The Effectiveness of the Keyword Method and the Context Method on L2 Vocabulary Learning

Hyun-Jeong Nam

Daegu Haany University

Nam, Hyun-Jeong. (2015). The effectiveness of the keyword method and the context method on L2 vocabulary learning. *Modern English Education*, *16*(3), 99-116.

Keyword method (KWM) has received significant attention over the past few decades to explore the memory and retrieval of second language vocabulary. A great deal of research has yielded valuable insights into the KWM; however, a consensus has not been reached on its effectiveness on L2 vocabulary memory. This study explored the effectiveness of the keyword method in comparison with the context method. In addition, in order to answer the reason for inconsistent results in the seemingly similar studies in the past, the present study included four learning-test conditions (1. KWM - word definition recall 2. KWM - sentence completion in L2 3. context method - word definition recall 4. context method - sentence completion in L2). The results show that there was no statistically significant difference in the efficacy of the KWM and the context method in the immediate and the delayed recall. There was, however, a statistically significant difference among the different learning—test conditions, observed both in the immediate and delayed recall test. The recall was more effective in the conditions in which the relation between the learning method and the test type was consistent. The findings suggest implications for the design of future empirical research.

[keyword method/context method/test task type/
핵심어기법/문맥활용/실험과제유형]

I. INTRODUCTION

Vocabulary has long been a prominent focus of pedagogy and research in second language acquisition. Along with various teaching methods and techniques developed and applied to second/foreign language teaching for the last few decades, there has been a great

deal of research concerning vocabulary learning and retention.

   Among the research exploring memory and retrieval of second language vocabulary, keyword method (hereafter KWM) (Atkinson, 1975; Atkinson & Raugh, 1975) has received significant attention over the past decades. Many studies have examined the effectiveness of the KWM in bilingual memory; however, conflicting results have been obtained in seemingly similar studies. Some researchers are convinced that the KWM has a positive effect on memory and vocabulary retention (Baleghizadeh & Ashoori, 2010; Beaton, Gruneberg & Ellis, 1995; Beaton et al., 2005; Dolean, 2014; Kasper, 1993; Sagarra & Alba, 2006; Shapiro & Waters, 2005; Zhanga & Schummb, 2000), while others have yielded opposing results (Campos, González & Amor, 2003; Wei, 2015; Wyra, Lawson & Hungi, 2007).

   Before the discussion of the effectiveness of the KWM in bilingual memory, one question that needs to be asked is why seemingly similar studies have produced inconsistent results. Given that experimental variation may induce different paths of lexical access and thus may differentially reflect bilingual memory organization (e.g., de Groot, 1995; Kroll, Bobb & Wodniecka, 2006), this possibility should therefore warrant further consideration. Thus the present study will also take account of the relation between learning (treatment) and test (task) type.

II. LITERATURE REVIEW

   The keyword method was first developed as a two-stage procedure (Atkinson & Raugh, 1975). That is, the L2 learner's L1 which has an acoustical similarity to the target L2 word, is chosen and then the L1-based keyword is linked to the target L2 word in the learner's mental image. Later, Hulstijn (1997) introduced the KWM in three stages; keyword selection, the link between the keyword and the target L2, and the creation of a visual image. Compared with the conventional implementation, Hulstijn's (1997) seems more flexible and practical in that the L2 learner can use both L1 and L2 which are similar to the target L2 word not only acoustically but also orthographically (p. 204). He further claims that "[i]f possible, keywords should be taken not from L1 but from L2 vocabulary with which the L2 learner is already familiar (Hulstijn, 1997, p.205).

   There have been many studies in favor of efficacy of the KWM (Baleghizadeh & Ashoori, 2010; Beaton, Gruneberg & Ellis, 1995; Beaton et al., 2005; Dolean, 2014; Kasper, 1993; Sagarra & Alba, 2006; Shapiro & Waters, 2005; Zhanga & Schummb, 2000). It is also suggested that visual image has a beneficial effect on the KWM. For example, Shapiro and Waters (2005) observed that the KWM worked better for high imagery words. Moreover, Dolean (2014) obtained positive results even from the visual image which was

unrelated to the association between the keyword and the target word.

On the other hand, the effectiveness of KWM has been challenged by contradictory evidence (Campos, González & Amor, 2003; Wei, 2015; Wyra, Lawson & Hungi, 2007). Most recently, Wei (2015) compared three types of methods; the word part technique, the KWM, and the self-strategy learning. It was found in the translation test that the KWM was not as effective as the word part technique and the self-strategy learning. In addition, Campos, González and Amor (2003) observed that the rote rehearsal method was superior to KWM in bilingual memory.

It should be noted, however, that many studies seem to compare KWM with rather conventional methods or no strategy method. For example, KWM was compared with the memorization of L1 definition in H. Kim and N. Lee (2002); with the word list method in Baleghizadeh and Ashoori (2010); with the rehearsal method in Zhanga and Schummb (2000); and with no strategy method (Avila & Sadoski, 1996). Not many studies of KWM have included its comparison with the contextual method (Levin, Levin, Glasman & Nordwall, 1992). Wang and Thomas (1995) found the promising evidence supporting the efficacy of the context method over KWM in the long term memory. Given that context may shape the "meaning potential" and the "actual meaning" (Fauconnier, 1997, p.37) as meaning construction comes from the contextual relations (Cruse, 1986, p.16), the importance of contextual cues in L2 vocabulary learning (Saeed, 2003) should also be considered in the study of bilingual memory.

While a great deal of research has yielded valuable insights into the KWM, a consensus has not been reached on the KWM's effectiveness on L2 vocabulary memory. The reason may possibly be, as suggested by some researchers (Beaton et al., 2005; Hall, 1988; Wyra et al., 2007), the inconsistency of the study designs. That is, the words utilized in the various experiments may have different properties and further, the treatment and test types may vary in those studies.

It should be noted that different type of tests may extract the lexical knowledge obtained in the treatment differently. The effectiveness of the learning method used in the treatment of the study may manifest itself best when learners are tested in the same way as they learned. The study of J. Park and Y. Jung (2013) may be a good example of the direct correlation between the learning method used in the treatment and the task requirement in the test. In their study, the learners in the definition group learned the L1 translation of the target L2 word while participants in the context group were trained to translate L2 example sentences in context into L1. The results that the definition recall test was more successful than the sentence-cued recall test, however, may be interpreted differently if the relation between the treatment and the test is considered. That is, it is possible to assume that the treatment for the context group focusing on translation into L1 may not adequately contribute to the context-based learning effect in the L2 sentence completion test. In a

same vein, Wyra et al. (2007) found that providing the participants with retrieval training lead to successful recall in the test. Unfortunately, little attention has been given to research pertaining to this issue in Korea. In order to answer the reason for inconsistent results in the seemingly similar studies in the past, the present study included four learning-test conditions. The research questions are addressed as follows:

1) Which learning method is more effective; the KWM or the context method?
2) Are there any changes regarding the effectiveness of the learning method from the immediate recall to the delayed recall?
3) Does the consistency of the learning method and the test type affect the results?

III. METHOD

1. Design

In the experiment of Rodríguez and Sadoski's (2000), most relevant to the present study comparing the KWM and the context method, the participants were provided with the L1-associated keywords which were acoustically similar to the target L2 words. However, in the present study, as suggested by Hulstijn (1997, p.205), the participants were allowed the leeway to build their own mental image using both acoustical similarity to L1 and semantic relatedness to other L2 words they were already familiar with. These customized keywords may be more effective than the keyword made by the researchers and uniformly applied to all the participants. In the context method, L2 definitions were also included as a form of a sentence in the examples in order for the learners to better grasp the meaning in the context.

Different from the study of Rodríguez and Sadoski's (2000) in which the recall test exclusively required the participants to produce the L1 translation of the target L2, the present study included both directions: L1-L2 and L2-L1. The detailed design of the study is shown in Figure 1.

<div style="border:1px solid black; padding:10px;">

Condition 1: KWM -Word definition recall (L1-L2 and L2-L1)

Learning: KWM
    empennage- (비행기·비행선의) 꼬리 날개
Word definition recall: L1-L2
    (비행기·비행선의) 꼬리 날개 : _____

Learning: KWM
    Manchet - 최고 품질의 밀가루로 만든 흰 빵

</div>

Word definition recall: L2-L1
    Manchet : _____

Condition 2: KWM-Sentence completion in L2

Learning: KWM
    empennage- (비행기·비행선의) 꼬리 날개

Sentence completion in L2
    _____ is the rear part of an aircraft
    After the plane crash, the only part that remained intact was the ____.
    The plane _____looks like a fish tail

Condition 3: Context method-Word definition recall (L1-L2 and L2-L1)

Learning: context method
    -Maggot is a small soft creature with no arms or legs that later changes into a fly and are found in old meat and dead bodies.
    -The rotten meat is full of maggots.
    -The maggot therapy may seem gross and disgusting, but it could save patients from pain.

Word definition recall: L1-L2
    구더기 : _____

Learning: context method
    -Cordite is a chemical that can explode, used in weapons
    -After each fire mission, I can still smell the odor of cordite.
    -Like gunpowder, cordite is classified as a low explosive because of its slow burning rates

Word definition recall: L2- L1
    Cordite : _____

Condition 4: Context method - Sentence completion in L2

Learning: context method
    -Maggot is a small soft creature with no arms or legs that later changes into a fly and are found in old meat and dead bodies.
    -The rotten meat is full of maggots.
    -The maggot therapy may seem gross and disgusting, but it could save patients from pain.

Sentence completion in L2

> *- _____ is a small soft creature with no arms or legs that later changes into a fly and found in old meat and dead bodies.*
> *-The rotten meat is full of _____.*
> *-The _____ therapy may seem gross and disgusting, but it could save patients from pain.*

FIGURE 1 Learning Method -Test Type

2. Participants

A total of 44 participants (20 subjects in the KWM group and 24 subjects in the context group) were involved in the present study. The subjects were Korean male and female freshmen from different majors taking a general English course at a local university in Korea. They were randomly assigned to either group.

3. Materials and Procedure

The preliminary study was conducted 4 weeks before the experiment in order to obtain the refined target L2 word list. The participants were provided with a list of candidate words which was adopted from Rodríguez and Sadoski (2000) and Nation′s (2001) vocabulary level test (the 10,000 word level). Then they were asked to mark the words they already knew. After the exclusion of the words that they previously learned, the final list of target L2 words, a total of 16 words (8 nouns and 8 verbs, see Appendix) were obtained for the study.

In the learning session, the participants in the KWM group were provided with a list of target L2 words with the L1 meaning (definition) and asked to provide their own key word for each target L2 word. They were allowed to create their own mental image using acoustically similar L1 and/or semantically related L2 word they were already familiar with. The participants in the context group learned three example sentences for each target L2 word. The both groups were given 30 minutes.

The test session consisted of two recall tests. The immediate recall was conducted after the learning session was complete and there was a one week interval between the immediate and the delayed recall. Both groups were tested both in the word definition recall (translating L1 into L2 /L2 into L1 at a word level) and in the sentence completion (producing the target L2 word) consecutively, regardless of the learning method in the treatment. In order to avoid any possible undue repetition effect from the first test which may induce better results in the second test, the test type consistent with the learning method was implemented first for both groups.

4. Data Collection and Analysis

The responses were collected manually and organized using Microsoft Excel program according to the types of the learning method and the test. Responses with minor spelling errors were included in the counting.

For analysis, the data were fed to the statistics program SPSS 22. First, descriptive statistics were used for general information such as Mean and Std. Second, Independent-samples T-test was used to examine the difference between the KWM and the context method in the test (research question 1) and also for the difference between the immediate recall and the delayed recall (research question 2). Third, a one-way between-groups analysis of variance (ANOVA) was used to explore the difference among the different learning—test types in the immediate recall and the delayed recall (research question 3).

IV. RESULTS

As shown in Table 1 and Table 2, the total mean scores of recall both from the word definition recall and the sentence completion were compared to examine any difference between the KWM and the context method in the immediate recall test (research question 1).

TABLE 1

KWM vs. Context Method in Immediate Recall

	Learning type	N	Mean	Std. deviation	Std. error
Immediate recall	KWM	40	13.85	2.74	.43
	Context method	48	14.52	2.75	.40

An independent-samples t-test was conducted to compare the learning methods. There was no significant difference in the recall for KWM (M=13.85, SD=2.74) and context method (M=14.52, SD=2.75; t (86)=-1.14, p=.26, two-tailed).

TABLE 2

Independent-Samples T-Test

		Levene's test for equality of variances		T-test for equality of means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean difference	Std. error difference	95% confidence interval of the difference	
									Lower	Upper
Immediate recall	Equal variance assumed	.66	.42	-1.14	86.00	.26	-.67	.59	-1.84	.50
	Equal variance not assumed			-1.14	83.28	.26	-.67	.59	-1.84	.50

In Table 3 and Table 4, the total mean scores of recall both from the word definition recall and the sentence completion in the delayed recall are compared based on the learning type.

TABLE 3

KWM vs. Context Method in Delayed Recall

	Learning type	N	Mean	Std. deviation	Std. error
Delayed recall	KWM	40	8.53	3.62	.57
	Context method	48	9.13	5.08	.73

An independent-samples *t*-test was conducted to compare the learning methods in the delayed recall. There was no significant difference in recall for KWM (M=8.53, D=3.62) and context method (M=9.13, SD=5.08; *t* (84.12)=-.65, *p*=.52, two-tailed).

TABLE 4

Independent-Samples T-Test

		Levene's test for equality of variances		T-test for equality of means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean difference	Std. error difference	95% confidence interval of the difference	
									Lower	Upper
Delayed recall	Equal variance assumed	8.59	.00	-.63	86.00	.53	-.60	.96	-2.51	1.31
	Equal variance not assumed			-.65	84.12	.52	-.60	.93	-2.45	1.25

Any changes of the effectiveness of the learning method from the immediate recall to the delayed recall were not observed (research question 2). Similar to the immediate recall test, there was no statistically significant difference in the learning methods in the delayed recall. This implies that there was no significant change in superiority of one learning method over the other from the immediate recall to the delayed recall.

Table 5, 6, 7, and 8 show whether the consistency of the learning method and the test type contributes to the better results (research question 3) in the immediate recall.

TABLE 5

Learning Method—Test Type: Immediate Recall

	N	Mean	Std. deviation	Std. error	95% confidence interval for mean		Minimum	Maximum
					Lower bound	Upper bound		
1 (KWM-WDR)	20.00	15.25	1.65	0.37	14.48	16.02	9.00	16.00
2 (KWM-SCL)	20.00	12.45	2.93	0.65	11.08	13.82	8.00	16.00
3 (CON-WDR)	24.00	14.25	2.91	0.59	13.02	15.48	6.00	16.00
4 (CON-SCL)	24.00	14.79	2.62	0.53	13.69	15.90	5.00	16.00
Total	88.00	14.22	2.75	0.29	13.63	14.80	5.00	16.00

TABLE 6

ANOVA

	Sum of squares	df	Mean square	F	Sig.
Between groups	91.74	3.00	30.58	4.53	0.01
Within groups	567.16	84.00	6.75		
Total	658.90	87.00			

TABLE 7

Robust Tests of Equality of Means

	Statistic [a]	df1	df2	Sig.
Welch	4.64	3	45.53	.01
Brown-Forsythe	4.62	3	74.91	.01

a. Asymptotically F distributed.

108

Hyun-Jeong Nam

TABLE 8

Tukey HSD Post Hoc Tests: Multiple Comparisons

(I) Learning method-test type	(J) Learning method-test type	Mean difference (I-J)	Std.error	Sig.	95% confidence interval	
					Lower bound	Upper bound
1 (KWM-WDR)	2 (KWM-SCL)	2.80*	.82	.01	.65	4.95
	3 (CON-WDR)	1.00	.79	.58		3.06
	4 (CON-SCL)	.46	.79	.94		2.52
2 (KWM-SCL)	1 (KWM-WDR)	-2.80*	.82	.01		
	3 (CON-WDR)	-1.80	.79	.11		.26
	4 (CON-SCL)	-2.34*	.79	.02		
3 (CON-WDR)	1 (KWM-WDR)	-1.00	.79	.58		1.06
	2 (KWM-SCL)	1.80	.79	.11		3.86
	4 (CON-SCL)	-.541	.75	.89		1.42
4 (CON-SCL)	1 (KWM-WDR)	-.46	.79	.94		1.60
	2 (KWM-SCL)	2.34*	.79	.02	.28	4.40
	3 (CON-WDR)	.54	.75	.89		2.51

Note. *. The mean difference is significant at the .05 level.

Note. KWM: keyword method, CON: Context method, WDR: word definition recall, SCL: sentence completion in L2

A one-way between-groups analysis of variance was conducted to explore the difference among the different conditions (learning method—test type) in the immediate recall. The assumption of homogeneity of variance was violated; therefore, the Welch F-ratio is reported. There was a statistically significant difference at the $p<.05$ level in the immediate recall for the four learning—test conditions: F (3, 45.53) =4.64, p=.01. The effect size was .14[1], which indicates a large effect. Post-hoc comparisons using the Tukey HSD test indicated that the mean score for Condition 1 (*M*= 15.25, *SD*=1.65) was significantly different from Condition 2 (*M*=12.45, SD=2.93) and Condition 2 (*M*=12.45, *SD*=2.93) was significantly different from Condition 4 (*M*=14.79, *SD*=2.62). Condition 3 did not differ significantly from either Condition 1, 2 or 4.

---

[1] Eta squared = Sum of squares between groups / Total sum of squares

In Table 9, 10, and 11 the effect of the consistency of the learning method and the test type on better results was examined in the delayed recall.

TABLE 9

Learning Method—Test Type: Delayed Recall

	N	Mean	Std. deviation	Std. error	95% confidence interval for mean		Minimum	Maximum
					Lower bound	Upper bound		
1 (KWM-WDR)	20	9.95	3.24	.72	8.44	11.46	4.00	14.00
2 (KWM-SCL)	20	7.10	3.49	.78	5.47	8.73	1.00	14.00
3 (CON-WDR)	24	12.17	4.02	.82	10.47	13.86	3.00	16.00
4 (CON-SCL)	24	6.08	4.16	.85	4.33	7.84	1.00	16.00
Total	88	8.85	4.46	.48	7.91	9.80	1.00	16.00

TABLE 10

ANOVA

	Sum of squares	df	Mean square	F	Sig.
Between groups	533.16	3.00	177.72	12.44	.00
Within groups	1199.92	84.00	14.29		
Total	1733.08	87.00			

TABLE 11
Tukey HSD Post Hoc Tests: Multiple Comparison

(I) Learning-test	(J) Learning-test	Mean difference(I-J)	Std.error	Sig.	95% confidence interval	
					Lower bound	Upper bound
1 (KWM-WDR)	2 (KWM-SCL)	2.85	1.20	.09	-.28	5.98
	3 (CON-WDR)	-2.22	1.14	.22	-5.22	.78
	4 (CON-SCL)	3.87*	1.14	.01	.87	6.87
2 (KWM-SCL)	1 (KWM-WDR)	-2.85	1.20	.09	-5.98	.28
	3 (CON-WDR)	-5.07*	1.14	.00	-8.07	-2.07
	4 (CON-SCL)	1.02	1.14	.81	-1.98	4.02
3 (CON-WDR)	1 (KWM-WDR)	2.22	1.14	.22	-.78	5.22
	2 (KWM-SCL)	5.07*	1.14	.00	2.07	8.07
	4 (CON-SCL)	6.08*	1.09	.00	3.22	8.94
4 (CON-SCL)	1 (KWM-WDR)	-3.87*	1.14	.01	-6.87	-.87
	2 (KWM-SCL)	-1.02	1.14	.81	-4.02	1.98
	3 (CON-WDR)	-6.08*	1.09	.00	-8.94	-3.22

Note. *. The mean difference is significant at the .05 level.

A one-way between-groups analysis of variance was conducted to examine the difference among the different conditions (learning method—test type) in the delayed recall. There was a statistically significant difference at the $p<.05$ level in the delayed recall for the four learning—test conditions: F (3, 84) =12.44, p=.00. The effect size was .31[2], which means a large effect. Post-hoc comparisons using the Tukey HSD test indicated that the mean score for Condition 1 (M=9.95, SD=3.24) was significantly different from Condition 4 (M=6.08, SD=4.16), Condition 2 (M=7.10, SD=3.49) was significantly different from Condition 3 (M=12.17, SD=4.02), and Condition 3 (M=12.17, SD=4.02) was significantly different from Condition 4 (M=6.08, SD=4.16).

---

[2] Eta squared = Sum of squares between groups / Total sum of squares

V. DISCUSSION

The results concerning research question one suggest that there was no statistically significant difference in the efficacy of the KWM (*M*=13.85, *SD*=2.74) and the context method *M*=14.52, *SD*=2.75; *t*(86)=-1.14, *p*=.26, two-tailed) in the immediate recall. Although there was a decline in the recall in the delayed test KWM (*M*=13.85 to *M*=8.53) and context method (*M*=14.52 to *M*=9.13), similar results were obtained in the delayed recall. Thus, research question two can be answered; there was no significant change in superiority of one learning method over the other from the immediate recall to the delayed recall. Given that there have been mixed results from the studies either in favor of or in opposition to these methods, these findings should be discussed in depth in consideration of the relation between the learning method and the test type as follows.

Some evidence was found to answer research question three regarding the influence of the consistency of the learning method and the test type on the results. The present study included four learning-test conditions (1. KWM - word definition recall 2. KWM -sentence completion in L2 3. context method - word definition recall 4. context method- sentence completion in L2). Given that the word definition recall test reflects the KWM and the sentence completion in L2 reflects the context method, the learning method - test type is consistent in condition 1 and condition 4. The results show that there was a statistically significant difference in the extent of recall among the different conditions, and the difference was observed both in the immediate recall and the delayed recall test.

In the immediate recall, the effect of the consistency of the learning method and the test type was evident. The mean of the recall (*M*=15.25) was the highest in condition 1 (KWM-WDR) followed by the condition 4 (CON-SCL) (*M*=14.79), both of which have consistency in the relation between the learning method and the test type.

In the delayed recall, the results of the consistency of the learning method and the test type seem rather complex. The mean of the recall (*M*=12.17) was the highest in condition 3 (CON-WDR) followed by the condition 1 (KWM-WDR) (*M*=9.95), only one of which have consistency in the relation between the learning method and the test type.

The reason that the seemingly inconsistent condition 3 (CON-WDR) yielded the best results may be explained in recognition of L1 mediation in the learning. That is, the trace of L1 translation was found on the context method learning sheets which were retrieved after the learning session. It revealed that the majority of the participants in the context method group (22 out of 24), although requested to use the L2 exclusively, in fact did attempt to study the target L2 word with their own L1 translation as well as the given L2 example sentences. Given that L1 lemma influence seems inevitable for Korean L2 learners in L2 vocabulary acquisition (J. Jeon, 2013; H. Nam, 2011), their learning from both L1 and L2 may have contributed to the best recall in condition 3.

The reason that the most effective condition 1 (KWM-WDR) in the immediate recall turned to the second best in the delayed recall can be explained as follows. First, it may be possible that the association between the keyword and the target word became weak in the delayed recall and thus could not be as effective as in the immediate recall. This can be supported by many researchers who questioned the effect of KWM on long-term memory (Rodríguez & Sadoski, 2000; Wang & Thomas, 1995; Wieczynski & Blick, 1996). Second, different from many studies requiring the participants to produce L1 translation, the present study included both directions (L1-L2 and L2-L1) in the word definition recall. When the participants were asked to produce L2 word based on the L1 meaning, it is possible that the keyword triggered a part of the target L2 word, leading to the failure of the complete recall in the test. Some examples of this were obtained from some participants' answer sheets. By tracking the participant's learning sheet it was found that the participant memorized the target L2 word 'maggot' by creating the mental image "sticking something in the maggot" using the keyword "(구더기에) 막 꽂아", but in the test which required him to produce the target L2 word, the keyword triggered a segment of the target L2 word and in turn the wrong word "makkot" was produced.

VI. CONCLUSION AND PEDAGOGICAL IMPLICATIONS

As suggested by Kasper (1993), it may be necessary that KWM be promoted in foreign language classes. It should also be noted, however, that L2 vocabulary learning should not exclusively rely on L1 translation equivalents without context (Hulstijn, 1997, p.214). Given that contextual information involving connotations and pragmatic appropriateness should also be stored in the lexical entries forming the semantic networks in L2 lexicon (Aitchison, 2003; Clark, 1993; Levelt, 1989; Miller & Fellbaum, 1991), the effect of context method (H. An & C. Min, 2014; Wang & Thomas, 1995) should also be implemented in English classes in Korea. L2 vocabulary learning integrating KWM and the context method may be most effective when the two methods complement each other (Pressley, Levin, & McDaniel, 1987; Rodríguez & Sadoski, 2000).

The findings that the L2 vocabulary recall was more effective in the consistent learning—test condition will have great implications for the design of the future empirical research. That is, the most refined results should be obtained when the test best reflects the treatment in the experiment. This can be explained in terms of the concept of "network training" (Plunkett & Marchman, 1993, p. 21; Dell, 2000, p. 345) in cognitive linguistics. The node in the network is built through the learning and it also becomes more activated by the repetition, which in turn makes the retrieval easier. Therefore, the target L2 word should be retrieved more effectively when the same network trained in the learning is

utilized in the test. This is in line with H. Nam (2015) that the more the L1 was utilized to memorize the target English idioms, the more idioms were produced with L1 primes than L2.

In conclusion, there was no statistically significant difference in the efficacy of the KWM and the context method in the immediate recall and the delayed recall. There was, however, a statistically significant difference among the different learning-test conditions, observed both in the immediate and delayed recall test. The recall was more effective in the conditions in which the relation between the learning method and the test type was consistent.

The findings of this study should be taken as suggestive rather than definitive due to the limitation of the study. Based on the class observations the participants in the study were highly motivated L2 learners, and thus their recall was considerably high especially in the immediate recall. The future studies may yield different results based on the learners' motivation, working memory, and other variables.

REFERENCES

Aitchison, J. (2003). *Words in the Mind: an introduction to the mental lexicon* (3rd ed.).Oxford: Blackwell.

An, Hyo-Seon, & Min, Chan-Kyoo. (2014). The effects of context variability on vocabulary learning. *Studies in English Education*, *19*(2), 127-149.

Atkinson, R. C. (1975). Mnemotechnics in second-language learning. *American Psychologist, 30,* 821-828.

Atkinson, R. C., & Raugh, M. R. (1975). An application of the mnemonic keyword method to the acquisition of Russian vocabulary. *Journal of Experimental Psychology: Human Learning and Memory, 104*, 126-133.

Avila, E., & Sadoski, M. (1996). Exploring new applications of the keyword method to acquire English vocabulary. *Language Learning, 46*, 379-395.

Baleghizadeh, S., & Ashoori, A. (2010). The effect of keyword and word list methods on immediate vocabulary retention of EFL learners. *Pakistan Journal of Social Sciences, 30*(2), 251-261.

Beaton, A., Gruneberg, M., & Ellis, N. (1995). Second retention of foreign vocabulary learned using the keyword method: a ten-year follow-up. *Language Research, 11* (2), 112-120.

Beaton, A., Gruneberg, M., Hyde, C., Shufflebottom, A., & Sykes, R. (2005). Facilitation of receptive and productive foreign vocabulary learning using the keyword method: the role of image quality. *Memory*, *13*(5), 458-471.

Campos, A., González, M. A., & Amor, A. (2003). Limitations of the mnemonic-keyword method. *The Journal of General Psychology*, *130*(4), 399-413.

Clark, E. V. (1993). *The lexicon in acquisition*. Cambridge: Cambridge University Press.

Cruse, D. A. (1986). *Lexical semantics.* Cambridge: Cambridge University Press.

De Groot, A. M. B. (1995). Determinants of bilingual lexicosemantic organization. *Computer Assisted Language Learning*, *8*, 151-180.

Dell, G. S. (2000). Commentary: counting, connectionism, and lexical representation. In M. B. Broe & J. B. Pierrehumbert (Eds.), *Papers in laboratory phonology V* (pp. 335-348). Cambridge: Cambridge University Press.

Dolean, D. D. (2014). Using the keyword method in the classroom: Is the interacting imagery necessary? *System*, *45*, 17-26.

Fauconnier, G. (1997). *Mappings in thought and language*. Cambridge: Cambridge University Press.

Hall, J. W. (1988). On the utility of the keyword mnemonic for vocabulary learning. *Journal of Educational Psychology*, *80*, 554-562.

Hulstijn, J. H. (1997). Mnemonic methods in foreign language vocabulary learning: Theoretical considerations and pedagogical implications. In J. Coady (Ed.), *Second language vocabulary acquisition* (pp. 203-224). Cambridge: Cambridge University Press.

Jeon, Ji-hyun. (2013). A study of the relationship between L1 lemma influence and L2 vocabulary acquisition. *English Language Teaching*, *25*(1), 43-62.

Kasper, L. F. (1993). The keyword method and foreign language vocabulary learning: A rationale for its use. *Foreign Language Annals*, *26*, 244-251.

Kim, Heon-Soo, & Lee, Nan. (2002). The effects of key-word method and concept-presentation method on Japanese vocabulary learning. *The Korean Journal of Educational Psychology*, *16*(4), 353-371.

Kroll, J. F., Bobb, S., & Wodniecka, Z. (2006). Language selectivity is the exception, not the rule: Arguments against a fixed locus of language selection in bilingual speech. *Bilingualism: Language and Cognition*, *9*(2), 119-135.

Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge: MIT Press.

Levin, J. R., Levin, M. E., Glasman, L. D., & Nordwall, M. B. (1992). Mnemonics vocabulary instruction: Additional effectiveness evidence. *Contemporary Educational Psychology*, *17*(2), 156-174.

Mang, Eun-Kyung. (2001). The effects of motivation on language learning. *Modern English Education, 2*(2), 157-170.

Miller, G. A., & Fellbaum, C. (1991). Lexical and conceptual semantics. In B. Levin & S. Pinker (Eds.), *Lexical & conceptual semantics* (pp. 197-229). Cambridge, MA: Blackwell Publishers Inc.

Nam, Hyun-Jeong. (2011). Konglish phenomenon: L1 activation in L2. *English Teaching*, *66*(4), 191-211.

Nam, Hyun-Jeong. (2015). Idiom processing: What primes idioms? *The Journal of Linguistic Science*, *72*, 153-170.

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

Park, Jung-Hwa, & Jung, Yang-Soo. (2013). The effects of different types of word lists and tasks on Korean college students' English vocabulary learning. *The Journal of Studies in Language, 29*(2), 245-267.

Plunkett, K., & Marchman, V. (1993). From rote learning to system building: acquiring verb morphology in children and connectionist nets. *Cognition*, *48*, 21-69.

Pressley, M., Levin, J. R., & McDaniel, M. A. (1987). Remembering versus inferring what a word means: Mnemonic and contextual approaches. In M. G. McKeown & M. E. Curtis (Eds.), *The nature of vocabulary acquisition* (pp. 107-127). Hillsdale, NJ: Lawrence Erlbaum Associates.

Rodríguez, M., & Sadoski, M. (2000). Effects of rote, context, keyword, and context/keyword methods on retention of vocabulary in EFL classrooms. *Language Learning*, *50*(2), 385-412.

Saeed, J. I. (2003). *Semantics* (2nd ed.). Oxford: Blakwell.

Sagarra, N., & Alba, M. (2006). The key is in the keyword: L2 vocabulary learning methods with beginning learners of Spanish. *The Modern Language Journal*, *90*(2), 228-243.

Shapiro, A. M., & Waters, D. L. (2005). An investigation of the cognitive processes underlying the keyword method of foreign vocabulary learning. *Language Teaching Research*, *9*(2), 129-146.

Wang, A. Y., & Thomas, M. H. (1995). Effect of keywords on long-term retention: Help or hindrance? *Journal of Educational Psychology, 87,* 468-475.

Wei, Z. (2015). Does teaching mnemonics for vocabulary learning make a difference? Putting the keyword method and the word part technique to the test. *Language Teaching Research*, *19*(1), 43-69.

Wieczynski, D. M., & Blick, K. A. (1996). Self-referencing versus the keyword method in learning vocabulary words. *Psychological Reports*, *79*, 1391-1394.

Wyra, M., Lawson, M. J., & Hungi, N. (2007). The mnemonic keyword method: The effects of bidirectional retrieval training and of ability to image on foreign language vocabulary recall. *Learning and Instruction, 17*, 360-371.

Zhanga, Z., & Schummb, J. S. (2000). Exploring effects of the keyword method on limited English proficient students' vocabulary recall and comprehension. *Reading Research and Instruction*, *39*(3), 202-221.

APPENDIX
List of Words Used in the Experiments

1. empennage
2. Maggot
3. goblet
4. skillet
5. Squirm Dissipate
6. Loot Rasp
7. Manchet
8. Cordite
9. Rubble
10. Barracks
11. Scrawl
12. Dabble
13. Pacify
14. Cringe

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

Hyun-Jeong Nam
Department of English & TESOL
Daegu Haany University
Haany daero1, GyeongSan-shi, Gyeongsangbuk-Do, 712-715
Tel:(053)819-1365
E-mail:namh@dhu.ac.kr