaction, and Day, Omura, & Hiramatsu, 1991; Dupuy & Krashen, 1993; Hulstijn, 1992; Knight, 1994; Pitts, White, & Krashen, 1989; Sherfer, 1993; Watanabe, 1997, for incidental vocabulary learning from reading). These studies have shown that it is possible for L2 learners to learn L2 vocabulary in an incidental learning condition where no particular instructional emphasis is given to learning vocabulary, while L2 learners’ primary focus is on comprehending oral or written input. However, when an explicit comparison of incidental and intentional vocabulary learning from context was made, greater gains by intentional learning
groups over incidental learning groups have been reported (e.g., Hulstijn, 1992; Konopak, et al., 1987).

As shown above, there has been a paucity of both L1 and L2 studies on learning vocabulary incidentally in a spoken context (i.e., listening), compared to those studies in a written context (i.e., reading). This is clearly due to the fact that “vocabulary research has existed as a strand of reading research while studies of oral language development tend to have been the domain of developmental psychologists interested in how young children begin to acquire language” (Beck & McKeown, 1991, p. 799). As a result, incidental vocabulary learning research has been predominantly conducted in a written context, especially in conjunction with reading comprehension research, where the relationship between vocabulary knowledge and reading comprehension has long been of great pedagogical and research interest. It has been demonstrated in the L1 reading research literature that there is a strong correlation between vocabulary knowledge and reading comprehension, and a similar claim has also been made in the L2 reading research literature (e.g., Davis, 1944; Singer, 1965; Spearit, 1972; Thurstone, 1946). Given this strong relationship, the quality of the written context from which vocabulary is learned incidentally through normal reading can be expected to affect reading comprehension and ultimately vocabulary learning (Sternberg & Powell, 1983).

2. L2 Vocabulary Learning from Modified Text

Drawing from the available research findings concerned with learning vocabulary in a written context, Ellis (1994b, pp. 22-23) identifies four general factors potentially affecting L2 vocabulary learning from oral as well as written input: (1) intrinsic word properties (i.e., part of speech, distinctiveness of word form, length of word form, correlation between form and meaning, and imageability); (2) input factors (i.e., frequency, saliency through focus, availability of contextual cues, and input complexity); (3) interactional factors (i.e., resulting in more input and resulting in elaborated output); and (4) learner factors (i.e., background knowledge, procedural knowledge, immediate phonological memory, and the learner’s L1). Of these four factors, input factors, including, especially, the availability of helpful contextual cues and the complexity of the written input in which unknown words are embedded, can be a strong candidate for the main determinant of successful L2 vocabulary learning from written input.

However, it has been shown that an unmodified written context, where vocabulary is embedded, is not “considerate” (Armbuster, 1984) and sometimes even of little or no use to vocabulary learning (Beck, McKeown, & McCaslin, 1983; Schatz & Baldwin, 1986), not to mention the fact that inconsiderate text often significantly misleads its readers and hinders successful reading comprehension. With respect to reading authentic text, Scherfer (1993, p.
1148) contends that "it is very possible that a learner will never encounter the specific context or all the necessary different contexts that allow him [sic] to infer the grammatical information necessary in order to really know a word." It is on this basis that much L1 reading research on text revision has been actively conducted to help children learn new concepts in school textbooks more accurately (see e.g., Sawyer, 1991, for a comprehensive review of L1 research on text modification). In a similar vein, L2 research on text modification, variously called text adaptation, abridgment, elaboration, rewriting, or simplification (cf. Simensen, 1990), has been conducted in an attempt to investigate how comprehensible oral or written input to NNSs affects their comprehension of the target language and development of their interlanguage systems (see e.g., Yano, et al., 1994, for a comprehensive review of L2 research on text modification), which is a continuation of one of the most fruitful and promising SLA research endeavors: Research on the role of input and interaction in SLA.

Long, through his pioneering work on this specialized research area, argues that "[s]atisfactory evidence of the a [input modification] → b [comprehensible input] and b → c [acquisition] relationships would allow the linguistic environment to be posited as an indirect causal variable in SLA" (1985, p. 378). His argument is concerned with how linguistic or conversational adjustments by NSs to input addressed to NNSs are related to L2 comprehension and acquisition. The first step of his argument has already been supported by a number of studies which have shown that various types of input or interactional modifications facilitate L2 comprehension to a certain extent, but the effects of such modifications depend relatively on the specific type of input modification.

In contrast to this situation, the SLA literature on this area (see, e.g., Ellis, 1994a; Gass, 1997, 2003; Larsen-Freeman & Long, 1991; Long, 1985, 1996; Lynch, 1996) indicates that more research is still needed to produce empirical evidence for the second step (i.e., Does comprehensible input promote acquisition?) of Long's three-step argument (1985). For instance, Lynch (1996, p. 73) points out that "[t]here have been some Step 2 studies, looking at whether successful modifications promote progress in the target language, but exploration of this area is still at an early stage." Thus, it is essential to demonstrate the validity of the second step to verify Long's claim, upon which much SLA research into text modification and vocabulary learning is based.

Discussing the pros and cons of using simplified, though not elaborated, text in foreign language instruction, Coady (1991), however, points out "a classic dilemma: the simplified text runs the risk of creating a distorted version of L2 lexical patterning in actual discourse; the realistic text with its authentic lexical cohesion and pragmatics can be too much realism too soon with the consequence that the learner is lost" (p. 26). Thus, in an attempt to avoid this dilemma, text elaboration, a process of "adding redundancy to discourse through the use of repetition, paraphrases and appositionals, and by making semantic structure more explicit" (Long, 1996) is
preferred as a way to modify text to text simplification, a process of using “shorter, syntactically less complex utterances or sentences, use of a narrower range of verb tenses, fewer modifiers, and frequently some loss of semantic context” (Long, 1996). This preference is due to the fact that the former is a pedagogically more useful way of making L2 reading materials conducive to learning of L2 vocabulary or syntax than the latter, in that the former does not delete or eliminate complex syntactic structures or low-frequency lexical items which are essential for L2 development.

To show how the definition of text elaboration is actually applied to text elaboration processes, two sample paragraphs each from the unmodified NS baseline experimental passage and the elaborated experimental passage used in the present study are presented below, where target words are boldfaced only in the sample paragraphs.

Example 1. Unmodified NS baseline experimental passage:

But experts contend that human ingenuity may only have postponed the day of reckoning. As the longstanding efforts to revive the bay suggest, technology is better at cleaning up industrial messes than at solving the intractable problems caused by millions of people going about their everyday activities.

Example 2. Elaborated experimental passage:

But experts contend, or put forward the following as their opinion, that human ingenuity, or cleverness in solving problems, may only have postponed the day of reckoning, the time when human beings have to pay for what they have done to the environment. What the long-standing efforts to revive the bay, or to bring back the bay to its original state, suggest is that technology is better at cleaning up industrial messes than it is at solving the intractable, or not easily controllable, problems. These problems are caused by millions of people who are going about their everyday activities.

As can be seen from a direct comparison of Examples 1 and 2 above. Example 1 is elaborated in the following ways. First, those lexical items in Example 1 (i.e., contend, ingenuity, day of reckoning, and intractable), determined to be difficult and thus selected as target words, are left intact, which might have been removed if text simplification had been applied as in typical commercially available EFL graded readers, and elaborated mainly by the use of paraphrasing and apposition. In terms of lexical elaboration, a non-target word (e.g., revive) is also elaborated in order not to give the participants in the study, particularly those in the intentional learning condition, an impression that only some words (i.e., target words) are systematically given their meanings, compared to others (i.e., non-target words) in the passage they read, and thus need to
be paid special attention. Second, in Example 2, some words which have been deleted through a process of ellipsis are provided (i.e., it is in a comparative construction and who are in a relative construction) to make syntactic relations between words explicit, although, as a result, the passage takes on some redundancy. Lastly, wh-cleft is used (i.e., What the long-standing efforts to revive the bay, ..., suggest) to make a thematic relation clear in Example 2. As a result, what is readily noticeable from Example 2, compared to Example 1, is that Example 2 is much longer and more redundant than Example 1. Despite its increased overall text length and redundancy, text elaboration applied to Example 2 is hypothesized to have made Example 1 more comprehensible by providing definitions of difficult words and making opaque syntactic or thematic structures transparent and, in doing so, increased the likelihood of its readers being able to learn vocabulary incidentally while their primary focus remains understanding its content.

To test this hypothesis, a study can be designed to investigate how text modification in the areas of discourse, syntax, semantics, and vocabulary, especially manipulating degrees of explicitness and richness of contextual information on unknown L2 lexical items, affects L2 comprehension and vocabulary learning through reading. Similar L1 studies have considered whether text is considerate or inconsiderate and how its (in)considerateness or reader-friendliness (unfriendliness) affects elementary school children's vocabulary learning (e.g., Gordon, Schumm, Coffland, & Doucette, 1992; Herman, et al., 1987; Konopak, 1988, 1989). In L2 research, only a few studies have empirically investigated this hypothesis, examining a relationship between text elaboration and vocabulary learning, whether incidental or intentional, (e.g., Hyunjoo Chung, 1995; Silva, 2000; Urano, 2000; Watanabe, 1997), although a different kind of text elaboration, glossing (i.e., inserting L1 or L2 vocabulary glosses into various positions in authentic text [e.g., Jacobs, 1991; Jacobs, Dufon, & Hong, 1994]) was used (e.g., Myong Hui Ko, 1995; Watanabe, 1997).

Previous L2 studies investigated the effects of text modification only on reading comprehension by experimentally manipulating various text characteristics, including context surrounding vocabulary to be learned, to examine optimal textual conditions for improving reading comprehension. Hyunjoo Chung (1995) and Watanabe (1997), however, have advanced one step beyond an investigation of the effects of text modification on reading comprehension (i.e., a relationship between input/interaction modification and L2 comprehension), to an inquiry into its effects on learning (i.e., a relationship between L2 comprehension and L2 learning).

3. Cross-fertilization between Research on Reading and Vocabulary

In L2 learning and teaching, there has been a fruitful cross-fertilization between research on reading and vocabulary, as has been shown especially in the perspectives taken in Huckin, et al.
(1993). Some studies have investigated L2 vocabulary learning through reading without an incidental learning research design (e.g., Duignan, 1990; Krantz, 1991) or without a focus on text modification (e.g., Day et al., 1991; Dupuy & Krashen, 1993; Hulstijn, 1992; Knight, 1994; Pitts et al., 1991). However, there have been only a few psycholinguistically motivated experimental studies on reading-based L2 vocabulary learning which have investigated both how L2 learners comprehend written input in unmodified or modified text and, while doing so, incidentally learn L2 vocabulary (e.g., Hyunjoo Chung, 1995; Watanabe, 1997). Given the great interest in input modification and its relationship to L2 comprehension and ultimately to L2 learning (cf. Long, 1996), it will be worthwhile to further investigate interactional effects of text elaboration and incidental vocabulary learning in a well-designed experimental study.

Thus, in the following study, a written context (i.e., reading), where comprehensible input via input elaboration is contained, is preferred to an oral context (i.e., listening) on the following two reasons. The first is that understanding L2 lexical items in reading, compared to listening, is expected to provide much more processing time for L2 learners in an experimental study than those in a listening comprehension study with a comparable design. In reading, words are fixed in time and L2 learners can make as many trips to experimental passages as they need as long as this option of going back to the passage is permitted and factors which can inadvertently affect the study can be controlled more easily and better than those in a study with spoken input. Secondly, vocabulary has been studied extensively in conjunction with reading comprehension research and has been a good candidate for such a learning-from-context measure (see e.g., Gordon, et al., 1992; Herman, et al., 1987; Konopak, 1988, 1989, for relevant L1 studies and Hyunjoo Chung, 1995; Watanabe, 1997, for relevant L2 studies). Thus, vocabulary is chosen as a measure of intentional and incidental L2 learning from elaborated written context.

4. Research Questions

1. Do adult EFL learners learn vocabulary through reading?
2. Does text elaboration facilitate EFL vocabulary learning?
3. Are there any differences in intentional and incidental EFL vocabulary learning from reading elaboratively modified or unmodified text?

III. METHOD

1. Participants

The participants were 211 adult Korean EFL learners, all freshmen drawn from the
Departments of Business Administration and English Language Education and the Colleges of Law and Medicine, at a university in Seoul, Korea. The participants sampled from their intact classes were randomly assigned to four treatment groups (n = 140) and two control groups (n = 71). Information on participants' EFL proficiency, based on the administration of a standardized test (e.g., the TOEFL), was not available, so their overall EFL proficiency was measured instead on a cloze test.¹

2. Materials

1) Passages for Treatment Groups²

The two treatment passages were based on an excerpt from an article, “More People, More Pollution,” from U.S. News & World Report (September 12, 1994). The unmodified NS baseline version was prepared by combining eight paragraphs from the article, which were taken from the first ten paragraphs in the original article, deleting the 5th and 6th paragraphs to make an excerpt containing fewer than 700 words. An NS of American English with teaching experience in EFL reading was asked to evaluate whether the excerpt followed logical development and sounded natural. The elaborated version was prepared from the unmodified NS baseline version on the basis of the suggestions for text elaboration in Parker and Chaudron (1987) and Yano et al. (1994). Example sentences in these two studies provided some practical ways to elaborate unmodified text. In creating the elaborated version, the unmodified NS baseline version was elaborated mainly in the area of lexis (i.e., vocabulary elaboration) (e.g., Chaudron, 1982; Toya, 1993), but attention was also paid to aspects of syntax, semantics, and discourse.

2) Target Words for Treatment Groups

The target words (TWs) in the treatment passages were selected in the following way. First, the researcher initially selected 37 words which could be perceived to be difficult and unfamiliar to college freshmen in the Korea EFL situation. Then, two other NSs of Korean with experience in teaching EFL in Korea were asked to circle words which, if they had been college freshmen, would have been felt to be difficult and unknown to them while reading the unmodified NS

¹ A univariate one-way ANOVA, performed on the cloze test scores from 211 participants, showed no significant differences in overall EFL proficiency among the six groups (F(5, 205) = .425, p = .8309).
² The readability of both treatment and control passages was analyzed using the Flesch-Kincaid formula. The result for the modified (i.e., elaborated), unmodified (e.g., unelaborated), and control passage, was 13.0, 12.0, and 12.2, respectively.
baseline version. On the basis of the pilot testing, 33 words were finally selected to be included in a revised version which was used in this study.

3) Passage for Control Groups

The control passage was an excerpt from an article, "Alaska’s Billion-Dollar Quandary," from *Time* (September 28, 1992). The first six paragraphs and the last paragraph in the article were combined to create the control passage. The same NS of American English was asked again to judge the natural transition of idea development in the control passage. As the document statistics show, the control passage was at nearly the same difficulty level as the unmodified NS baseline passage and possessed similar text characteristics to the unmodified treatment passage.

4) Target Words for Control Groups

The TWs in the control passage were selected in the same way as those in the treatment passages. A total of 37 words were initially chosen to be piloted in consultation with the same two NSs of Korean. On the basis of the pilot study, 33 TWs were finally chosen. Both the treatment and control passages were presented without the titles and summary-type one-sentence introductions that appeared in the original articles.

3. Instrumentation

1) Reading-related Measures

Two versions of an eight-item multiple-choice (MC) reading comprehension (RC) test were used as a device to fully engage the participants in a reading activity before an immediate vocabulary posttest, which was believed to approximate to a normal reading-for-meaning process. One RC test, based on the unmodified and elaborated passages, was for the four treatment groups and the other test, based on the control passage, was for the two control groups. Each item in the two types of RC tests consisted of one correct answer and three distractors. The instructions were given in Korean and the participants were asked to mark their answers on the answer sheets provided together with the RC tests. The reliability of the RC test for the treatment groups was very low, .55 by using Cronbach’s alpha. The purpose of administering the RC measure just before the immediate vocabulary posttest was to expose the participants to TWs in context while reading and not to sensitize the participants in the incidental learning condition, where no mention of any type of postreading vocabulary assessment was made, to paying their
attention to TWs intentionally. Thus, since the comprehension tests were actually part of the treatment condition, the results from the RC tests were not included in the main statistical analyses.

2) Vocabulary-related Measures

As measures of vocabulary learning through intentional and incidental exposure to TWs in context, two versions of supply-definition vocabulary posttests were used in this study. An immediate vocabulary posttest was administered right after the RC tests and a delayed vocabulary posttest was administered exactly one week later. These vocabulary posttests asked the participants to write the meaning of each target word in Korean in the blanks provided while TWs were presented out of context. In order to elicit any partial vocabulary knowledge from the participants, it was explicitly stated in the instructions in Korean that even when they were not sure of the word meaning, they were encouraged to guess the word meaning. Each target word's part of speech, as functioned in the treatment and control passages, was also provided in Korean right next to the target word in parentheses to give the participants clues to its meaning by showing how each target word in question behaved syntactically in the reading passages they had read. There were 66 items in both vocabulary posttests, 33 TWs each drawn from the treatment and control passages. The presentation order of 66 TWs was randomized in the immediate and delayed vocabulary posttests. As the participants were exposed to only the half of the total number of TWs \( (k = 66) \) in each posttest, it was expected that a comparison of the posttest scores of the treatment groups and the control groups on the TWs they were not exposed to in the treatment (i.e., passage reading) could function as a surrogate vocabulary pretest only when it could be shown that both treatment groups and control groups were similar in terms of their EFL proficiency. Thus, for example, when the treatment groups' performance on the 33 TWs from the treatment passages (i.e., the treatment groups were exposed to these TWs) and the control groups' performance on the same 33 TWs (i.e., the control groups were not exposed to these TWs) were compared, it could be shown how the treatment groups would have performed on a vocabulary pretest if it had been administered.

4. Procedures

1) Treatment

The participants were randomly assigned to six groups by random distribution of six different versions of the treatment and control passages (i.e., reading passages for control/incidental
(CON/INC), control/intentional (CON/INT), unmodified/incidental (UNM/INC), unmodified/intentional (UNM/INT), elaborated/incidental (ELA/INC), and elaborated/intentional (ELA/INT) groups) and the presence or absence of the written instructions in the reading passages to focus on words (i.e., intentional vs. incidental learning) while reading. All the participants, regardless of their group membership, were told that they were participating in a RC experiment and, as soon as finishing reading the passage, they would have a test on their comprehension of the passage they read by a MC test. They were also told that the reading passage would be collected after 20 minutes and that, because of this, they had to read the entire passage carefully. Only the participants in the intentional learning condition were told that there would be an additional vocabulary test based on the words they would come across in the passage they would read, so it was expected that their intention would be sensitized to focus on words while reading, whereas no mention of vocabulary was made at all to the participants in the incidental learning condition. In this respect, it could be said that the participants in the incidental learning condition were put into a pure incidental learning experiment. 20 minutes were given to passage reading and 8 minutes, answering an MC RC test.

2) Immediate Vocabulary Posttest

As soon as finishing the RC test, the participants were given an immediate unexpected or expected vocabulary posttest, determined by which learning condition each participant had been in the treatment. 20 minutes were allowed. No indication was given to the participants that another vocabulary test would be given later during the semester.

3) Delayed Vocabulary Posttest

One week after the immediate vocabulary posttest in the treatment session, another vocabulary posttest was administered. This delayed vocabulary posttest was designed to investigate whether there were any long-term effects of text elaboration on either intentional or incidental vocabulary learning.

5. Scoring

A four-point scoring scale ranging from 0 to 3 was used in this study, where 0 represented no response or an incorrect answer; 1, a distant partial answer; 2, a very close partial answer; and 3, a reasonably complete correct answer. The same scoring scale was applied to score both immediate and delayed vocabulary posttests. Two people including the researcher scored the
immediate vocabulary posttest scores from the 40 participants, which were randomly selected on the basis of the first 40 non-overlapping numbers out of 60 randomly generated numbers. As the interrater reliability in terms of percentage agreement was .93, half of the remaining data were scored by each.

6. Design

The design of this study was a 3 x 2 factorial multivariate analysis of variance (MANOVA) design with text modification type (three levels: unmodified, elaborated, control) and learning condition (two levels: intentional, incidental) as between-subjects factors. The independent variables in this study were (1) text modification type and (2) learning condition; the dependent variables were two supply-definition vocabulary posttests.

7. Analyses

A two-way MANOVA was performed only on the participants’ performance on the TWs from the treatment passages to investigate whether there would be main effects for the two independent variables (i.e., text modification type and learning condition). As there were significant main effects, two separate univariate ANOVAs of the two dependent variables were carried out. Then, a post-hoc comparison procedure, Scheffé’s test, was used to determine which pairs of means were significantly different. Additionally, the Tukey test was used to further compare the means of the treatment groups of research interest. The alpha level was set at .05 for all statistical tests.

V. RESULT

1. Vocabulary Posttests

Only the data from the participants’ performance on the 33 target words from the treatment passages were considered in analyzing the immediate and delayed vocabulary posttests. The reliabilities of these two 33-item vocabulary posttests were .7919 for the immediate vocabulary posttest and .7671 for the delayed vocabulary posttest by using Cronbach’s alpha. The means and standard deviations of all the participants were 27.924 and 12.254, respectively, for the immediate vocabulary posttest, and 27.100 and 12.050, respectively, for the delayed vocabulary posttest. (See Tables 1 and 2 for the means and standard deviations for the immediate and
delayed vocabulary posttests.)

### TABLE 1
**Means and Standard Deviations of Immediate Vocabulary Posttest Scores**

<table>
<thead>
<tr>
<th>Conditions</th>
<th>n</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control/Incidental</td>
<td>37</td>
<td>21.000</td>
<td>9.321</td>
</tr>
<tr>
<td>Control/Intentional</td>
<td>34</td>
<td>24.647</td>
<td>11.703</td>
</tr>
<tr>
<td>Unmodified/Incidental</td>
<td>35</td>
<td>26.029</td>
<td>10.150</td>
</tr>
<tr>
<td>Unmodified/Intentional</td>
<td>36</td>
<td>31.972</td>
<td>13.289</td>
</tr>
<tr>
<td>Elaborated/Incidental</td>
<td>34</td>
<td>30.235</td>
<td>12.780</td>
</tr>
<tr>
<td>Elaborated/Intentional</td>
<td>35</td>
<td>33.914</td>
<td>11.528</td>
</tr>
</tbody>
</table>

### TABLE 2
**Means and Standard Deviations of Delayed Vocabulary Posttest Scores**

<table>
<thead>
<tr>
<th>Conditions</th>
<th>n</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control/Incidental</td>
<td>37</td>
<td>22.486</td>
<td>9.591</td>
</tr>
<tr>
<td>Control/Intentional</td>
<td>34</td>
<td>24.294</td>
<td>10.939</td>
</tr>
<tr>
<td>Unmodified/Incidental</td>
<td>35</td>
<td>25.914</td>
<td>11.379</td>
</tr>
<tr>
<td>Unmodified/Intentional</td>
<td>36</td>
<td>29.333</td>
<td>14.004</td>
</tr>
<tr>
<td>Elaborated/Incidental</td>
<td>34</td>
<td>28.324</td>
<td>13.232</td>
</tr>
<tr>
<td>Elaborated/Intentional</td>
<td>35</td>
<td>32.400</td>
<td>10.669</td>
</tr>
</tbody>
</table>

A two-way MANOVA was used in analyzing the data from the participants’ performance on the immediate and delayed vocabulary posttests. The results of the MANOVA tests of statistical significance indicate that the main effects of text modification type and learning condition were found to be significant ($F_s = 5.716$ and $3.861$; $p < .05$, respectively) by the Wilks’ Lambda Criterion, while there was no interaction between text modification type and learning condition ($F = .322$, $p < .86$).

Due to the significant MANOVA tests for main effects, it was appropriate to examine the separate univariate analyses of the two dependent variables (Hummel & Sligo, 1971). Tables 3 and 4 display the results of applying univariate $F$ tests to each of the dependent variables.

### TABLE 3
**Two-Way Univariate ANOVA Table for Text Modification Type, Learning Condition, and Interaction of Text Modification Type and Learning Condition on Immediate Vocabulary Posttest**

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Text Modification Type (A)</td>
<td>2</td>
<td>3114.547</td>
<td>1557.274</td>
<td>11.719</td>
<td>.0001*</td>
</tr>
<tr>
<td>Learning Condition (B)</td>
<td>1</td>
<td>1031.116</td>
<td>1031.116</td>
<td>7.760</td>
<td>.0058*</td>
</tr>
<tr>
<td>A x B</td>
<td>2</td>
<td>61.248</td>
<td>30.624</td>
<td>.230</td>
<td>.7944</td>
</tr>
<tr>
<td>Error</td>
<td>205</td>
<td>27240.569</td>
<td>132.881</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*p < .05
TABLE 4
Two-Way Univariate ANOVA Table for Text Modification Type, Learning Condition, and Interaction of Text Modification Type and Learning Condition on Delayed Vocabulary Posttest

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Text Modification Type (A)</td>
<td>2</td>
<td>1728.766</td>
<td>864.383</td>
<td>6.289</td>
<td>.0022*</td>
</tr>
<tr>
<td>Learning Condition (B)</td>
<td>1</td>
<td>506.810</td>
<td>506.810</td>
<td>3.688</td>
<td>.0562</td>
</tr>
<tr>
<td>A x B</td>
<td>2</td>
<td>47.841</td>
<td>23.921</td>
<td>.174</td>
<td>.8404</td>
</tr>
<tr>
<td>Error</td>
<td>205</td>
<td>28174.886</td>
<td>137438</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*p < .05.

As Table 3 indicates, the results of the two-way univariate ANOVA on the immediate vocabulary posttest show significant main effects for text modification type ($F = 11.719$, df = 2/205, $p < .05$) and learning condition ($F = 7.76$, df = 1/205, $p < .05$), whereas the interaction of text modification type and learning condition was not significant ($F = .23$, df = 2/205, $p < .7944$). The results of the same analysis on the delayed vocabulary posttest in Table 4 also reveal that there were a significant main effect only for text modification type ($F = 6.289$, df = 2/205, $p < .05$) and a nonsignificant main effect for the interaction of text modification type and learning condition ($F = .174$, df = 2/205, $p < .8404$), with the factor, learning condition, approaching significance ($F = 3.688$, df = 1/205, $p < .0562$).

Where significant $F$ ratios were obtained, the most conservative multiple comparison procedure, Scheffé's test, was used to identify which means were significantly different from one another among the levels of the two independent variables in the immediate and delayed vocabulary posttests. (See Tables 5, 6, 7, and 8).

TABLE 5
Results of Scheffé's Test for Text Modification Type on Immediate Vocabulary Posttest

<table>
<thead>
<tr>
<th>Mean Difference</th>
<th>Critical Difference</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>CON vs. UNM</td>
<td>6.296</td>
<td>4.771</td>
</tr>
<tr>
<td>CON vs. ELA</td>
<td>9.355</td>
<td>4.805</td>
</tr>
<tr>
<td>UNM vs. ELA</td>
<td>3.059</td>
<td>4.805</td>
</tr>
</tbody>
</table>

*p < .05.

TABLE 6
Results of Scheffé's Test for Text Modification Type on Delayed Vocabulary Posttest

<table>
<thead>
<tr>
<th>Mean Difference</th>
<th>Critical Difference</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>CON vs. UNM</td>
<td>4.296</td>
<td>4.852</td>
</tr>
<tr>
<td>CON vs. ELA</td>
<td>7.039</td>
<td>4.887</td>
</tr>
<tr>
<td>UNM vs. ELA</td>
<td>2.743</td>
<td>4.887</td>
</tr>
</tbody>
</table>

*p < .05.
An examination of the means in Tables 5 and 9 indicates that the mean for the elaborated group was greater than the mean for the unmodified group and the latter was greater than the mean for the control group. The differences between the control group and the unmodified group, and the control group and the elaborated group, were large enough to be significant at the $p < .05$ level. The difference between the unmodified group and the elaborated group was not significant. In the delayed vocabulary posttest, only the difference between the control group and the elaborated group was significant ($p < .05$) and no other differences were significant as shown in Tables 6 and 10.

Tables 7 and 11 reveal that the mean difference between the incidental group and the intentional group in the immediate vocabulary posttest was significant at the $p < .05$ level. The same result was obtained in the delayed vocabulary posttest as Tables 8 and 12 indicate.
TABLE 11
Means Table for Different Levels of Learning Condition on Immediate Vocabulary Posttest

<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>INC</td>
<td>106</td>
<td>25.623</td>
<td>11.356</td>
</tr>
<tr>
<td>INT</td>
<td>105</td>
<td>30.248</td>
<td>12.734</td>
</tr>
</tbody>
</table>

TABLE 12
Means Table for Different Levels of Learning Condition on Delayed Vocabulary Posttest

<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>INC</td>
<td>106</td>
<td>25.491</td>
<td>11.590</td>
</tr>
<tr>
<td>INT</td>
<td>105</td>
<td>28.724</td>
<td>12.340</td>
</tr>
</tbody>
</table>

Additionally, the Tukey test was applied to the data from the immediate and delayed vocabulary posttests to examine whether there would be significant differences in the mean comparisons of the two groups where learning conditions had been controlled for (i.e., the UNM/INC group versus the ELA/INC group and the UNM/INT group versus the ELA/INT group in both immediate and delayed vocabulary posttests) and those where text modification types had been controlled for (i.e., the UNM/INT group versus the UNM/INC group and the ELA/INT group versus the ELA/INC group in both immediate and delayed vocabulary posttests). As Table 13 reveals, no mean differences between the treatment groups in the immediate vocabulary posttest (i.e., ELA/INC minus UNM/INC, ELA/INT minus UNM/INT, ELA/INT minus ELA/INC, and UNM/INT minus UNM/INC) exceeded the critical values for the Tukey test and thus found to be not significant at the \( p < .05 \) level. The same pairs of mean comparisons in the delayed vocabulary posttest were not significant, either, at the \( p < .05 \) level (see Table 14).

TABLE 13
Results of Tukey Test for Group on Immediate Vocabulary Posttest

<table>
<thead>
<tr>
<th></th>
<th>Mean Difference</th>
<th>Critical Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>UNM/INC vs. ELA/INC</td>
<td>4.207</td>
<td>8.43</td>
</tr>
<tr>
<td>UNM/INC vs. UNM/INT</td>
<td>5.944</td>
<td>8.32</td>
</tr>
<tr>
<td>ELA/INT vs. ELA/INC</td>
<td>3.679</td>
<td>8.43</td>
</tr>
<tr>
<td>ELA/INT vs. UNM/INT</td>
<td>1.942</td>
<td>8.31</td>
</tr>
</tbody>
</table>

TABLE 14
Results of Tukey Test for Group on Delayed Vocabulary Posttest

<table>
<thead>
<tr>
<th></th>
<th>Mean Difference</th>
<th>Critical Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>UNM/INC vs. ELA/INC</td>
<td>2.409</td>
<td>8.56</td>
</tr>
<tr>
<td>UNM/INC vs. UNM/INT</td>
<td>3.419</td>
<td>8.44</td>
</tr>
<tr>
<td>ELA/INT vs. ELA/INC</td>
<td>4.076</td>
<td>8.57</td>
</tr>
<tr>
<td>ELA/INT vs. UNM/INT</td>
<td>3.067</td>
<td>8.45</td>
</tr>
</tbody>
</table>
VI. DISCUSSION

1. Incidental Vocabulary Learning from Context

The first specific research question examines whether the absence or presence of exposure to written input containing unknown target lexical items in the form of reading affects EFL vocabulary learning. The participants in the four treatment groups had only a single exposure to the TWs in the treatment passages, whether unmodified or elaborated, they had read. Taking this fact into consideration, it was expected, as in the case of the previous studies on L1 and L2 incidental vocabulary learning from reading, that a single exposure to new words would lead to at least some partial learning of some of those words, but the strength of such an exposure and learning relationship would be quite weak, but significant.

As the results of Scheffé’s test in Table 5 and the mean scores in Table 9 indicate, the participants in the two control groups ($M = 22.746, SD = 10.612$) who had not been exposed to the TWs from the treatment passages during the treatment session learned a significantly smaller number of those TWs than the treatment groups (i.e., the unmodified groups and the elaborated groups) ($Ms = 29.042$ and $32.101$; $SDs = 12.137$ and $12.213$, respectively) with such exposure, on the immediate vocabulary posttest. This result shows that EFL learners can learn the meaning of unknown or previously partially known words incidentally from reading although the amount of incidental vocabulary learning from reading is not very big. For instance, the two treatment groups, regardless of the text modification type, under the incidental learning condition (i.e., the UNM/INC groups and the ELA/INC groups) learned about 28% (i.e., an average of nine words) of the 33 TWs from the treatment passages (see Table 1).

However, the fact that the control groups were able to score about 23% of the 33 TWs implies that some of the TWs in the treatment passages might have already been known to these groups. The similar reasoning can be made as to the interpretation of the treatment groups’ performance on the TWs. Although there was significant vocabulary learning from context, not every target word could be safely said to have been encountered for the first time by the participants in the treatment groups. Thus, a strong claim can not be made that all the learning of the target lexical items was only possible from exposure to the treatment passages. A pretest-posttest control group design with an appropriately long interval between administering a pretest and a posttest might have revealed how many words were already known to the participants in the study.

2. Text Elaboration and EFL Vocabulary Learning

The second research question is concerned with whether an elaborated text will lead to better
vocabulary learning than an unmodified text. As the results of Scheffé's test in Table 5 indicate, there was no statistically significant difference between the unmodified groups and the elaborated groups in vocabulary learning although there was a mean score difference of 3.059, showing the elaborated groups' better performance. In this study, the elaborated version was modified from a baseline NS text which did not contain many helpful contextual clues to difficult words in its passage, which provide its readers with information which can facilitate successful contextual guessing of those difficult words. However, the participants in the elaborated groups were asked to read the elaborated passages which contained about twice as many words as those in the unmodified passages in the same 20 minutes although the two versions were evaluated to be at similar difficulty. Thus, although the participants in the elaborated groups were given contextual help in the form of text elaboration, they did not seem to have been able to make the fullest use of this advantage due to the big difference in passage length. This seems to have produced a nonsignificant finding concerning the second research question.

3. Intentional and Incidental Vocabulary Learning from Elaborated Text

The third specific research question asks which vocabulary learning condition, incidental or intentional, is more facilitative of EFL vocabulary learning from reading an elaborated text. As Table 1 shows, the participants reading the elaborated version under the intentional learning condition \((M = 33.914, SD = 11.528)\) performed better on the immediate vocabulary posttest than those reading the same version under the incidental learning condition \((M = 30.235, SD = 12.780)\). This result also obtains in the delayed vocabulary posttest (ELA/INT, \(M = 32.400, SD = 10.669\) vs. ELA/INC, \(M = 28.324, SD = 13.232\)) as indicated in Table 2. However, the mean differences between the two groups on both immediate and delayed vocabulary posttests were not statistically significant as Tables 13 and 14 show.

4. Text Elaboration and Vocabulary Retention

Additionally, in order to investigate whether facilitative effects of text elaboration on vocabulary learning immediately after the treatment will also be retained one week later (i.e., a question of vocabulary retention), the mean scores of the unmodified and the elaborated groups on both immediate and delayed vocabulary posttests were compared. For the elaborated groups, there was a mean score drop of 1.719 from the immediate posttest to the delayed posttest, whereas for the unmodified groups, there was a mean score drop of 1.394 from the immediate posttest to the delayed posttest. Given the fact that after the treatment, both the unmodified and
elaborated groups were not exposed again to the TWs in a meaningful context, a small drop in the mean scores can be interpreted as there was a positive effect of text elaboration on vocabulary retention. This comparison was also intended as a check of whether one group's significant performance over the other is purely due to a recency effect.

VII. CONCLUSION

The present study investigated whether text elaboration would facilitate intentional and incidental EFL vocabulary learning from reading with 211 Korean college freshmen in an EFL classroom instructional setting, using a 3 x 2 factorial MANOVA design. The main research findings seem to suggest that: (1) incidental vocabulary learning through reading was possible with these learners although the amount of such learning was small, but significant; (2) elaborated text better facilitated vocabulary learning from reading than unmodified text although the difference between the two text types was not statistically significant; and (3) intentional vocabulary learning from reading was better than incidental vocabulary learning, regardless of the text modification type, on both immediate and delayed vocabulary posttests.

However, these research findings have to be interpreted with caution due to the limitations of the present study. First, although there were significant differences in vocabulary learning between the control groups and both treatment groups on the immediate vocabulary posttest, and between the control groups and the elaborated groups on the delayed vocabulary posttest, the control groups without the treatment managed to provide the meanings of some target words from the treatment passages they had not been exposed to. This can be interpreted as they might have exhibited their previous knowledge of these words. A posttest only control group design without a pretest adopted in this study paid a careful attention not to provide the participants in the control groups any slightest information as to the words targeted in the treatment. The fact that the control groups managed to score 23% of the target words without exposure to these words made it less powerful to draw a strong support for the incidental-vocabulary-learning-from-reading hypothesis. There can be ways to deal with this problem. For example, a pretest-posttest control group design with a relatively long interval between the pretest and the treatment, using a supply-definition measure, which does not provide information as to the meanings of the target words, can be employed in future research investigating a similar research question. Additionally, in terms of selecting the target words in the experiment, objective word frequency information provided in EFL learners' dictionaries can be consulted along with experienced EFL teachers' subjective assessment of word difficulty in order to select the target words participants have the least familiarity with, which will help interpret reasonably
any net learning resulting from incidental exposure to EFL vocabulary through reading.

Second, as discussed in the Materials section, the reading passages chosen for this study were more difficult and longer in terms of readability than those used in the previous studies. Coupled with these inherent passage difficulty and length factors, the time constraint imposed on implementing this study (i.e., only two 50-minute sessions were available) could have put the elaborated groups at a disadvantage as they were asked to read about twice as much as the other groups in the same 20 minutes. This time constraint factor could have produced some nonsignificant results, overriding the advantage of text elaboration for an expected significant increase in EFL vocabulary learning. In this respect, shorter passages and more time to read can be given to a group reading elaborated text in future studies.

Third, the previous text modification studies have focused on whether modified text can lead to better listening or reading comprehension and have not gone, with the exception of Hyunjoo Chung (1995) and Watanabe (1997), as far as to investigating whether an advantage of text modification for comprehension can be also shown in learning. The present study only dealt with the learning issue due to the time constraint factor mentioned above and its concomitant constraint on the number of items on a reading comprehension test. Thus, it was not easy to produce a highly reliable measure of reading comprehension to test whether text elaboration enhances both reading comprehension and vocabulary learning, which is a killing-two-birds-with-one-stone-advantage of text elaboration for comprehension and learning. With the time factor controlled for, a reliable reading comprehension test with a reasonably enough number of items can be used in future research.

Despite these limitations, the present study provided some positive support to: (1) incidental vocabulary learning through reading is possible and can be enhanced with text elaboration; (2) intentional vocabulary learning is superior to incidental vocabulary learning; and (3) although not directly addressed in the present study, extensive EFL reading, compared to intensive EFL reading, can facilitate both vocabulary learning and reading comprehension skills.

The suggestions for the future studies in this line of research are that: (1) there is a need for longitudinal studies into long-term effects of text elaboration on EFL vocabulary learning and reading comprehension in an extensive EFL reading program, which will ask EFL learners to read a set of elaborated reading materials over an extended period of time, allowing them multiple exposures to a group of targeted lexical items to be learned in meaningful and various contexts; (2) there is a need for process-oriented single-subject case studies into multifaceted processes involved in incidental learning of EFL vocabulary from reading an elaborated text, which, with the help of think-aloud research methodology including both introspection and retrospection (e.g., Fecht & Kasper, 1987), will shed light on how L2 lexis is initially learned and retained (see, e.g., a special issue of Second Language Research (Vol. 11, No. 2, June 1995)
on single-subject research on lexical acquisition edited by Paul Meara); and (3) in addition to L2 vocabulary learning, incidental learning of other areas of L2 (e.g., grammar and discourse) from reading an elaborated text need to be investigated.

REFERENCES


Applicable levels: tertiary education, adult education
Key words: vocabulary acquisition, reading, incidental and intentional learning, text elaboration

Youngkyu Kim
Department of Second Language Studies
University of Hawai‘i at Mānoa
1890 East-West Road
Honolulu, HI 96822, USA
Tel: 1-808-956-2785
Fax: 1-808-956-2802
Email: youngkyu@hawaii.edu

Received in September 2003
Reviewed in October 2003
Revised version received in February 2004