Modern English Education, Vol. 18, No. 4, Winter 2017
http://dx.doi.org/10.18095/meeso.2017.18.4.04

L1 vs. L2 Glosses, Learner Perceptions, and L2 Vocabulary Learning*

Hyung-Sun Kim**
Chosun University
Ung-Yong Choi
Chonnam National University

Kim, Hyung-Sun & Choi, Ung-Yong. (2017). L1 vs. L2 glosses, learner perceptions, and L2 vocabulary learning. *Modern English Education*, *18*(4), 67-89.

Studies on vocabulary learning have shown conflicting results on the effect of L1 and L2 glosses. This study attempted to examine the different effects that L1-Korean and L2-English glosses would bring about, employing the new method of the Tracing Measure of Retention (TMR), which traced the individualized target words one by one through immediate and delayed tests. The learners' perceptions and preferences of the glossing language were also investigated through a questionnaire survey. The participants were 152 undergraduate students from various majors at a Korean university, who were divided into four proficiency levels as per their CSAT scores. While the L1 and L2 glosses did not create meaningful differences overall, the high intermediate level learners performed better with the L1 glosses at a statistically significant level. The participants voted overwhelmingly for the L1 glosses, but among those who experienced the L2 glosses, the L1 preference was not extreme. The tendency was evident in the high intermediate learners, which may be taken as a sign that they are in transition to the stage of direct conceptual processing of L2 words. While general effectiveness of L1 glossing was acknowledged, the need to consider learners' state to benefit from L2-mediated input enhancement was also highlighted.

**[L1 glosses/L2 glosses/learner perceptions/vocabulary learning/retention/
모국어 주석/외국어 주석/학습자 인지/어휘학습/어휘유지]**

---

* Part of the data of this study was used in H. S. Kim and U. Y. Choi (2017) for illustration purpose of the newly developed measure of vocabulary learning, TMR.

** First author: Hyung-Sun Kim, Corresponding author: Ung-Yong Choi

I. INTRODUCTION

Acknowledging that linguistic proficiency is attained on the basis of lexical competence (Godwin-Jones, 2010; Zhang & Li, 2011), researchers have extensively explored ways to facilitate vocabulary learning, to the extent that L2 learning is equated with its vocabulary learning. Three approaches in vocabulary development seem to be often discussed as suggested by Hunt and Beglar (2002) in their review of the research and teaching trend of vocabulary: incidental learning, explicit instruction, and independent strategy development. They suggested that while all of the approaches have a role in learners' vocabulary development, most words are learned incidentally, for example, through reading (Hemmati & Asmawi, 2015; Hulstijn, Hollander, & Greidanus, 1996; Nation, 2001) and/or listening (Brown, Waring, & Donkaewbua, 2008; Elley, 1989).

As noted from the considerable amount of studies on vocabulary learning through reading, many researchers appear to be convinced that exposure to written language, rather than that to oral language, is the primary contributor to learners' lexical competence (Cunningham & Stanovich, 2001; Verhoeven & Perfetti, 2011). It would be especially so in an EFL environment such as in Korea, where opportunities for exposure to authentic oral English are limited. On the other hand, researchers have found that L2 vocabulary learning through reading may lead to some drawbacks (Huckin & Coady, 1999; Hulstijn, 1993; Hulstijn, Hollander, & Greidanus, 1996). For example, learners may infer the meaning of unknown words incorrectly, particularly for items with deceptive morphological structures (Laufer, 1997) and idioms that are not construed from the meaning of their components. In addition, while the incidental learning of word meanings occurs in terms of small increments (Frishkoff, Perfetti, & Collins-Thompson, 2011), chances of incremental learning may not be enough for learners to build solid lexical knowledge because essential vocabulary does not always occur frequently in texts.

Other researchers such as Schmidt (1994) suggested that incidental learning should involve some extent of consciousness along with its mainly subconscious process, in the same way as noticing plays a key role in learning. In fact, many studies have dealt with the consciousness element, along with examining ways to promote incidental vocabulary learning by means of input enhancement or modification (Ellis & He, 1999; Laufer & Hulstijn, 2001; Schmitt, 2010). Providing marginal glosses was one of the most common ways for input modification to promote vocabulary learning due to the convenience in implementation and the benefits: (1) it assists learners to acquire less frequent but essential vocabulary; (2) it helps learners be less distracted from the reading when they encounter unknown words in texts than, for example, consulting a dictionary; and (3) it promotes learners' comprehension of the text by providing clues for the content of the text (Nation, 2013).

One of the points which have not gained much attention is learners' perception or attitude toward the glossing and the glossing languages. When learners evaluate learning activities highly, teachers would believe it is because the activities have generated their high performances. However, some research studies (Cashin & Downey, 1992; O'Connell & Dickinson, 1993; Ryan & Harrison, 1995) indicated that it is not always the case; the learner perceptions of learning in their classes were more highly correlated with the course satisfaction than with the learner performances in their experiments. Likewise, in addition to measuring their progress on test scores, how a treatment is perceived and interpreted by learners should be considered to make their gains from the learning activity sustainable.

This study—as a follow-up study of H. S. Kim and U. Y. Choi (2017) that suggested a new way to measure vocabulary learning—also deals with the common topic of marginal glosses and centers on different effects which could possibly result from the use of different languages. This issue has been explored in many empirical studies, which have not always yielded agreeing results. One of things that differentiate this study from previous research with similar methodology is that learners' perception was additionally considered to determine how their attitudes toward glosses and/or glossing languages are interrelated with their learning. This study has attempted to answer the following research questions:

1) How did L1 and L2 glosses differently influence L2 learners' vocabulary learning, overall and/or by proficiency level, as measured by immediate and delayed tests?
2) How did learners perceive L1 and L2 glosses? Are their perceptions related with their vocabulary learning as measured by immediate and delayed tests?

II. LITERATURE REVIEW

Glosses refer to short definitions (Nation, 2013), brief explanations, or translations into L1 for unusual or difficult words. Their typical form would be in textual cues, but they involve any "attempts to supply what is perceived to be deficient in a reader's … knowledge" (Roby, 1999, p. 96), such as pictorial, aural, mixed-type, and multimedia interactive glosses. In the L2 context, glosses have been employed in the side or bottom of pages basically to promote L2 reading comprehension by limiting continual dictionary usage that could interrupt the reading process and cut the flow (Lomica, 1998).

From the viewpoint of vocabulary learning, glosses can be considered as a type of focus on form because they may draw learners' attention to lexical items while reading (Laufer & Girsai, 2008). Thus, previous studies on glossing have often concentrated on the effect of glosses as planned focus on form for L2 vocabulary learning.

Yoshii and Flaitz (2002) tested three types of annotation—text, picture, and combining the two—in an Internet reading environment and reported that the combination of textual and pictorial glosses resulted in the best performance of the learners. Hulstijn, Hollander, and Greidanus (1996) focused on textual cues examining the marginal glossing effect on vocabulary learning. They suggested that the frequency of the occurrence of unknown words would foster learning only in the presence of external information of their meaning, i.e., when glosses are provided or learners consult dictionaries. They indicated that readers often ignored unknown words or made incorrect guesses, which diminished the benefit of frequency.

The effect of the glossed and non-glossed texts was examined in M. H. Ko (2012) with Korean participants. She reported a significant positive result with the glossed ones but found no meaningful differences between the L1 and L2 glossed texts, as was the case in Yoshii (2006). Rouhi and Mohebbi's (2012) findings with Iranian participants were partially contradictory to the studies. They experimented with three groups of the learners—L1 glosses, L2 glosses, and no glosses—using the Scaffoglossing software, and found that the glossed group outperformed the non-glossed group and within the glossed group, the L1 glosses outscored the L2 glosses.

These divergent results were also found in earlier studies that compared L1 and L2 glosses in various L2 settings: Laufer and Hill (2000) reported that the Hong Kong learners of L2 English performed significantly better with the L2 glosses; in Miyasako's (2002) study with the Japanese participants, the L2 group also performed significantly better than the L1 group; Chen's (2002) study with the Taiwanese participants reported no significant difference between the L1 and L2 glossed groups; Jacobs, Dufon, and Hong (1994) carried out an experiment with 85 Spanish learners who were native speakers of English and reported that there was no significant difference between the L1 vs. L2 gloss effects.

In fact, there are few studies involving the learners' perception of the glossing languages. The above-mentioned Jacobs et al. (1994) exceptionally asked the participants, who were English-speaking Spanish learners, about their preference between the L1 and L2 glossing. They reported that the learners liked the L2 (Spanish) glosses better. Brooks-Lewis (2009) investigated the reverse setting in her more comprehensive and wider-scoped study on learners' perceptions of using L1 in L2 learning, with 256 Spanish-speaking English learners in Mexico. Based on the research results generated using "a qualitative-interpretative methodology" (p. 221), she made an assertive conclusion that using learners' L1 is "a manner of demonstrating respect for the learner, of taking the learner's viewpoint into account, and is a step towards meeting the learner half-way" (p. 234). Although the research settings of Jacobs et al. (1994) and Brooks-Lewis (2009) carried different research focuses, the contrast between the English-speaking Spanish learners and the Spanish-speaking English learners and the

results pointing to opposite conclusions seem noteworthy as they may be reflective of the learners' different perceptions towards L2.

According to Cook (2003, as cited in Brooks-Lewis, 2009), the traditional way of teaching an L2 was always through the learner's own L1. The L2 education intended to have learners equally conscious of the syntax and resources of the L1 as those of the target language (Howatt & Widdowson, 2004). However, the belief has been growing that words in different languages cannot be equivalent in semantics and that direct associations between words in L2 and their referents are essential to L2 learning (Howatt & Widdowson, 2004). L1 intervention, such as L1 glosses, often came to be conceived as interference and/or habits to be overcome as observed in many immersion language programs which ban L1 use in classrooms. This study will additionally consider this issue through the results of the questionnaire survey, mainly focusing on the glossing language effect on the participants' immediate vocabulary learning and retention.

III. Method

1. Participants

The participants were 152 undergraduate students taking a compulsory English course at a Korean university: students from Premed (29), Liberal Arts (62), and Natural Science and Engineering (61). Most of them were in their first year in the college. A couple of students who missed either of the immediate and delayed tests were excluded. The participants' English proficiency levels were identified by means of the results of the English section of the CSAT (College Scholastic Ability Test), which is administered by the ministry of education in Korea as a prerequisite to proceed to college education. Out of the total nine levels, the participants were distributed from Levels 1 to 7: Level 1 (24 students, 15.79%), Level 2 (19 students, 12.50%), Level 3 (35 students, 23.02%), Level 4 (43 students, 28.29%), Level 5 (18 students, 11.84%), Level 6 (5 students, 3.29%), Level 7 (1 student, .66%) and data missing (7 students, 4.60%). At the beginning of the course, this research project was explained to the students, and informed consents were obtained on collecting data including necessary personal information for the research purpose only.

2. Procedures

The treatment task which was intended to elicit the participants' incidental vocabulary learning was an in-class reading of a text from Jordens and Zeter (2013), which was glossed either in their L1 or L2. The reading material from the textbook was chosen for (1)

the topic, the early civilization of mankind, for which a fair unfamiliarity to the participants can be assumed since there were no participants majoring history, and (2) the readability level, the eleventh as per the Flesch Grade Level Readability Formula.[1] The text was arranged into two-page lengths, three paragraphs on each; glosses for thirty words were provided in the right margin, segregated according to paragraph in the sequence of occurrence in each paragraph. A participant's individual target words were identified from the total pool of target words (see Appendix). They consisted of the words in the glossary of the textbook and additional words of the researchers' choice to help facilitate reading. Figure 1 shows the procedure flow of the experiment using the words.

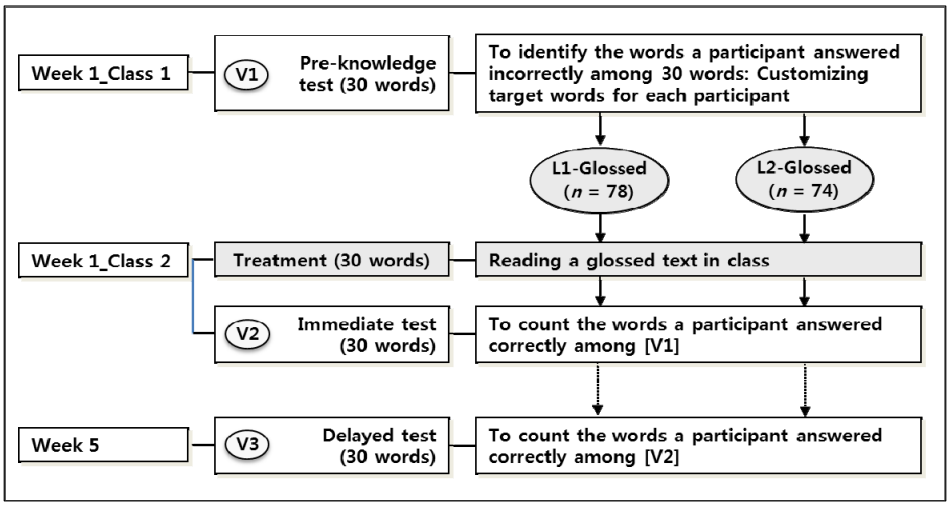

FIGURE 1 Procedure Flow of the Experiment Based on TMR

In Week 1_Class 1, the participants were given a ten-minute pre-knowledge test requiring them to write the meaning of the thirty words in Korean. This step was to identify unknown words for each participant; as an individual participant answered a different set of words incorrectly, each one would have a unique set of customized target words. In Week 1_Class 2, they were required to read the text with glossed vocabulary, either in L1 or L2, for thirty-minutes to answer a set of reading comprehension quizzes that, they were told falsely, would follow. The purpose was to have the participants focus on reading; instead they were given the immediate vocabulary test. They were asked to choose the correct meaning of a given word out of four choices provided in either L1 or L2. For example, for *wedge-shaped*, they were supposed to choose a. *in the form of one pointed end and one thicker end*, over b. *in the form of two sets of parallel lines*, c. *in the form of*

[1] http://www.readabilityformulas.com/

sparrows' footprint, and d. *in the form of a spider's web*. For the word, irrigation, they were supposed to choose a. *관개* over b. *수로*, c. *정작*, and d. *비옥*. After four weeks, the delayed test, without prior notice, was administered with the same repertoire of words in a different arrangement to measure the participants' vocabulary retention.

On the last page of the delayed test, the participants were asked to complete a mini-questionnaire made up of four items concerning their perceptions of the glossed reading task: two five-point Likert scale items ranging from 1-strongly disagree to 5-strongly agree (Q1 *I understood the text well*; Q2 *I benefited from the glosses.*), a multiple-choice question (Q3 *How did you use the glosses?*), and a binary-choice question (Q4 *I preferred L1/L2 glosses.*). They were also invited to make a short remark on their preference of glossing language.

3. Data Collection and Analysis

As shown in Figure 1, the three tests (pre-knowledge, immediate, and delayed tests) and the reading task all involved thirty words on the surface. However, each participant's score needed to be based on a different number of words as this study followed the Tracing Measure of Retention model (TMR) (H. S. Kim & U. Y. Choi, 2017) to measure the vocabulary learning. Table 1 illustrates that the TMR traces the target words customized for each participant to immediately learned words (Cases 1 and 2), and to retained words (Case 1); out of the eight possible cases, only Case 1 was counted as successful vocabulary learning through the reading task. Each item of the tests was input in an Excel worksheet for each participant; each item was scored as 1 or 0, making 30 cells for each test. The test scores of the immediate and delayed test were calculated utilizing the Excel formulas which were provided by H. S. Kim and U. Y. Choi (2017).

The students' mini-questionnaires were also coded on the Excel worksheet as a separate section so that they could be converted to data for the SPSS analysis together with the results of the tests. Their open comments were sorted into several categories that represented their overall perception of the L1 and L2 glosses.

As extensive studies have pointed out high performing learners' different behaviors in strategically planning their learning (H. Jeong, Biswas, Johnson, & Howard, 2010; Hofer & Yu, 2003; Nemati, 2013; Tekmen & Daloğlu, 2006), the participants were grouped into four different levels of English competence by making use of their CSAT levels to identify proficiency effects on vocabulary learning if any: Advanced (Level 1: 24 students), High Intermediate (Levels 2 and 3: 54 students), Intermediate (Level 4: 43 students), and Beginning (Levels 5, 6, & 7: 24 students). The number of participants of each level was considered so that the grouping would not create strong bias high or low which could be found in an extremely small group due to the consequent large difference in the number of

group members. In addition, the results of the multiple comparisons of the CSAT levels were referred to in order to ensure the appropriateness of the grouping, which divided Levels 3 and 4—both together constitute an intermediate level typically—into different levels; the Bonferroni post hoc tests for the participants' performances in the pre-knowledge, immediate, and delayed tests indicated differences between Levels 3 and 4.

TABLE 1

Vocabulary Learning in TMR

	Pre-knowledge Test	Immediate Test	Delayed Test	TMR		Pre-knowledge Test	Immediate Test	Delayed Test	TMR
Case 1	X	O	O	Yes	Case 5	O	O	O	No
Case 2	X	O	X	No	Case 6	O	O	X	No
Case 3	X	X	O	No	Case 7	O	X	O	No
Case 4	X	X	X	No	Case 8	O	X	X	No

Note. O: A participant's answer was correct. X: A participant's answer was incorrect.

IV. RESULTS

1. L1 vs. L2: Overall Comparisons of Their Influence on Vocabulary Learning

There were 78 students who received the L1-glossed material and completed both immediate and delayed tests following the pre-knowledge test (L1 group, hereafter); 74 students received the L2-glossed material and took all the tests students (L2 group, hereafter). Table 2 gives (1) the mean number of unknown words (the words a participant answered incorrectly in the pre-knowledge test: {V1}), (2) among {V1}, the ratio of the words a participant answered correctly in the immediate test (V2), i.e., their immediate learning (V2/V1), and (3) among {V2}, the ratio of the words a participant answered correctly in the delayed test (V3), which was weighted to reflect the immediate learning ratio of each participant (V3/V2*V2/V1). While the labels of immediate and delayed tests generally mean the test scores in other studies, they indicate the learning ratios based on a set of words individually targeted for each test in this study. Therefore, the weighting was necessary in order to make the values of the immediate and delayed tests comparable. The mean numbers of the unknown words out of 30 words were 21.96 and 21.93 for the L1 and L2 groups respectively, which indicated that the two groups' pre-knowledge levels were similar. However, the L1 group showed relatively higher learning ratios in both immediate and delayed test, recording 72.65% and 57.28% respectively.

TABLE 2

Descriptive Statistics: Normalized Scores

	n	*M* (Number of Unknown Words[a])	Immediate Test (%)[b]		Delayed Test (%)[c]	
			M	*SD*	*M*	*SD*
L1	78	21.96	72.65	16.34	57.28	19.70
L2	74	21.93	68.63	15.63	54.20	17.64
Total	152	21.95	70.69	16.08	55.79	18.73

Note. a: V1; b: V2/V1; c: V3/V2*b

Figure 2 and Table 3 show, overall, both L1 and L2 groups' vocabulary learning ratios dropped significantly after a four-week time interval (p = .000 for both groups). However, the two groups' performances in both immediate and delayed tests were not statistically different from each other (p = .124 for the immediate test and p = .313 for the delayed test) as seen in Table 4; the glossing languages influenced neither their immediate learning nor retention of vocabulary significantly.

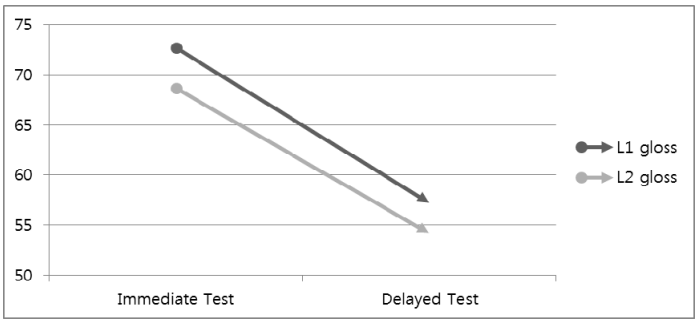

FIGURE 2 Immediate Learning vs. Retention of Vocabulary

TABLE 3

Paired-Samples *t*-Test: Immediate Test vs. Delayed Test

Immediate Test – Delayed Test	Paired Differences			*t*	*df*	*p* (2-tailed)
	M	*SD*	Standard Error Mean			
L1	15.37	8.94	1.01	15.18	77	.000
L2	14.43	7.67	.89	16.18	73	.000

TABLE 4

ANOVA: L1 Gloss vs. L2 Gloss

Immediate Test	Sum of Squares	*df*	Mean Square	*F*	*p*
Between Groups	613.015	1	613.015	2.394	.124
Within Groups	38,416.559	150	256.110		
Total	39,029.574	151			
Delayed Test	Sum of Squares	*df*	Mean Square	*F*	*p*
Between Groups	359.976	1	359.976	1.026	.313
Within Groups	52,625.621	150	350.837		
Total	52,985.597	151			

2. L1 vs. L2: Their Influence on Vocabulary Learning by Proficiency Levels

The glossing language effects were examined for each proficiency group. One overall tendency was that, regardless of the glossing languages, the advanced group outperformed the high intermediate group, which outperformed the intermediate, and then the beginning groups, in both the immediate and delayed tests. Table 5 shows that the CSAT levels influenced vocabulary retention significantly ($F = 17.050$, $p = .000$).

The post hoc results in Table 6 indicate the retention ratios of the advanced and high intermediate groups were higher than those of the intermediate and beginning groups with a statistical significance. The results were congruous with the previous studies which reported students' language competence as an indicator of successful vocabulary learning (Hofer & Yu, 2003; Jenkins, Stein, & Wysocki, 1984; Nemati, 2013; Tekmen & Daloğlu, 2006). Tables 7 and 8 show the breakdown figures of the L1 and L2 groups separately, for the mean values of the immediate and delayed tests according to the participants' CSAT levels.

TABLE 5

ANOVA: Vocabulary Retention by CSAT Levels

	Sum of Squares	*df*	Mean Square	*F*	*p*
Between Groups	7,704.771	3	2,568.257	17.050	.000
Within Groups	21,238.402	141	150.627		
Total	28,943.172	144			

TABLE 6

Post Hoc Tests: Multiple Comparisons of Retention by CSAT Levels (Bonferroni)

(I)SAT Level	(J)SAT Level	Mean Difference (I-J)	Standard Error	*p*	95% Confidence Interval	
					Lower Bound	Upper Bound
Advanced	High Intermediate	6.021	3.011	.285	-2.036	14.078
	Intermediate	15.701	3.127	.000	7.333	24.069
	Beginning	21.327	3.543	.000	11.846	30.808
High Intermediate	Advanced	-6.021	3.011	.285	-14.078	2.036
	Intermediate	9.680	2.508	.001	2.967	16.392
	Beginning	15.306	3.011	.000	7.249	23.363
Intermediate	Advanced	-15.701	3.127	.000	-24.069	-7.333
	High Intermediate	-9.680	2.508	.001	-16.392	-2.967
	Beginning	5.626	3.127	.445	-2.742	13.994
Beginning	Advanced	-21.327	3.543	.000	-30.808	-11.846
	High Intermediate	-15.306	3.011	.000	-23.363	-7.249
	Intermediate	-5.626	3.127	.445	-13.994	2.742

TABLE 7

CSAT Levels & Immediate Vocabulary Learning

CSAT Level	L1			L2		
	n	*M*	*SD*	*n*	*M*	*SD*
Advanced	11	87.07	8.89	13	84.25	9.02
High Intermediate	29	78.29	14.12	25	69.73	13.84
Intermediate	18	66.36	13.01	25	63.93	13.92
Beginning	15	61.20	18.11	9	56.88	16.97
Total	73	73.16	16.55	72	68.73	15.71

TABLE 8

CSAT Levels & Vocabulary Retention

CSAT Level	L1			L2		
	n	*M*	*SD*	*n*	*M*	*SD*
Advanced	11	76.45	14.31	13	74.75	10.02
High Intermediate	29	66.20	15.78	25	56.11	14.49
Intermediate	18	46.56	15.66	25	48.23	14.28
Beginning	15	42.15	17.42	9	38.88	15.98
Total	73	57.96	20.16	72	54.59	17.50

In the L1 vs. L2 comparison, the L1 group of each level outscored their L2 counterpart in both tests overall: not statistically significant for the three levels (advanced: p = .449 & .744 for the immediate and delayed tests respectively; intermediate: p = .560 & .723; and beginning: p = .563 & .645) and significant for the high intermediate level. Table 9 shows the results of the independent samples t-test; within the high intermediate level, the L1 group outperformed the L2 group with statistical significance in both the intermediate (p = .029) and delayed (p = .018) tests.

TABLE 9

t-Tests for Equality of Means, L1 vs. L2: High Intermediate Level Participants (n = 54)

		Levene's Test for Equality of Variances		t-Test for Equality of Means		
		F	p	t	df	p (2-tailed)
Immediate Test	Equal Variances Assumed	.264	.609	2.242	52	.029
	Equal Variances Not Assumed			2.245	51.110	.029
Delayed Test	Equal Variances Assumed	.215	.645	2.432	52	.018
	Equal Variances Not Assumed			2.448	51.771	.018

3. Learner Perceptions of L1 and L2 Glosses: Questionnaire Survey

This section reports on the results of the student questionnaires which asked their perceptions and preferences of the glossing languages, L1 and L2. Q1 (*I understood the text well.*) was to see if any interrelationship exists between the participants' self-perceived reading comprehension and their incidental vocabulary learning. Q2 (*I benefited from the glosses.*) was to determine a more direct relationship between the participants' appreciation of the glosses and their actual vocabulary learning, if any. Q3 (*How did you use the glosses?*), asked them to choose from four options: (1) I focused on the text and referred to the glosses for the unknown words only; (2) I read all the glosses while focusing on the text; (3) I read the glosses for each paragraph first, and then read the paragraph; and (4) I used the glosses in other ways such as _____. Q4 asked them to indicate their preferences between the L1 and L2 glosses and to state reasons in a few words.

The questionnaire item asking the participants' preference of the L1 or L2 glosses was given fourth in order to have them deliberate on their choice through the three previous questions. Table 10 compares the L1 and L2 groups on their preference of the L1 and L2 glosses; Tables 11 and 12 show the statistical significance in the differences between groups. Out of 148 participants who responded to the question, a total of 104 participants

preferred the L1 glosses: 61 (81.33%) and 43 (58.90%) participants of the L1 and L2 groups respectively.

TABLE 10

Participants' Preference of L1-Glossed Text

Glosses Preferred	*n*	Glosses Used	
		L1	L2
L1	104 (70.27%)	61 (81.33%)	43 (58.90%)
L2	44 (29.73%)	14 (18.67%)	30 (41.10%)
Total	148 (100.00%)	75 (100.00%)	73 (100.00%)

TABLE 11

Overall Preference of the Glossing Languages: L1 Group vs. L2 Group

	Value	*df*	Asymptotic Significance (2-sided)	Exact Significance (2-sided)	Exact Significance (1-sided)
Pearson Chi-Square	8.908	1	.003		
Fisher's Exact Test				.004	.002

TABLE 12

Preference of Glossing Languages in High Intermediate Level: L1 Group vs. L2 Group

	Value	*df*	Asymptotic Significance (2-sided)	Exact Significance (2-sided)	Exact Significance (1-sided)
Pearson Chi-Square	5.368	1	.021		
Fisher's Exact Test				.025	.021

The difference in their L1 preference between the two groups was statistically significant (X^2 = 8.908, p = .003). That is, the L2 group's L1 preference was not as strong as the L1 group's. An interesting finding was that the significant difference was mainly caused by the high intermediate level participants: out of 29 participants who used the L1 glosses, 24 (82.75%) favored L1 glosses; and out of 23 who used the L2 glosses, only 10 (43.48%) favored L1 glosses Table 12 shows its statistical significance; the chi-square value was 5.368, significant at .05 level (X^2 = .153, 3.210, & 1.364 for the advanced, intermediate, & beginning levels respectively, all not statistically significant).

As for the optional response about the reasons for their preferences, 74 out of 148 participants made short comments: 40 participants for L1 glosses, 30 for L2 glosses, and 4 for no difference. Those responses for L1 glosses were broadly categorized into three

categories: (1) L2 glosses are difficult and/or involve another unknown L2 word (16 responses); (2) L1 glosses enable an intuitive understanding of a word (10 responses); (3) L2 glosses cut the reading flow (4 responses). Those for L2 glosses were grouped into four categories: (1) L2 glosses are easier (17 responses); (2) L2 glosses improve L2 skills (12 responses); (3) L2 glosses provide L2-like meanings of a word (6 responses); and (4) L2 glosses in L2 reading are more efficient as only one language is involved in the reading (5 responses).

The rest of this section covers the participants' self-perceived reading comprehension in the context of vocabulary learning. The results of the immediate and delayed tests were examined separately to look into the learner perception in reading comprehension relevant to immediate vocabulary learning and/or its retention.

Tables 13 and 14 show that the participants' achievement on the immediate test consistently coincided with the degree of their self-perceived reading comprehension at a statistically significant level: overall $F = 8.413$, $p = .000$; for the L1 group $F = 3.260$, $p = .026$; for the L2 group $F = 7.858$, $p = .000$.

TABLE 13

Self-Perceived Reading Comprehension & Immediate Vocabulary Learning

"I understood the text well."	L1			L2		
	n	*M*	*SD*	*n*	*M*	*SD*
5 (Strongly Agree)	13	84.24	15.08	7	83.13	9.41
4 (Agree)	27	70.96	15.27	30	74.91	12.45
3 (Neutral)	30	70.38	16.29	29	62.81	13.59
2 (Disagree)	7	64.89	15.22	5	56.48	19.06
1 (Strongly Disagree)				3	48.57	17.38
Total	77	72.42	16.33	74	68.63	15.63

TABLE 14

ANOVA: Immediate Vocabulary Learning & Self-Perceived Reading Comprehension ($n = 151$)

		Sum of Squares	*df*	Mean Square	*F*	*p*[a]
Immediate Test	Between Groups (Combined)	7,240.367	4	1,810.092	8.413	.000
SRC	Within Groups	31,414.000	146	215.164		
	Total	38,654.367	150			

Note. a. L1 group ($n = 77$, $F = 3.260$, $p = .026$); L2 group ($n = 74$, $F = 7.858$, $p = .000$)

Tables 15 and 16, in regard to the delayed tests, show that the greater their self-perceived reading comprehension, the higher their performance on the delayed test: overall $F = 8.656$, $p = .000$; for the L1 group $F = 3.331$, $p = .024$; for the L2 group $F = 9.183$, $p = .000$.

TABLE 15

Self-Perceived Reading Comprehension & Vocabulary Retention

"I understood the text well."	L1			L2		
	n	*M*	*SD*	*n*	*M*	*SD*
5 (Strongly Agree)	13	71.45	21.23	7	73.21	12.33
4 (Agree)	27	55.58	16.71	30	61.29	12.64
3 (Neutral)	30	52.83	19.34	29	47.03	16.60
2 (Disagree)	7	51.97	17.45	5	36.18	18.02
1 (Strongly Disagree)				3	38.48	11.73
Total	77	56.86	19.47	74	54.21	17.65

TABLE 16

ANOVA: Vocabulary Retention & Self-Perceived Reading Comprehension ($n = 151$)

		Sum of Squares	*df*	Mean Square	*F*	*p*[a]
Delayed Test	Between Groups (Combined)	9,930.857	4	2,482.714	8.656	.000
SRC	Within Groups	41,876.347	146	286.824		
	Total	51,807.204	150			

Note. a. L1 group ($n = 77$, $F = 3.331$, $p = .024$); L2 group ($n = 74$, $F = 9.183$, $p = .000$)

Interestingly, it turned out that the participants' self-perceived usefulness of the glosses as reflected in their response to Q2 was not meaningfully related to their performances on the immediate and delayed tests. Their method of using the glosses, as inquired by Q3, also did not significantly influence their vocabulary learning.

V. DISCUSSION AND CONCLUSION

This study has examined the different influences of the different glossing languages, the L1 and L2, on the participants' immediate vocabulary learning and retention, employing the new method to measure vocabulary gains, TMR. Further, it has investigated how the participants' perceptions were related with their learning. The results are discussed in the order of the research questions.

1. How did L1 and L2 Glosses Differently Influence L2 Learners' Vocabulary Learning, Overall and/or by Proficiency Level?

Regarding the findings through the immediate and delayed test analysis, the first noticeable point was that the vocabulary attrition phenomenon which has been confirmed in many research studies was also observed in this study for both cases of L1 and L2 glossing. The participants' vocabulary loss in four weeks' period was statistically significant regardless of the glossing language. Lexical knowledge is considered to be more vulnerable to attrition than other aspects of language (Schmitt, 2010). Nelson (1978) indicated that, even after language attrition, there are residues that can be stimulated another time, i.e., savings. According to his study, savings indicate an advantage of relearning old items over learning new items. By means of relearning previously known items, the dormant lexical knowledge will become activated enough to be retrievable. As such, additional treatments after the initial learning, probably within a four-month period, would be necessary for effective and economical vocabulary learning.

Secondly, with reference to the glossing language effects, the difference of the L1 and L2 glosses did not create significant differences overall in neither immediate learning nor retention of the vocabulary. This result concurred with M. H. Ko (2012) and Yoshii (2006). but other studies yielded different results; either the L1 group outperformed (Rouhi & Mohebbi, 2012) or the L2 group outperformed (Laufer & Hill, 2000; Miyasako, 2002).

M. H. Ko (2012) suggested that the different results can be attributed to the participants' L2 proficiency levels; that is the highly proficient learners would perform better with L2 glosses, as was the case of the high level Hong Kong learners in Laufer and Hill (2000), for example. Indeed, the glossing language effects were scrutinized as per the learners' proficiency level in this study. Overall, the L1 glosses tended to elicit better results than the L2 glosses although not statistically significant. However, it was only among the high intermediate participants, which consisted of Levels 2 and 3 of the CSAT, that the L1 condition generated better results than the L2 condition at a statistically significant level. It indicates that in the highest proficiency group, which consisted of the CSAT Level 1 participants, there was no meaningful difference between the L1 and L2 glosses. Taking account of this result, attributing learners' better performance—in either L1 or L2 gloss condition—to a high level proficiency of the language does not seem to be enough explanation. Indeed, one's L1 competence will always be better than or equivalent to her/his L2 competence at least. Then, as a result, the L1 condition should have always produced better performance than the L2 condition in every single research study.

A reasonable explanation for the result was found in Kroll and Tokowicz (2001). They suggested that L2 learners may rely on the counterpart L1 words for the conceptual processing of L2 words more during their early stage of L2 acquisition than after the early

stage. Further, they argue that the challenge L2 learners face is two-fold: first, creating representation for L2 which allows access to meaning; second, developing a control mechanism to modulate the action or influence of L1. The result of this study could be understood to deviate slightly from their claims as the L1 and L2 groups in the lower two levels along with the highest level did not produce significantly different results. That is, the former maintained lower performances regardless of the glossing languages while the latter demonstrated higher performances with both L1 and L2 glosses.

However, it is not implausible to differentiate performance levels and learning stages. Kroll and Tokowicz (2001) emphasized the role of L1 as a temporary scaffold for L2 development that would be later substituted by direct conceptual representation in the L2. The participants in the high intermediate level, which was the second highest in this study, benefited from L1 glosses significantly more than from L2 glosses. A more convincing explanation for this finding would be that they may be in transition to the stage of direct conceptual processing of L2 words, using the temporary scaffold of L1 until they could eventually modulate the L1 factor. At the point, the L1 or L2 gloss condition would not create considerable differences as was the case of the highest level participants in this study.

The third noticeable result was that, as confirmed in many other vocabulary research studies, the higher the learners' proficiency level, the greater their incidental gain, both in the immediate vocabulary learning and retention, consistently without exceptions. Vocabulary was often taken as a painful memorization chore by learners in which only perseverance would bring a success (H. S. Kim & Y. H. Na, 2010). This finding demonstrates the importance of having L2 learners reach a sort of "threshold" level lexical competence as early as possible, at which point their learning can be accelerated.

2. How Did Learners Perceive L1 and L2 Glosses? Are Their Perceptions Related With Their Vocabulary Learning?

A noteworthy point regarding the findings through the questionnaire survey on the learners' perceptions was that while the participants overall preferred the L1 glosses, the degree of preference differed significantly between the L1 and L2 groups. Those who experienced the L2 glosses did not dislike the L2 gloss condition as much as those who experienced the L1 glosses did. This tendency was particularly evident in the high intermediate learners; among those who read the L2-glossed texts, more than half of participants (56.52%) responded that they would prefer L2 glosses. It can be interpreted that while the high intermediate level learners were indeed able to handle both L1 and L2 glosses, they had a preconception that L2 glosses would be more difficult and inconvenient, which actually affected their performances in the immediate and delayed tests. However, their negative preconception toward the L2 glosses seemed to be changed and moderated

by actual experience. Their changed attitude was also found in part of the participants' voluntary comments: *I found L2 glosses easier and more explanatory*; *The word meanings in L2 seem to be more appropriate*; *L2 glosses made me focus on L2*; *Reading L2 glosses was an additional L2 reading practice*; and *My English skills will be improved faster with L2 glosses*, etc. The results could also be taken as an additional signal that they are on their way to the stage of direct processing of L2 as previously discussed.

Another finding that may carry strong implication was that those with higher degrees of self-perceived reading comprehension achieved higher short-term and long-term gains in vocabulary learning. The ANOVA revealed the statistical significance of the result; the values of the five-point scale in the participants' response and their performance scores were directly proportional. It means that the participants who gained more vocabulary through the reading activities considered that their level of text comprehension was higher or vice versa. In any case, this result can be taken as suggesting evidence of the appropriateness of scaffolded reading activities to promote vocabulary learning. The benefit of providing learners with more chances to read L2 texts with enhanced readability by glosses, for an example, was highlighted by the results.

3. Conclusive Remarks

This study intended to compare the effect of L1 and L2 glossed reading activities on vocabulary learning; it rather confirmed that what really matters was the time effect over glossing language effect. Being aware of the importance of relearning in lexical competence, Ansarin and Khojasteh (2013) emphasized that "recycling and revision" (p. 17) activities such as vocabulary games should follow the initial vocabulary learning programs. As Yang and Chen (2007) proposed, employing computer-assisted methods could enhance learning possibilities by making the follow-up activities more manageable.

Returning to the main issue of L1 vs. L2 glosses, the difference in the glossing language was noticeable only in the high intermediate group, who performed better with the L1-glossed texts. That is, the advanced level participants benefited from the glosses regardless of L1 or L2 while the lower level participants did not benefit as much from any glosses to the extent that the glossing language did not create significant differences. Yet, although not statistically significant, the participants performed relatively better with the L1 glosses. Therefore, the overall effectiveness of the L1 cues in L2 learning needs to be acknowledged, particularly for intermediate and beginner level learners. On the basis of the finding of this study, it is recommended to incorporate L2 cues as scaffolding for learners above intermediate level, which could lead them to obtain tools to access concepts of L2 words independently from their L1 (Kroll & Tokowicz, 2001). It is likely that learners prefer L1 glossing, but that may not be because they cannot process L2 cues but rather

because they are not used to taking them. In any case, L2 teachers need to be attentive to learners' status and examine if they are ready to gain benefit from L2-mediated input enhancement in any form.

This study bears limitations. The immediate and delayed tests were all administered at recognition level only, in the form of multiple-choice tests. As Hulstijn (2001) observed, both automatic word recognition and automatic word retrieval are critical in fluent language use. Including retrieval level tests to measure the participants' vocabulary development could have elicited more insightful discussions. Another limitation is about the retention period; the delayed test in this study was given after a four-week period, which appeared to be the most common time gap among previous vocabulary studies. Future research could examine how long the retention period would be in order for a word to be identified as retained, probably by measuring learners' vocabulary knowledge at more than two time intervals. As for the survey questionnaire, the item asking the learners' self-perceived reading comprehension would have been better elaborated into more detailed questions or complemented by oral interviews of the participants to elicit more insightful implications. The factor turned out to be closely related to vocabulary learning, and future study might want to scrutinize this area.

Lastly, as for the methodology, in order to sort out the learners' purer gains from the learning activity in concern, this study employed the new method of tracing a target word one by one. It is hoped that it will be a small contribution for vocabulary research grounded in accurate measurement.

REFERENCES

Ansarin, A. A., & Khojasteh, M. R. B. (2013). Retention in meaning-based vocabulary instruction. *SKY Journal of Linguistics, 26*, 7-19.

Brooks-Lewis, K. A. (2009). Adult learners' perceptions of the incorporation of their L1 in foreign language teaching and learning. *Applied Linguistics, 30*(2), 216-235.

Brown, R., Waring, R., & Donkaewbua, S. (2008). Incidental vocabulary acquisition from reading, reading-while-listening, and listening to stories. *Reading in a Foreign Language, 20*(2), 136-163.

Cashin, W. E., & Downey, R. G. (1992). Using global student rating items for summative evaluation. *Journal of Educational Psychology, 84*(4), 563-572.

Chen, H. (2002, March). *Investigating the effects of L1 and L2 glosses on foreign language reading comprehension and vocabulary retention*. Paper presented at the meeting of the Computer-Assisted Language Instruction Consortium, Davis, CA.

Cook, G. (2003). *Applied linguistics*. Oxford: Oxford University Press.

Cunningham, A. C., & Stanovich, K. E. (2001). What reading does for the mind. *Journal of Direct Instruction, 1*(2), 137-149.

Elley, W. B. (1989). Vocabulary acquisition from listening to stories. *Reading Research Quarterly, 24*(2), 174-187.

Ellis, R., & He, X. (1999). The roles of modified input and output in the incidental acquisition of word meanings. *Studies in Second Language Acquisition, 21*(2), 285-301.

Frishkoff, G. A., Perfetti C. A., & Collins-Thompson K. (2011). Predicting robust vocabulary growth from measures of incremental learning. *Scientific Studies of Reading, 15*(1), 71-91.

Godwin-Jones, R. (2010). Emerging technologies from memory palaces to spacing algorithms: Approaches to second language vocabulary learning. *Language Learning & Technology, 14*(2), 4-11.

Hemmati, T., & Asmawi, A. B. (2015). Incidental vocabulary learning and retention through reading a graded reader among Iranian EFL learners. *The Online Journal of New Horizons in Education, 5*(1), 72-86.

Hofer, B. K., & Yu, S. L. (2003). Teaching self-regulated learning through a "Learning to Learn" course. *Teaching of Psychology, 30*(1), 30-33.

Howatt, A. P. R., & Widdowson, H. G. (2004). *A history of English language teaching* (2nd ed.). Oxford: Oxford University Press.

Huckin, T., & Coady, J. (1999). Incidental vocabulary acquisition in a second language. *Studies in Second Language Acquisition, 21*(2), 181–193.

Hulstijn, J. H. (1993). When do foreign-language readers look up the meaning of unfamiliar words? The influence of task and learner variables. *The Modern Language Journal, 77*(2), 139-147.

Hulstijn, J. H. (2001). Intentional and incidental second language vocabulary learning: A reappraisal of elaboration, rehearsal and automaticity In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 258-286). Cambridge: Cambridge University Press.

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

Hunt, A., & Beglar, D. (2002). Current research and practice in teaching vocabulary. In J. C. Richards & W. A. Renandya (Eds.), *Methodology in language teaching: An anthology of current practice* (pp. 258-266). Cambridge: Cambridge University Press.

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

Jenkins, J. R., Stein, M. L., & Wysocki, K. (1984). Learning vocabulary through reading.

American Educational Research Journal, 21(4), 767-787.

Jeong, Hogyeong, Biswas, G., Johnson, J., & Howard, L. (2010). Analysis of productive learning behaviors in a structured inquiry cycle using hidden Markov models. *Proceedings of the Educational Data Mining 2010, Pittsburgh, USA*, 81-90.

Jordens, R., & Zeter, J. (2013). Reading the world now 3. Seoul: Compass Publishing.

Kim, Hyung-Sun, & Choi, Ung-Yong. (2017). An improved measure of retention in vocabulary learning and its application. *Modern English Education, 18*(1), 1-24.

Kim, Hyung-Sun, & Na, Yoon-Hee. (2010). Vocabulary learning and task-induced involvement. *Korean Journal of Applied Linguistics, 26*(4), 183-211.

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

Kroll, J. F., & Tokowicz, N. (2001). The development of conceptual representation for words in a second language. In J. Nicol (Ed.), *One mind, two languages: Bilingual language processing* (pp. 49-71). Malden, MA: Blackwell.

Laufer, B. (1997). The lexical plight in second language reading: Words you don't know, words you think you know, and words you can't guess. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition: A rationale for pedagogy* (pp. 20-34). Cambridge: Cambridge University Press.

Laufer, B., & Girsai, N. (2008). Form-focused instruction in second language vocabulary learning: A case for contrastive analysis and translation. *Applied Linguistics, 29*(4), 694-716.

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.

Laufer, B., & Hulstijn, J. (2001). Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. *Applied Linguistics, 22*(1), 1-26.

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

Miyasako, N. (2002). Does text-glossing have any effects on incidental vocabulary learning through reading for Japanese senior high school students? *Language Education & Technology, 39*, 1-20.

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

Nation, I. S. P. (2013). *Teaching & learning vocabulary*. Boston: Heinle Cengage Learning.

Nelson, T. O. (1978). Detecting small amounts of information in memory: Savings for nonrecognized items. *Journal of Experimental Psychology: Human Learning and Memory, 4*(5), 453-468.

Nemati, A. (2013). Vocabulary learning strategies: A short way to long term retention.

Linguistic and Literature Studies, 1(1), 8-14.

O'Connell, D. Q., & Dickinson, D. J. (1993). Student ratings of instruction as a function of testing conditions and perceptions of amount learned. *Journal of Research and Development in Education, 27*(1), 18-23.

Roby, W. B. (1999). What's in a gloss? *Language Learning & Technology, 2*(2), 94-101.

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.

Ryan, J., & Harrison, P. (1995). The relationship between individual instructional characteristics and the overall assessment of teaching effectiveness across different instructional contexts. *Research in Higher Education, 36*(5), 577-594.

Schmidt, R. (1994). Implicit learning and the cognitive unconscious: Of artificial grammars and SLA. In N. Ellis (Ed.), *Implicit and explicit learning of language* (pp. 165-210). London: Academic Press.

Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual*. New York: Palgrave Macmillan.

Tekmen, E. A. F., & Daloğlu A. (2006). An investigation of incidental vocabulary acquisition in relation to learner proficiency level and word frequency. *Foreign Language Annals, 39*(2), 220-243.

Verhoeven, L., & Perfetti, C. A. (2011). Introduction to this special issue: Vocabulary growth and reading skill. *Scientific Studies of Reading, 15*(1), 1-7.

Yang, S. C., & Chen, Y. J. (2007). Technology-enhanced language learning: A case study. *Computers in Human Behavior, 23*(1), 860-879.

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

Yoshii M., & Flaitz, J. (2002). Second language incidental vocabulary retention: The effect of text and picture annotation types. *CALICO Journal, 20*(1), 33-58.

Zhang, B., & Li, C. (2011). Classification of L2 vocabulary learning strategies: Evidence from exploratory and confirmatory factor analyses. *RELC Journal, 42*(2), 141-154.

APPENDIX

The Pool of Target Words

1. inhibit	7. fertile	13. artisan	19. deceased	25. worldly
2. pursuit	8. dike	14.hieroglyphic	20. rectangular	26. cast
3. nomadic	9. irrigation	15. elaborate	21.grid	27. vessel
4. crop	10. canal	16. embalm	22. inscription	28. weave
5. domesticate	11. cuneiform	17. corpse	23. utensil	29. merge
6. emerge	12. wedge-shaped	18. gigantic	24. oracle	30. diverge

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

Hyung-Sun Kim
Division of Undeclared Majors
Chosun University
309 Pilmun-daero, Dong-gu, Gwangju 61452
Tel: 062-230-6948
Email: hskim4@chosun.ac.kr

Ung-Yong Choi
College of Business Administration
Chonnam National University
77 Yongbong-ro, Buk-gu, Gwangju 61186
Tel: 062-530-1476
Email: uychoi@jnu.ac.kr

Received 21 September 2017
Reviewed 24 October 2017
Accepted 11 November 2017