Modern English Education, Vol. 18, No. 3, Fall 2017
http://dx.doi.org/10.18095/meeso.2017.18.3.03

The Relationship Between Gloss Type and L2 Proficiency in Incidental Vocabulary Learning*

Myong-Hee Ko
Soonchunhyang University

Ko, Myong-Hee. (2017). The relationship between gloss type and L2 proficiency in incidental vocabulary learning. *Modern English Education, 18*(3), 47-69.

This study investigates the relationship between L2 proficiency and different types of glosses. Just under 330 university students in Korea were divided into higher and lower levels of English fluency, and were further divided into four gloss conditions (no-gloss, L1 gloss, L2 gloss, and L1+L2 gloss). They read passages, completed an immediate vocabulary test and a survey, and took a delayed vocabulary test after four weeks. According to the findings, the L1 gloss and the L1+L2 gloss were effective for lower-level learners, and the L2 gloss and the L1+L2 gloss were effective for higher-level learners. However, on the survey, the L1+L2 gloss type was the most preferred gloss type of the participants. Considering the statistical analysis of the interaction effect between L2 proficiency and gloss type together with the results of the student opinion survey, L1+L2 glosses are likely the best gloss type for reading materials at any proficiency level.

**[vocabulary acquisition/gloss/incidental learning/
어휘 습득/어휘 주석/우연적 학습]**

I. INTRODUCTION

The role of vocabulary in second/foreign language (L2) learning is very important since a word is the fundamental component of any language. The misuse of a word can cause breakdowns in communication since words carry speakers' or writers' intended messages (Gass, 1988; Nation, 1982; Thornbury, 2002). Nation (2013) indicates that if a L2 reader

* This work was supported by the National Research Foundation of Korea Grant funded by the Korean Government (NRF—2014S1A5A8019034).
This work was supported by the Soonchunhyang University Research Fund.

www.kci.go.kr

knows around 9,000 word families, s/he may know close to 98% of the words in most texts and be able to read without a dictionary. Acquiring 9,000 words, the coverage of high- and mid-frequency vocabulary, is not easy for L2 learners. Accordingly, the search for an effective way to teach or learn a word is a crucial one for educators seeking ways to improve learners' overall L2 competence.

Along with explicit teaching of L2 vocabulary, incidental vocabulary learning while reading has been investigated, and it has had a positive impact on SLA literature (Cheng & Good, 2009; Gan, 2014; H. Sung, 2016; Yanguas, 2009; Yoshii, 2006). There has been concern, however, that L2 learners may experience difficulty in guessing the meanings of unknown words in texts due to limited contextual clues and learners' lack of L2 competence. As a result, researchers introduced glossing as a way of coping with these constraints (Coady, 1993; Huang & Liou, 2007; Stein, 1993; Thornbury, 2002).

A *gloss*, as used in L2 learning, is a definition or synonym of a key term in a text, which is given, usually at the bottom of a page, to help L2 readers decode an unknown or insufficiently understood word while reading a particular page of text (Bowles, 2004; Lomicka, 1998; Nation, 2001). In the case of online reading, glossing extends to pictures, video, and sound (Lomicka, 1998).

The studies of glossing, in connection with L2 vocabulary learning, were built around diverse research purposes and designs. The effect of a combination of first language (L1) and second language (L2) glosses (L1+L2 gloss), however, has rarely been examined. Moreover there is not enough research regarding the relationship between gloss type and L2 proficiency.

Therefore, the present study aims to reveal the interaction between L2 proficiency and various gloss types (no gloss, L1 gloss, L2 gloss, L1+L2 gloss) in L2 vocabulary learning in order to have a precise understanding of how L2 proficiency affects a learner's response to one or another gloss type. The findings of the present study offer useful information for instruction and materials development in L2 vocabulary education.

II. LITERATURE REVIEW

1. Three Models of Lexical Presentation

Gloss comparison studies in L1 and L2 may be explained adequately in line with recent studies on the lexical information process of bilinguals with different L2 proficiency (Kroll & Sunderman, 2003). According to Potter, So, Eckardt and Feldman (1984), there are three ways to process L2 words, depending on learner L2 proficiency: (a) the word association model; (b) the intermediate model; (c) the concept mediation model. Follow-up studies

www.kci.go.kr

(e.g., Chen & Leung, 1989; Kroll & Stewart, 1994; Kroll & Sunderman, 2003; Kroll, Van Hell, Tokowicz, & Green, 2010) also supported the hypothesis that L2 learners' lexical representation lies on the continuum between the activation of L1 and concepts of L2 when encountering L2 stimulus.

The word association model views that L2 words are closely associated with the corresponding L1 equivalents. This model suggests that there are word-to-word mappings that link L1 and L2 for lexical processing. When beginners are given L2 stimulus, they activate the L1 word association model and search for the corresponding L1 word in their memory repertoire. In other words, beginning language learners use L1 words as a medium to process L2 words.

As learner proficiency improves, learners move into the second stage, *the intermediate model*, which is the middle stage of lexical processing for L2 learners. L2 learners still operate L2 lexical processing using their L1; however, they gradually increase independent performance, a direct link between the L2 word and its concept. In other words, this model includes both the word-to-word link and word-to-concept link. At this stage, less skilled L2 learners may experience two types of lexical processing. While encountering L2 stimulus, they may experience unintended L1 activation and operation of the L2 conception.

As L2 learner proficiency further develops, learners gradually enter the third stage, *the concept mediation model*, where they are able to process a semantic concept of an L2 word when facing L2 stimulus. At this stage, learners are able to operate the L2 words independently without relating them to L1 equivalents. In other words, when given L2 stimulus, proficient L2 learners use the L2 concept link that is directly connected with the L2 word.

In short, the three models of lexical presentation depict the role of L1 as learners develop their L2 proficiency. At the early stage of learning, the lexical link between L1 and L2 is crucial. However, as their L2 proficiency improves, the direct link between an L2 word and an L2 semantic concept becomes indispensable.

2. Gloss Studies on L2 Vocabulary Learning

Incidental vocabulary learning refers to acquiring unknown words as a by-product of another activity. According to Krashen (1989), native speakers learn most words through repeated encounters while engaging in listening, speaking, reading, and writing. He also claimed that L2 learners are able to acquire a great number of words incidentally through extensive reading. However, some researchers (Huang & Liou, 2007; Laufer, 1996; Nation, 2001; Thornbury, 2002) have indicated concern about the effectiveness of L2 learners' incidental vocabulary learning with different language proficiency and educational backgrounds. For example, L2 learners may be constrained when retrieving the meaning of unknown words in a text due to the lack of their L2 ability. Besides, if the learner lives in

www.kci.go.kr

an English as a Foreign Language (EFL) context, s/he may have limited exposure to the target words. Reading can be a good avenue for L2 vocabulary development; however, it certainly has shortcomings for L2 readers. Incorporating glosses in reading texts is one way of coping with these shortcomings for L2 learners.

Researchers have employed various experimental designs to study the effects of gloss in L2 vocabulary learning. Several analyzed the effectiveness of L1 glosses (Bowles, 2004; Cheng & Good, 2009; Gettys, Imhof, & Kautz, 2001; Hulstijn, Hollander, & Greidanus, 1996). Researchers compared single gloss and multiple-choice (MC) gloss (Gan, 2014; Nagata, 1999; Rott, Williams, & Cameron, 2002; Watanabe, 1997). Several studies investigated textual gloss together with pictorial gloss, or various types of pictorial glosses in a multimedia environment (Moazzeni, Bagheri, Sadighi, & Zamanian, 2015; Rouhi & Mohebbi, 2012; Yanguas, 2009).

A few studies also compared L1 and L2 glosses through paper-based texts and a multimedia environment. Regarding gloss comparison studies in L1 and L2, Jacobs, Dufon and Hong (1994), Yoshii (2006), and M. H. Ko (2012) compared three gloss conditions (no gloss, L1 gloss, L2 gloss) using students in a Spanish course at an American university, Japanese university students, and Korean university students respectively. Jacobs et al. (1994) and M. H. Ko (2012) showed that a significant difference between the no-gloss and glossed conditions, but not between the L1 and L2 gloss conditions on the immediate vocabulary test. However, they found different findings on the delayed vocabulary test. In the case of Yoshii (2006), the definition-supply test showed no significant difference between L1 and L2 gloss conditions on the immediate and delayed vocabulary tests; however, a significant effect was found between the L1 and L2 gloss conditions on the multiple choice (MC) test.

The previous studies of Jacobs et al. (1994), Yoshii (2006), and M. H. Ko (2012) still leave researchers unclear regarding the effect of gloss type on L2 proficiency. One major reason may be due to a lack of consistent proficiency measurement devices across studies. When determining participant proficiency, Jacobs et al. (1994) assumed participants' level based on the indicated level of their L2 course, Yoshii (2006) assumed the level of his participants by guessing, and M. H. Ko's (2012) measurement was based on a cloze test. Consequently, these studies have reported mixed results.

Laufer and Hill (2000) were interested in the relationship between participants' online gloss look-up patterns and vocabulary learning in a multimedia environment. They compared three types of gloss (L1 gloss, L2 gloss, L1/L2 with other information), and also they analyzed participants' selection of both L1 and L2 glosses while reading online. There were two different groups of participants: 40 Hong Kong students with 570 on the paper-based TOEFL, and 32 Israeli students who scored above the mean on their university entrance exam. All the participants read a 120-word passage with 12 target words, and took an immediate vocabulary

www.kci.go.kr

test that required writing the meaning of these words, either in L1 or L2. For the Israeli students, there was no significant difference among the four look-up conditions. They used the L1 gloss most often, but they learned the most words when they looked up both L1 and L2 glosses together. The Hong Kong students, on the other hand, performed significantly lower with the L1 gloss option. Their most frequently used gloss type was the L1/L2 gloss with other information, but they acquired the most words when they looked up the L2 gloss.

Recently Hu, Vongpumivitch, Chang, and Liou (2014) examined the effects of L1 glosses and L2 glosses with L2 proficiency in a multimedia environment. Seventy-eight junior high school students in Taiwan were divided into high and low levels using an intra-school English test, and read two online passages with either L1 glosses or L2 glosses through two experiments with one week interval. In other words, they were assigned to four gloss treatments: (a) higher-level having L1 gloss before L2 gloss; (b) higher-level having L2 gloss before L1 gloss; (c) lower-level having L1 gloss before L2 gloss; (d) lower-level having L2 gloss before L1 gloss. After reading each passage, they were given an immediate vocabulary test, and delayed vocabulary test after two weeks, consisting of a definition-supply and a cloze section. In the case of the cloze test, the two higher-level groups performed significantly better than the two lower-level groups in both immediate and delayed vocabulary tests. In the case of the immediate definition-supply vocabulary test, both in the first and second experiment, two higher-level groups performed significantly better than the lower-level groups. In the case of the delayed definition-supply vocabulary test, in the first experiment, the higher-level receiving L2 glosses before L1 glosses performed significantly better than the two lower-level groups, and, in the second experiment, the higher-level receiving L1 glosses before L2 glosses performed significantly better than the two lower-level groups. In short, those who encountered L2 glosses in the text were able to maintain significantly more words than the two lower-level groups. It means that, in the case of higher-level students, there was a gloss-language effect on vocabulary retention.

3. Rationale of the Study

Overall, the effect of the L1+L2 gloss type has hardly been investigated in previous gloss studies. Besides, the relationship between L2 proficiency and gloss type has rarely been examined except for Hu et al. (2014) and Laufer and Hill (2000), but neither of them considered the effect of L1+L2 gloss in connection with L2 proficiency. Therefore more empirical evidence is needed in order to draw clearer conclusions regarding the effect of L1+L2 glosses including single L1 glosses and L2 glosses, in relation with L2 proficiency on incidental vocabulary learning.

In light of this need, the present study attempts to reveal the relationship between L2 proficiency and gloss type, using a between-subject design. To accomplish this goal, first,

www.kci.go.kr

L1+L2 glosses are included in the gloss type, along with L1 glosses and L2 glosses. Second, L2 proficiency is clearly incorporated as an independent variable into the present study's design. Finally, a learner opinion survey is used in the present study to find out participants' preferences regarding glossing and gloss type. Learner opinion is important information, along with the statistical data, because learners' perspectives reflect their attitudes or values, which have an important impact on L2 vocabulary learning. Along with the results of the present study's statistical analysis, student reports on gloss preference may lead to better conclusions on gloss comparison studies in relation to L2 proficiency.

The present study has a 2 × 4 factorial design; it has two proficiency levels (higher and lower) and four gloss type (no gloss, L1 gloss, L2 gloss, and L1+L2 gloss). There are four research questions:

1) Are there any interaction effects between L2 proficiency and gloss type on an immediate vocabulary test?
2) Are there any interaction effects between L2 proficiency and gloss type on a delayed vocabulary test?
3) Is there a time effect in vocabulary scores between an immediate and a delayed vocabulary tests?
4) What are learner opinions regarding different gloss types?

III. METHOD

1. Participants

In total 329 undergraduate and graduate students in Korea were divided into two levels of English ability and then each group was further divided into four gloss conditions. One hundred sixty-six undergraduate students were assigned to a higher proficiency level. According to Watanabe's (1997) cloze test, their English proficiency was determined to be high-intermediate; the mean of the cloze test was 21 out of 25. The four groups of higher-level participants were equal based on one-way ANOVA ($p > 0.05$) analysis. Their gender, major, and grade level varied, but all were between 18 and 28 years old, and enrolled in college English reading courses.

One hundred sixty-three graduate students were assigned to a lower proficiency level. Their mean on Watanabe's (1997) cloze test was around 8.5 out of 25, and their English proficiency was determined to be low. The four groups of lower-level participants were equal based on one-way ANOVA ($p > 0.05$) analysis. They belonged to a special graduate school, designed for students who have full-time jobs. Most of the graduate students had bachelor's degrees from

www.kci.go.kr

much less competitive universities than the four groups of higher-proficiency undergraduate students. The graduate students were mixed in terms of gender, major, and age, and enrolled in English reading courses to fulfill a requirement for graduation.

2. Instruments

1) Cloze Test

Watanabe's (1997) cloze test was used, as in M. H. Ko (2012), to check the English proficiency level of the participants (Oller, 1972). This cloze test originally consisted of 277 words; in the present study 25 words were randomly omitted. According to the Flesch-Kincaid readability scale, the text belonged to a 9th-grade reading level.

2) Reading Texts

An adapted reading passage from *The Multicultural Workshop* (Blanton & Lee, 1995) was used from M. H. Ko (2012). The title of the reading was "The Struggle to Be an All-American Girl." It is a story of a Chinese-American girl who struggles with her cultural identity. This topic was chosen because it was judged as an unfamiliar one for the participants. The text consisted of 602 words and belonged to an 8th-grade level, according to the Flesch-Kincaid readability scale. The reading text was adapted into four different versions since four different gloss types were being studied: no-gloss, L1 gloss, L2 gloss, and L1+L2 gloss. The text with no-gloss had no aid; it was the same as the original text. For the glossed texts, glosses appeared at the bottom of the page, but the selected vocabulary was not boldfaced in the passage.

In total, 14 glossed words were used, and these consisted of adjectives, nouns, an adverb, and a verb: *dissuade, stern, maniacal, crisp, flanked, painstakingly, blotch, nagging, vendor, raunchy, lilting, pedestrian, frenzied, and gibberish*. The L1 and L2 gloss were designed to be succinct explanations of the target words. The most appropriate meaning of the target word was given based on the term's usage in the reading passage. In the case of a L1+L2 gloss, both a L1 gloss and a L2 gloss were provided simultaneously (e.g., flanked: placed at the side/측면에 위치한). All of the final 14 target words appeared once in the glossed texts. There were originally 16 target words in M. H. Ko (2012) but, after piloting, two items were removed. According to M. H. Ko (2012), those target words were selected based on marks students had made in their texts when they encountered unknown words while reading. Those students who had similar backgrounds were instructed to underline any unknown words when reading the given text. They were the items that would cause difficulty for reading comprehension since they were unknown to 60% or more of the students.

www.kci.go.kr

3) Reading Comprehension Test

A reading comprehension test was given to meet their expectations, as indicated in the procedure section. It asked three things: (a) indicating the degree of understanding content, in percent; (b) writing main ideas of the story; and (c) writing details of the story.

4) Immediate Vocabulary Test and Delayed Vocabulary Test

The immediate vocabulary test was a MC test. The MC test format was thought to be appropriate to measure receptive knowledge of a word since, in the present study, the students acquired vocabulary incidentally after reading a text. The vocabulary test was revised several times as a result of pilot testing. Each of the test's 14 items had four MC options: one answer and three distractors. The following is an example.

e.g.) flanked:

a. placed at the back b. placed at the front

c. placed at the side d. placed at the center

Two versions of the test were used, to minimize favoring a particular language gloss. Combining the Korean and English versions, the total number of items in the MC test was 28 (14 items x 2 versions). In order to minimize the practice effect of these two versions of the test, vocabulary items and distractors were rearranged randomly.

For the delayed vocabulary test, items and distractors were also rearranged randomly. A quick survey question, "*Have you ever seen these words over the past four weeks?*" was inserted at the bottom of the test to check whether participants had been exposed to the target words over the previous four weeks.

5) Student Opinion Survey

A student opinion survey was created to report the participants' preferences regarding glosses. Two questions were asked for all studied groups in the sample: (a) whether one prefers having a gloss; (b) which type of gloss is preferred.

3. Procedures

A total of 329 participants went through three stages. On the first day, participants were given the cloze test to check their English proficiency. Then, they were classified into two groups based on their cloze test scores. Students who received equal to or more than 18 were defined as "higher level," and those who were equal to or less than 13 were defined as "lower level." One hundred sixty-six undergraduate students who were registered in a college English reading

www.kci.go.kr

course belonged to the higher level, and 163 graduate students who were registered in an English reading course for graduate students were lower level. A pre-vocabulary test was not given because its exposure would influence the results of the subsequent vocabulary tests.

After a week, the students of each proficiency level were randomly divided into one of four gloss conditions: no-gloss, L1 gloss, L2 gloss, and L1+L2 gloss. They were given instructions regarding the experimental process: (a) they would read a passage for 15 minutes, (b) return the reading material, and (c) have their reading comprehension measured with a test. To clarify that the students should focus on the content of the story rather than another feature, the students were shown the reading comprehension test before reading the passage. Then, after the students had completed their reading, the comprehension test was administered, as the students had expected. Then they were given an unexpected MC vocabulary test. The students took an English and a Korean version of the immediate vocabulary test. In order to counterbalance the two versions of the vocabulary tests, as indicated in the immediate and delayed vocabulary tests section, half the students in each gloss condition took the Korean version first and then the English version, while the other half took the test the English version first. Finally the students completed an opinion survey regarding their gloss preference.

Four weeks later, an unexpected delayed vocabulary test was administered to the same students. Participants were given the same amount of time to get the test done, and the ones who had taken the test in Korean first then took it in English first, and vice versa.

4. Analyses

The number of words gained on the immediate and delayed vocabulary tests was calculated by assigning one point to a correct answer--but only when the answer was correct on both the Korean and English versions of the test. The purpose of comparing the score on the Korean version of the vocabulary test to the English version was to determine with more certainty whether a student had learned words, since students may seem to understand an L2 word when relying on an L1 version of the test but actually do not understand it in an L2 context. Furthermore, by having two versions of the test one can be certain that students did not randomly answer a question correctly. For the student opinion survey, participants' responses were tallied so that trends could be identified.

IV. RESULTS

1. The Immediate Vocabulary Test

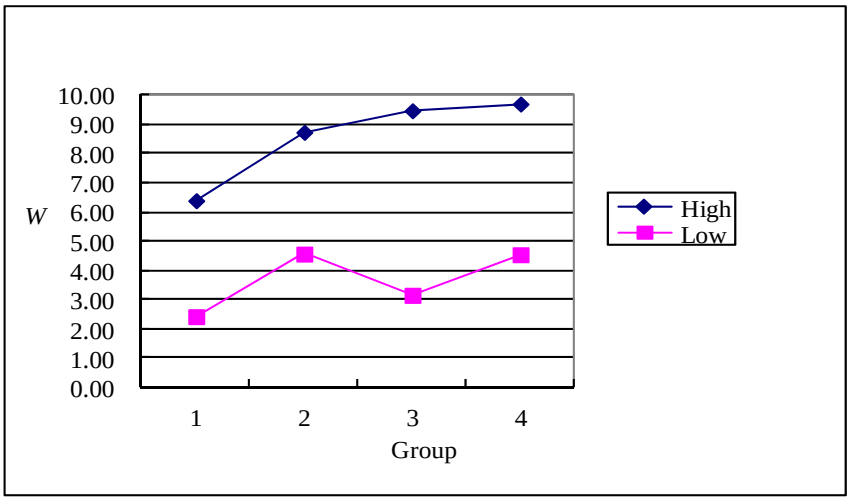
For all statistical analysis of the present study, the alpha level was set at 0.05. A factorial

www.kci.go.kr

ANOVA was conducted to check the interaction effect between L2 proficiency and gloss type on the immediate vocabulary test. Cronbach's alpha analysis demonstrated that the reliability of the immediate vocabulary test was 0.84.

A 2 x 4 between-subjects factorial ANOVA was calculated to compare the immediate vocabulary test scores of the students who had two levels of English proficiency and read a passage with four different gloss type. A significant interaction between L2 proficiency and gloss type was present ($F(3, 321) = 4.35$, $p < 0.05$). The main effect for proficiency was significant ($F(1, 321) = 357.23$, $p < 0.05$). The main effect for gloss type as well produced a significant effect ($F(3, 321) = 20.82$, $p < 0.05$). The interaction between L2 proficiency and gloss type had a strength rate of 3.90%, and L2 proficiency itself accounted for much more of the variance, 52.70%, while gloss type accounted for 16.30%.

Figure 1 illustrates L2 proficiency by gloss type on the immediate vocabulary test. It shows that L2 gloss seems to have been effective for higher-level learners but not effective for lower-level learners. In other words, depending on the L2 proficiency level, a learner needs different gloss treatment in order to learn L2 vocabulary more effectively.



W = words gained
Group = gloss conditions (1=no-gloss, 2=L1 gloss, 3=L2 gloss, 4=L1+L2 gloss)

FIGURE 1 L2 Proficiency × Gloss Type on the Immediate Vocabulary Test

Table 1 depicts descriptive statistics for the immediate vocabulary test, given the different gloss type used by students. The mean score is the total number of correct answers on both L1 and L2 vocabulary tests, and words gained refers to the number of words acquired, and they were calculated by dividing the mean by two (*M*/2), since there

www.kci.go.kr

were two versions of the vocabulary test. In the case of higher-level learners, the L1+L2 gloss group achieved the highest score (9.71), followed by the L2 gloss group with a slightly lower score (9.49). In the case of lower-level learners, the L1 gloss group achieved the highest score (4.57), followed by the L1+L2 gloss group with a very similar score (4.55).

TABLE 1

Means and Standard Deviations of the Immediate Vocabulary Test

Proficiency Level	Groups	*n*	Words Gained *M* /2	*SD*
Higher-Level	No gloss	40	6.40	2.27
	L1 gloss	43	8.74	2.57
	L2 gloss	41	9.49	2.54
	L1+L2 gloss	42	9.71	2.22
	Total	166	8.58	2.40
Lower-Level	No gloss	41	2.44	1.64
	L1 gloss	40	4.57	2.53
	L2 gloss	42	3.17	2.54
	L1+L2 gloss	40	4.55	2.33
	Total	163	3.68	2.26

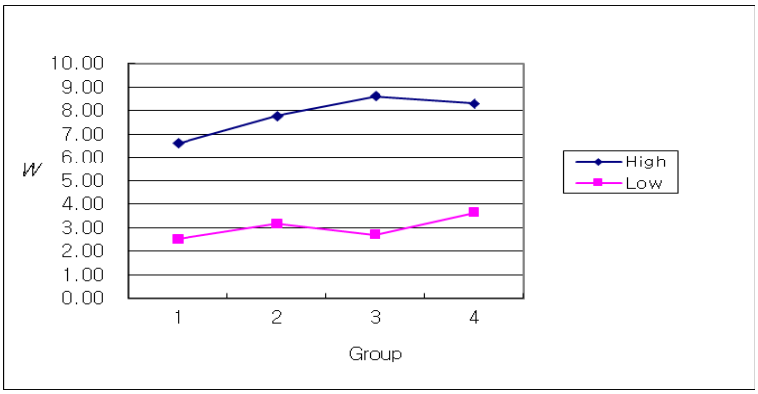
2. The Delayed Vocabulary Test

A 2 × 4 factorial ANOVA was also calculated to check the interaction effect between L2 proficiency and gloss type on the delayed vocabulary test. Cronbach's alpha analysis showed that the reliability of the delayed vocabulary test was 0.81.

The interaction between L2 proficiency and gloss type was not significant ($F(3, 321) = 2.61$, $p > 0.05$). The main effect for L2 proficiency was significant ($F(1, 321) = 404.05$, $p < 0.05$), and the main effect for gloss type was also significant ($F(3, 321) = 6.30$, $p < 0.05$). Compared to the statistical results of the immediate vocabulary test, those of the delayed vocabulary test were obviously different. The delayed vocabulary test was affected by both L2 proficiency and gloss type. L2 proficiency accounted for 55.70% of the difference and gloss type accounted for 5.60%, which means that L2 proficiency played a more important role than gloss type in explaining the variance.

Figure 2 presents L2 proficiency by gloss type for the delayed vocabulary test. Consistent with the immediate vocabulary test's results, the L2 gloss appears to have been effective for the higher-level learners but not for the lower-level learners.

www.kci.go.kr



W = words gained

Group = gloss conditions (1=no-gloss, 2=L1 gloss, 3=L2 gloss, 4=L1+L2 gloss)

FIGURE 2 Proficiency × Gloss Type on the Delayed Vocabulary Test

Table 2 depicts descriptive statistics for the delayed vocabulary test under different gloss conditions. The mean score is the total number of correct answers on both L1 and L2 vocabulary tests, and words gained refers to the number of words acquired. Word gained was calculated by dividing the mean by two (*M*/2), since there were two versions of the vocabulary test. The findings showed a somewhat different result compared to the immediate vocabulary test. In the case of the higher-level learners, the L2 group achieved the highest score (8.61), followed by the L1+L2 gloss group with a slightly lower score (8.31). In the case of the lower-level learners, the L1+L2 gloss group achieved the highest score (3.62), followed by the L1 gloss group with a slightly lower score (3.15).

TABLE 2

Means and Standard Deviations of the Delayed Vocabulary Test

Proficiency Level	Groups	*n*	Words Gained *M*/2	*SD*
Higher-Level	No gloss	40	6.60	2.04
	L1 gloss	43	7.77	2.84
	L2 gloss	41	8.61	2.74
	L1+L2 gloss	42	8.31	2.26
	Total	166	7.82	2.47
Lower-Level	No gloss	41	2.51	1.83
	L1 gloss	40	3.15	1.74
	L2 gloss	42	2.69	1.74
	L1+L2 gloss	40	3.62	1.80
	Total	163	2.99	1.78

Scheffe's post hoc analysis was conducted to locate the specific significant differences since L2 proficiency and gloss type indicated the main effects. Regarding L2 proficiency, it

www.kci.go.kr

was easy to locate the significant differences since there were only two levels (higher and lower). With respect to gloss type, there were no significant differences among the different gloss conditions, as shown in Table 3. However, having some kind of gloss was significantly more effective than not having any glosses at all.

TABLE 3

Gloss Conditions	Mean Difference	Standard Error	*Sig.*
No gloss vs. L1 gloss	-1.01	0.34	0.03*
No gloss vs. L2 gloss	-1.08	0.34	0.01*
No gloss vs. L1+L2 gloss	-1.49	0.34	0.00*
L1 gloss vs. L2 gloss	-0.07	0.34	0.99
L1 gloss vs. L1+L2 gloss	-0.48	0.34	0.56
L2 gloss vs. L1+L2 gloss	-0.41	0.34	0.69

*$p < .05$

3. Time Effect

The effect of time on the participants' vocabulary learning was measured with a paired *t*-test, considering each gloss condition, as shown in Table 4.

TABLE 4

A Paired *t*-Test of Vocabulary Scores Over Time

Proficiency Level	Groups	*n*	Words Gained on Immediate Test	Words Gained on Delayed Test	*t*	*Sig.*
Higher-Level	No gloss	40	6.40	6.60	-0.64	0.52
	L1 gloss	43	8.74	7.77	2.59	0.01*
	L2 gloss	41	9.49	8.61	2.72	0.01*
	L1+L2 gloss	42	9.71	8.31	5.77	0.00*
Lower-Level	No gloss	41	2.44	2.51	-0.25	0.80
	L1 gloss	40	4.57	3.15	4.34	0.00*
	L2 gloss	42	3.17	2.69	1.27	0.21
	L1+L2 gloss	40	4.55	3.62	3.30	0.00*

*$p < .05$

In the case of the higher-level group of learners, the no-gloss condition had no significant effect ($t(39) = -0.64$, $p > 0.05$). In the case of the lower-level learners, both the no-gloss and the L2 gloss conditions had no significant effect ($t(40) = -0.25$, $p > 0.05$ and $t(41) = 1.27$, $p > 0.05$, respectively).

www.kci.go.kr

4. Student Opinion Survey

To check student opinions on glossing and gloss type, the student opinion survey was analyzed after responses were tallied. As seen in Table 5, most of the students in both proficiency groups (about 95%) reported that they preferred having glosses. Only around 5% reported that they did not want glosses in reading materials. Those who wanted glosses preferred L1+L2 glosses the most (about 56.50%).

TABLE 5

Student Preference in Glossing and Gloss Type

	Preferred Glosses			Preferred No gloss
Higher-Level	94.74%			5.26%
	L1 gloss	L2 gloss	L1+L2 gloss	
	7.41%	36.11%	56.48%	
	Preferred Glosses			Preferred No gloss
Lower-Level	94.70%			5.30%
	L1 gloss	L2 gloss	L1+L2 gloss	
	19.20%	14.40%	66.40%	

Data from the L1+L2 gloss group were analyzed again because these students had actually experienced the L1+L2 glosses while reading. Table 6 presents the L1+L2 gloss group's preferences regarding gloss type. Interestingly, compared to the responses of the whole sample, more students in L1+L2 gloss condition preferred L1+L2 glosses in their reading materials, regardless of their proficiency levels.

TABLE 6

L1+L2 Group's Preference Regarding Gloss Type

	Gloss Type		
	L1 gloss	L2 gloss	L1+L2 gloss
Higher-Level	4.76%	23.81%	71.43%
Lower-Level	12.50%	10.00%	77.50%

V. DISCUSSION

1. Are There Any Interaction Effects Between L2 Proficiency and Gloss Type on an Immediate Vocabulary Test?

The first research question examined whether there were any interaction effects between L2 proficiency and gloss type on an immediate vocabulary test. The answer is that there

www.kci.go.kr

was an interaction effect between these two independent variables. As seen in Figure 1, overall, the groups in the higher-level condition outperformed the lower-level groups. Apparently, the L2 gloss was an effective gloss type for the higher-level learners, but not for the lower-level learners.

As seen in Table 1, in the higher-level group, students acquired the most words in the L1+L2 gloss condition (9.71 words), while the L2 gloss condition followed (9.49 words). L2 gloss or L1+L2 gloss appeared to both be effective among various gloss options for the higher-level learners. According to Table 5, about 56.50% of the participants in the higher-level group reported that they prefer having L1+L2 glosses in their reading materials. The higher-level learners in the present study showed their best performance with and preference for the L1+L2 gloss type. This indicates that they may have been in the learning stage between the intermediate model and concept mediation model, which is explained in the introduction of this paper. The higher-level learners in the present study were able to use an L2 semantic concept when encountering an L2 word, but they also usually needed a word-to-word link (Kroll & Sunderman, 2003).

In the case of Laufer and Hill (2000), Hong Kong students who had quite high proficiency performed best in the L2 gloss condition. These students' most frequent look-up option was the L1/L2 with other information. The Hong Kong students in Laufer and Hill (2000) seemed to lie in the domain of the concept mediation model, in which learners were able to use L2 concept links without associating them with L1 equivalents (Kroll & Sunderman, 2003; Kroll et al., 2010). Hu et al. (2014) discovered similar findings. The two kinds of immediate vocabulary tests revealed similar results; both in the first and second experiment, two higher-level groups performed significantly better than the lower-level groups. There were no significant differences between L1 glosses and L2 glosses within the same learner proficiency groups. Unfortunately, Laufer and Hill (2000) and Hu et al. (2014) could not explain the effects of L1+L2 glosses since they did not include the combination of L1 and L2 glosses in their research design.

In the case of lower-level group, as seen in Figure 1, the L2 gloss appeared to be the least effective gloss type among all the options. Lower-level learners in the L2 gloss condition performed very close to those in the no-gloss condition, which means that the effectiveness of L2 gloss was very trivial for the lower-level learners. Lower-level learners performed similarly given the L1 gloss and L1+L2 gloss conditions. For the lower-level learners, the L1-related gloss type appeared to be effective. The lower level learners of the present study may likely be described by the word association model, where they used L1 words as a medium to process L2 words. It seemed that their use of L1 was detrimental when encountering L2 stimulus. However, interestingly, participants in the present study reported that they would prefer having L1+L2 glosses in their reading materials. As shown in Table 5,

www.kci.go.kr

about 66% of the lower-level learners preferred L1+L2 gloss and only about 19% of them wanted L1 glosses only.

In the case of Laufer and Hill (2000), the Israeli students performed the best when they looked up both L1 and L2 glosses and less effectively in the L1 gloss condition. In this sense, the Israeli students appear to have had a slightly higher L2 proficiency than those of the present study. In the case of Hu et al. (2014), unsurprisingly, the low-level groups performed better with L1 glosses. However, they could not explain the effects of L1+L2 glosses due to not including the combination of L1+L2 glosses in their studies.

2. Are There Any Interaction Effects Between L2 Proficiency and Gloss Type on a Delayed Vocabulary Test?

The second research question examined whether there were any interaction effects between L2 proficiency and gloss type on a delayed vocabulary test. The answer is that there was no interaction effect between L2 proficiency and gloss type. The reason that the interaction effect between L2 proficiency and gloss type disappeared may have to do with time duration after exposure to the target words. The strength of the interaction effect between these two independent variables was not strong (3.90%) on the immediate vocabulary test. Besides, four weeks is rather a long period to sustain the incidentally acquired words. According to the student survey, almost all students had not recalled seeing the words between the immediate and delayed vocabulary tests. Consequently, the role of different gloss type might have been weakened, and this may have influenced the removal of the interaction effect over four weeks.

There were main effects for the two independent variables on the delayed vocabulary test. The strength of the variance for L2 proficiency was much greater, 55.70%, and the strength of the variance for gloss type was much smaller, 5.60%. The main effect of L2 proficiency was significant. Those at the higher level retained 7.82 words out of 14, whereas those at the lower level retained 2.99 words. It is obvious that higher level students could retain significantly more acquired words after four weeks.

The main effect for gloss type was significant; there was a significant difference between the no-gloss and the glossed conditions. Participants in the glossed conditions performed significantly better than those in the no-gloss group. In other words, regardless of proficiency levels, those in the L1 gloss, L2 gloss, and L1+L2 gloss conditions performed significantly better than those in the no-gloss condition. However, there were no significant differences among glossed conditions. These results have been interpreted to mean that glossing had a

www.kci.go.kr

positive effect on vocabulary retention, but that the gloss language did not influence the retention of words after four weeks.

In Hu et al. (2014), participants took the delayed vocabulary test after two weeks. In the case of the cloze test, the two higher-level groups performed significantly better than the two lower-level groups. It seems that there was a proficiency effect. In the case of definition-supply test, in the first experiment, only the higher-level receiving L2 gloss before L1 gloss performed significantly better than the two lower-level groups, whereas in the second experiment, the higher-level receiving L1 glosses before L2 glosses performed significantly better than the two lower-level groups. For the higher-level groups, it seems that there was a gloss-language effect. Compared to the present study, Hu et al. (2014) might have had different results in vocabulary retention, due to test methods and the different interval between the immediate and delayed tests. Unfortunately, comparison between Laufer and Hill (2000) and the present study regarding delayed vocabulary test was not able to make since they did not give any delayed vocabulary test.

3. Is There a Time Effect in Vocabulary Scores Between an Immediate and a Delayed Vocabulary Tests?

The third research question was whether there was a time effect between the two vocabulary tests. Laufer and Hill (2000) did not give a delayed vocabulary test. Hu et al. (2014) did not check the time effect between the immediate and delayed vocabulary tests. The present study checked whether the eight gloss conditions were correlated with different rates of forgetting acquired words between the immediate and delayed tests. Five out of eight gloss conditions showed a significant decline rate of mean scores between these two tests. It means that students in these gloss conditions forgot a significant number of their incidentally acquired words. However, interestingly, three conditions did not show statistical differences over four weeks: (a) the no-gloss condition at the higher level; (b) the no-gloss condition at the lower level; and (c) the L2 gloss condition at the lower level. The reason that these three gloss conditions did not show any significant decline in the rate of word knowledge over four weeks may be due to either vocabulary pre-knowledge or ineffective gloss type, not because they were effective for the retention of incidentally acquired words.

For the no-gloss group at the higher level, there was no significant decline between the immediate and delayed vocabulary tests. As discussed in M. H. Ko (2012), the participants may have had some knowledge of the target words before the present study, even though the 14 words selected were not known by 60% of participants in the pilot study, and almost all students reported on the survey question *"Have you ever seen these words over the past*

www.kci.go.kr

four weeks?" that they had not seen the target words. Thus, the participants' immediate test and delayed test scores probably demonstrated their vocabulary pre-knowledge.

Again, the no-gloss group at the lower level showed a very similar result. There was no significant decline over time. The reason this group did not indicate a significant difference may be because they already knew two or three of the words before the experiment occurred.

The outcome for the L2 gloss at the lower level is worth noting. The mean score of this group slightly dropped between the immediate and delayed vocabulary tests, but there was no significant time effect on vocabulary retention. The number of words gained by lower level learners under the L2 gloss condition was low enough that it was very close to that of the no-gloss group, meaning that providing L2 glosses in the reading materials to the lower-level learners was similar to not providing any glosses.

4. What Are Learner Opinions Regarding Different Gloss Type?

The fourth research question sought to reveal student opinions about glossing and preference for gloss type. Laufer and Hill (2000), and Hu et al. (2014) did not investigate student opinions regarding glossing. As seen in Table 5, around 95% of the students of the present study in both proficiency level groups reported that they preferred having glosses in L2 reading passages. In the case of higher-level learners, about 56.50% of the students preferred L1+L2 gloss, followed by about 36% who preferred L2 glosses. Only about 7% wanted L1 glosses in their reading materials. In the case of lower-level learners, about 66% preferred L1+L2 glosses, followed by about 19% who preferred L1 glosses. Around 14% of lower level learners wanted L2 glosses in their reading materials.

The preference of students in the L1+L2 gloss condition regarding gloss type was further analyzed because it was thought that since this group had experienced L1+L2 glosses while reading the passage, their data would be more relevant to the current study's main research question regarding what type of gloss is most useful. As seen in Table 6, about 71% of higher-level learners wanted L1+L2 glosses, and 77.50% of the lower-level learners wanted L1+L2 glosses. In short, those who had experienced L1+L2 glosses wanted L1+L2 glosses more.

The reason why higher-level learners want to have L1+L2 glosses may have been their current lexical processing ability. They appeared to fit between the intermediate and concept mediation model, and likely closer to the concept mediation model. Accordingly, they may have needed both L2 and L1 for the sake of effective processing of the L2 stimulus.

It is difficult to explain why lower-level learners wanted L1+L2 glosses. These participants appeared to be in the first stage of lexical processing, described by the word

www.kci.go.kr

association model, wherein they needed the help of a corresponding L1 word when encountering a L2 stimulus. However, although they were not fully capable of processing L2 glosses, they may have considered L2 glosses beneficial, perhaps because they had heard that such glosses are useful to add richness of semantic assistance. In other words, learners at the lower level might have thought that it is better to have both L1 and L2 glosses, just in case. Their psychological need for security may have led them to want both L1 and L2 input.

VI. CONCLUSION

The present study investigated the relationship between L2 proficiency and gloss type on L2 incidental vocabulary learning. It included the L1+L2 gloss type, along with other gloss treatment conditions, and also incorporated L2 proficiency into the research design. The results of the present study add valuable information to the current gloss database of L2 vocabulary learning and teaching.

The interaction effect between L2 proficiency and gloss type indicated that gloss type should be carefully chosen, grounded upon learner L2 proficiency. In the case of higher-level learners, the L2 gloss and L1+L2 gloss appeared to be more effective, whereas in the case of lower-level learners, the L1 gloss or L1+L2 gloss seemed to be more effective. However, the student opinion survey regarding gloss preference indicated that the participants preferred L1+L2 glosses to the other types, regardless of their L2 proficiency levels.

Regarding retention of incidentally acquired words, there were no statistical differences between immediate and delayed vocabulary tests over four weeks in the case of no-gloss conditions and L2 glosses for the lower-level group. The fact that L2 glosses for the lower-level group showed a similar result compared with the no-gloss condition may indicate that the L2 glosses were not effective in the case of lower-level learners.

Considering the statistical analysis of the interaction effect between L2 proficiency and gloss type, and the reports of the student opinion survey, L1+L2 glosses are likely the best gloss type for L2 reading materials at any proficiency level. Given the fact that readers from various levels are likely to encounter the same reading materials, providing one type of gloss may not be beneficial for all circumstances. In other words, not only intermediate level learners, but also high-beginners, low-intermediate learners, and even high-intermediate learners may encounter the same reading materials even if they are designed for the intermediate learners. Therefore, providing both L1 and L2 gloss may be the most efficient way to assist reading comprehension and facilitate vocabulary learning for L2 readers.

www.kci.go.kr

Recently various ELT reading materials including storybooks have utilized glossing in the texts, but most cases they included L2 glosses only. Paper-based reading materials might have difficulties in providing L1 and L2 glosses simultaneously, due to limited offline space. Along with technology advancement, online reading materials have been gradually increasing in ELT field. In the case of online learning environment, integrating L1+L2 glosses into online text may be easily feasible. Accordingly reading materials designers may be able to incorporate L1+L2 glosses into online texts through various technology-assisted platforms.

The present study suggests that further research on repeated exposure to target words in line with L2 proficiency and gloss type may uncover interesting findings. Repeated exposures to target words are likely to enable learners to store newly learned words in their long-term memory. Thanks to the development of technology, various kinds of technology-assisted platforms, such as computers, the Internet, smartphones, and software applications, are being developed for vocabulary learning. In particular, mobile or easily-portable devices could provide both carefully controlled repetition and a ubiquitous learning environment in which learners could learn and review target words at any time and in any place. Such further study may offer a more precise picture regarding the extent to which repeated exposure affects ways of learning L2 vocabulary in relation with L2 proficiency and gloss type.

REFERENCES

Blanton, L. L., & Lee, L. (1995). *The multicultural workshop*. Boston, MA: Heinle & Heinle.

Bowles, M. A. (2004). L2 glossing: To CALL or not to CALL. *Hispania, 87*(3), 541-552.

Chen, H.-C., & Leung, Y.-S. (1989). Patterns of lexical processing in a nonnative language. *Journal of experimental psychology: Learning, memory, and cognition, 15*(2), 316-325.

Cheng, Y.-H., & Good, R. L. (2009). L1 glosses: Effects on EFL learners' reading comprehension and vocabulary retention. *Reading in a Foreign Language, 21*(2), 119-142.

Coady, J. (1993). Research on ESL/EFL vocabulary acquisition: Putting it in context. In T. Huckin, M. Haynes, & J. Coady (Eds.), *Second language reading and vocabulary learning* (pp. 3-23). Norwood, NJ: Ablex Publishing.

Gan, X. (2014). Study on the effects of gloss type on Chinese EFL learners' incidental vocabulary acquisition. *Theory and Practice in Language Studies, 4*(6), 1251-1256.

www.kci.go.kr

Gass, S. M. (1988). Second language vocabulary acquisition. *Annual Review of Applied Linguistics, 9*, 92-106.

Gettys, S., Imhof, L. A., & Kautz, J. O. (2001). Computer-assisted reading: The effect of glossing format on comprehension and vocabulary retention. *Foreign Language Annals, 34*(2), 91-99.

Hu, S.-M., Vongpumivitch, V., Chang, J. S., & Liou, H.-C. (2014). The effects of L1 and L2 e-glosses on incidental vocabulary learning of junior high-school English students. *ReCALL, 26*(1), 80-99.

Huang, H.-T., & Liou, H.-C. (2007). Vocabulary learning in an automated graded reading program. *Language Learning & Technology, 11*(3), 64-82.

Hulstijn, J. H., Hollander, M., & Greidanus, T. (1996). Incidental vocabulary learning by advanced foreign language students: The influence of marginal glosses, dictionary use, and reoccurrence of unknown words. *Modern Language Journal, 80*(3), 327-339.

Jacobs, G., Dufon, P., & Hong, F. C. (1994). L1 and L2 vocabulary glosses in L2 reading passages: Their effectiveness for increasing comprehension and vocabulary knowledge. *Journal of Research in Reading, 17*(1), 19-28.

Ko, Myong-Hee. (2012). Glossing and L2 vocabulary learning. *TESOL Quarterly, 46*(1), 56-79.

Krashen, S. (1989). We acquire vocabulary and spelling by reading: Additional evidence for the input hypothesis. *Modern Language Journal, 73*(4), 440-464.

Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language, 33*(2), 149-174.

Kroll, J. F., & Sunderman, G. (2003). Cognitive processes in second language learners and bilinguals: The development of lexical and conceptual representations. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 104-129). Berlin: Blackwell publishing.

Kroll, J. F., Van Hell, J. G., Tokowicz, N., & Green, D. W. (2010). The revised hierarchical model: A critical review and assessment. *Bilingualism: Language and Cognition, 13*(3), 373-381.

Laufer, B. (1996). The lexical plight in second language reading. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition* (pp. 20-34). Cambridge: Cambridge University Press.

Laufer, B., & Hill, M. (2000). What lexical information do L2 learners select in a CALL dictionary and how does it affect word retention? *Language Learning & Technology, 3*(2), 58-76.

www.kci.go.kr

Lomicka, L. L. (1998). “To gloss or not to gloss”: An investigation of reading comprehension online. *Language Learning & Technology, 1*(2), 41-50.

Moazzeni, Z., Bagheri, M. S., Sadighi, F., & Zamanian, M. (2015). The effect of glossing technique on incidental vocabulary retention, a comparison between textual and mult-media glosses. *Modern Journal of Language Teaching Methods, 5*(1), 222-235.

Nagata, N. (1999). The effectiveness of computer-assisted interactive glosses. *Foreign Language Annals, 32*(4), 469-479.

Nation, I. S. P. (1982). Beginning to learn foreign vocabulary: A review of research. *RELC Journal, 13*(1), 14-36.

Nation, I. S. P. (2001). *Learning vocabulary in another language*. New York: Cambridge University Press.

Nation, I. S. P. (2013). *What should every EFL teacher know?* Seoul: Compass Publishing.

Oller, J. W. Jr. (1972). Scoring methods and difficulty levels for cloze tests of proficiency in English as a second language. *Modern Language Journals, 56*(3), 151-158.

Potter, M. C., So, K.-F., Eckardt, B. V., & Feldman, L. B. (1984). Lexical and conceptual representation in beginning and proficient bilinguals. *Journal of Verbal Learning and verbal behavior, 23*(1), 23-38.

Rott, S., Williams, J., & Cameron, R. (2002). The effect of multiple-choice L1 glosses and input-output cycles on lexical acquisition and retention. *Language Teaching Research, 6*(3), 183-222.

Rouhi, A., & Mohebbi, H. (2012). The effect of computer assisted L1 and L2 glosses on L2 vocabulary learning. *The Journal of Asia TEFL, 9*(2), 1-19.

Stein, M. J. (1993). The healthy inadequacy of contextual definition. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition* (pp. 203-214). New York: Cambridge University Press.

Sung, Hyunmi. (2016). Task-induced involvement load and low-level learners’ vocabulary learning. *Modern English Education, 17*(4), 1-24.

Thornbury, S. (2002). How words are learned? In J. Harmer (Ed.), *How to teach vocabulary* (pp. 13-31). Edinburgh Gate: Pearson Education Limited.

Watanabe, Y. (1997). Input, intake, and retention: Effects of increased processing on incidental learning of foreign language vocabulary. *Studies in Second Language Acquisition, 19*(3), 287-307.

Yanguas, I. (2009). Multimedia glosses and their effect on L2 text comprehension and vocabulary learning. *Language Learning & Technology, 13*(2), 48-67.

Yoshii, M. (2006). L1 and L2 glosses: Their effects on incidental vocabulary learning. *Language Learning & Technology, 10*(3), 85-101.

www.kci.go.kr

Examples in: English
Applicable Languages: English
Applicable Levels: Tertiary

Myong-Hee Ko
Soonchunhyang University
Hyangseol Sharing Department
22 Soonchunhyang-ro, Shinchang-myeon, Asan-si, Chungcheongnam-do 31538
Tel: 041-530-1783
Email: myongheeko@naver.com

Received 14 June 2017
Revised 2 August 2017
Accepted 14 August 2017

www.kci.go.kr