The Effects of Semantic Relatedness on EFL Vocabulary Recall and Retention

Shinyoung Bak
(International Graduate School of English)


Empirical research on the effects of semantic relatedness on EFL new vocabulary presentation has far yielded inconsistent results. Thus, this study aims to examine to what extent semantic relatedness influences EFL vocabulary recall and retention for primary school students in Korea. The current study was conducted with 107 primary school students over 4 weeks. To compare the effects of presenting semantically related words (SR) and unrelated words (SU), participants were divided into two groups, the semantically related words group (Group R) and the unrelated group (Group U). All the participants learned the same 40 English words paired with their Korean translation through one of two different methods. The results revealed that both presentation methods have positive effects on EFL vocabulary recall and retention. Between these two methods, SU was found to yield better results on recall and retention. This difference between groups could be explained from the perspective of interference theory, the distinctive hypothesis, or cross-association. Lastly, significant perceptual change concerning the effectiveness of related words presentation was found in Group R.

I. INTRODUCTION

Vocabulary teaching attracted little interest in second language (L2) acquisition before the 1990s (Meara, 2002). Many scholars and teachers seemed to believe that L2 learners would acquire vocabulary naturally just as they did in their first language (L1) (Coady, 1997). Moreover, it is frequently assumed that vocabulary learning is a simple memorization task which requires lower-level intellect and is easier than grammar learning (Coady, 1997). Therefore, vocabulary focused instruction has assumed little value.

However, the value of vocabulary learning seems to have found its place since the early 21th century. Read (2004) argued that natural vocabulary acquisition without any
instruction might be a false impression caused by the process of incidental vocabulary learning (Hulstijn, 2001). Wilkins (1972) clearly emphasized the role vocabulary plays in language by stating that “while without grammar very little can be conveyed, without vocabulary nothing can be conveyed” (as cited in Schmitt, 2010, p. 3). Currently, many of researchers are generally in agreement that incidental vocabulary learning is certainly another way for learners to acquire vocabulary knowledge besides intentional learning, and that incidental learning alone cannot explain all vocabulary growth (Nation, 2001; Read, 2004). This could also be true for intentional vocabulary learning because these two approaches are seen as complementary to each other in vocabulary learning (Hulstijn, 2001; Schmitt, 2000). Based on this general agreement, the main focus in L2 vocabulary research seems to have changed from the necessity of vocabulary learning to the most effective means of undertaking it (Folse, 2004). Regarding effective vocabulary learning, some studies have focused on the effects of various vocabulary presentation techniques for intentional learning such as pictures, example sentences, and L1 translation (Chin, 2002; Jeong, 2007; Kwon, 2004; Ramachandran & Rahim, 2004; Rodriguez & Sadoski, 2000; Webb, 2007).

Among these investigations, the effects of simultaneous presentation of semantically related words on L2 vocabulary learning still seem arguable to some extent because of discrepant findings of studies on this issue (Baleghizadeh & Naeim, 2011; Bolger & Zapata, 2011; Chin, 2002; Erten & Tekin, 2008; Finkbeinter & Nicol, 2003; Hashemi & Gowdasiaei, 2005; Jullian, 2000; Tinkham, 1993; Waring, 1997). Despite its as-yet-unproven effects, there is a good possibility that teachers and students may unquestioningly accept the use of semantically related vocabulary as an effective means of vocabulary presentation in Korea. Presenting semantically related words is probably one of the more popular techniques for some course books used in the English language classroom (Bolger & Zapata, 2011; Finkbeiner & Nicol, 2003; Khosravizadeh & Mollaei, 2011). In Korea, this technique can be found in some units of the text books of primary school students. In these units, target words which share a common feature of meaning are presented together under a certain topic such as occupation, family members, and weather. Because this method generally fits most well with a communicative language teaching (CLT) approach and a notional-functional syllabus (Folse, 2004; Tinkham, 1993), both of which have influenced the Korean national curriculum, this method of vocabulary presentation could be thought of as one of the more natural or useful methods by teachers and students. Under these genuine circumstances, the necessity arises to compare the effect of presentation of semantically related words and that of unrelated words.
II. LITERATURE REVIEW

In spite of the popularity of using semantic relatedness on L2 vocabulary learning, the results of the existing research on the effect of this method have been inconsistent as well as rather controversial. For a better understanding of this issue, the research findings and theories supporting each side of the argument are reviewed. Based on the review, the specific research questions of the present study are addressed.

1. Semantic Relatedness in Vocabulary Teaching

Generally, semantic is an adjective which means “used to describe things that deal with the meaning of words and sentences” (Collins Cobuild English dictionary for advanced learners, 2003). In the field of vocabulary teaching, semantically related words are the sets of words which have certain connections, share common meanings, or compose a network in meaning. Khosravizadeh and Mollaee (2011) grouped these connections under the concept semantic field and described this as an “interrelationship between the words” (p. 20). Within these connections, the meaning of the words in the same semantic field seems to be defined and limited by each other’s meaning (Khosravizadeh & Mollaee, 2011).

In the current study, the term lexical-set or semantically related words refers to a group of words which are related in meaning or can be arranged under a broader concept (Hashemi & Gowdasiaei, 2005). For instance, the following four words “blue, red, black, and white” share some semantic characteristics and could be arranged under a concept of “color.” If the word “blue” is presented with other words like “depressed, unhappy, or low,” these words could form another potential network. In this new network, the meaning of the same word “blue” would probably be interpreted differently because the other words in same group would constrain or clarify its meaning.

2. Presenting Semantically Related Vocabulary (SR)

The pragmatic benefit has possibly contributed to the popularity of lexical-sets in some widely used course books for English class (Erten & Tekin; 2008; Waring, 1997). Semantic clusters generally fit well with most English learning courses regardless of their approach and seem to be an equally convenient method for material developers (Folse, 2004; Tinkham, 1993). Moreover, course book authors who favor lexical-sets have believed that showing the connections among words promotes learners’ vocabulary concept learning (Finkbeiner & Nicol, 2003; Folse, 2004).

This belief is supported by certain theories and has been corroborated by the findings of some research studies. First, one of the rationales of presenting related words can be drawn
from Ausubel’s (1963) meaningful learning (Hashemi & Gowdasiaei, 2005; Khosravizadeh & Mollaei, 2011). Ausubel distinguished between rote learning, a passive process, and meaningful learning, the active process of relating new information or concepts to learners’ prior knowledge. To encourage this information incorporating process, using an advance organizer is regarded as one of the more useful strategies (Santrock, 2004; Slavin, 2003). When the words under a broader concept are presented synchronously, the concept acts as an advance organizer showing meaningful connections between words, and thus assisting the vocabulary concept learning process (Hashemi & Gowdasiaei, 2005).

Second, studies on the mental lexicon generate another rationale for this position. Word association analysis in L1 studies has yielded a general consensus about the existence of a mental lexicon, a well organized network of words and concepts (Collins & Loftus, 1975; Schmitt, 2010; Wolter, 2001). Aitchison (2003) stated that native speakers seem to have a similar word-web which has logical organization: including “coordination, collocation, superordination, synonym” (p.86). This word-web, termed a mental lexicon, helps native speakers recall and retain the vocabulary better as learners notice the connections among words (Aitchison, 2003). Similarly, Schmitt (1997) noted that grouping words benefits native speakers in the word list recall test. Advocators of using lexical-sets have insisted that these findings could be applied to L2 learning (Baleghizadeh & Naeim, 2011; Erten & Tekin, 2008; Schmitt, 2010, Wolter, 2006).

To be specific, a spreading activation model proposed by Collins and Loftus (1975) is one of the frequently cited theories to support the use of lexical-sets in vocabulary teaching (Bolger & Zapata, 2011; Hashemi & Gowdasiaei, 2005). In this model, the network consists of nodes representing words and lines between nodes representing connection between words. The length of the line shows how strongly the words are semantically associated (Randall, 2007). Once a certain node in a network is initiated, this activation spreads through the whole network, thus leading to the activation of other nodes in the network (Collins & Loftus, 1975). The spread activation primes the other nodes within the network and results in a faster process (Randall, 2007). Simultaneous presentation of related words possibly strengthens the links between words and facilitates vocabulary learning (Hashemi & Gowdasiaei, 2005). Additional theoretical support is found in the levels-of-processing theory (Morin & Goebel, 2001; Shapiro & Waters, 2005). Researchers have noted that recognized information can be processed at a variety of levels from shallow to deep, and that the amount of cognitive effort that is given to the process determines the quality of the retention (Craik & Lockhart, 1972 as cited in Erten & Tekin, 2008; Otten, Henson & Rugg, 2001). Proponents of lexical-sets have claimed that when the related words are presented at the same time, learners benefit from comparing, contrasting, and organizing or chunking the words (Chin, 2002; Hashemi & Gowdasiaei, 2005; Jullian, 2000; Randall, 2007; Seal, 1991).
Several empirical studies have explored how semantic relatedness affects vocabulary learning. Morin and Goebel (2001) investigated the effectiveness of semantic mapping as a vocabulary learning strategy. Four classes of English speaking college-level learners received Spanish vocabulary instruction through either a communicative-activity-only condition or a semantic-mapping-plus-communicative-activity condition. After a semester-long observation, researchers concluded that semantic mapping represents a useful strategy for beginners. Similarly, Chin (2002) explored how three learning strategies affected vocabulary learning and retention. Three groups of participants received the following treatment: in context, word list, and semantic mapping respectively. Subsequent to this, three types of post and delayed tests followed the treatment. In one type of post test, the fill in test, the context group produced the highest score but showed a significant decrease in the delayed test. On the other hand, the semantic mapping group yielded the second highest score in the posttest and a decline in their score was not found in the delayed test. Chin interpreted these results as signifying that semantic mapping is a beneficial strategy to assist in both learning and retaining vocabulary.

Some researchers have also asserted the value of semantic relatedness as a teaching technique. Julian (2000) noted that Spanish learners of English who were taught words in the same lexical-sets simultaneously enhanced their vocabulary knowledge and gained more accurate use of words. Hashemi and Gowdasiaei (2005) argued that learning semantically related vocabulary together facilitates learning. Iranian adult learners participated in their study and self-reported their vocabulary breadth and depth using the Vocabulary Knowledge Scale. The results showed that learners in the related group surpassed their counterparts in the unrelated group, regardless of their proficiency level. A similar study was conducted by Baleghizadeh and Naeim (2011) with a single adult learner. The learner learned two sets of forty-two words, the first set using a semantic mapping method and the other forty-two-word set of randomly grouped words without using a semantic mapping method. The positive effect of semantic mapping method was found on a meaning recall test and interview.

3. Presenting Semantically Unrelated Vocabulary (SU)

In contrast to the proponents of using lexical-sets, others have disputed the benefits of implementing this technique. Some researchers have attempted to advance the negative effects of simultaneous presentation of related words based on the interference theory (Baddeley, 1997; Erten & Tekin, 2008; Finkbeiner & Nicol, 2003; Papathanasiou, 2009a; Tinkham, 1993; Waring 1997).

Interference theory suggests that forgetting is caused not by a loss of memories, but rather because of other information around the targeted one (Santrock, 2004). Interference
occurs “when previously learned information is lost because it is mixed up with new and somewhat similar information” (Slavin, 2003, p. 189). Additionally, the level of similarity between information seems to be accepted as a crucial factor to determine the strength of the interference (Baddeley, 1997). The time gap between learning is suggested as another factor affects the level of interference (Bower, 2000; Roediger & McDermott, 2000).

In vocabulary learning, the more the presented words share meaning or the more closely the words are taught, the more interference is likely to occur with the words that are learned before or after (Tinkham, 1993). Along this same line, Nation (2006) proposed that learning related words together may contribute to the increased possibility of cross-association between the words. Adding to this claim, Waring (1997) reported that when learners were asked to match a set of artificial words to the their L1 (Japanese) words under a related words treatment condition, what caused major problems was not remembering the pseudo words themselves, but rather pairing them with their counterparts. Many researchers seem to agree that the result of simultaneous presentation of related vocabulary is increased difficulty in vocabulary learning (Erten & Tekin, 2008; Nation, 2001; Pigada & Schmitt, 2006; Tinkham, 1993, 1997). Similarly, Hunt and Mitchell (1982) hypothesized that what makes learning easier is the resultant distinctiveness or non-similarity of the information by making it more salient and noticeable (as cited in Papathanasiou, 2009b). This idea, known as the distinctiveness hypothesis has been applied to explain the reason of why presentation of words in the same semantic field hinders learning (Mirjalili, Jabbari, & Rezai, 2012; Tinkham, 1993, 1997). According to this hypothesis, words from the same lexical-set which share a common meaning yet present subtle differences are difficult to learn due to the lack of salient features.

Others have pointed out that though the semantic relatedness could lead to deeper level of cognitive processing or spreading activation, it may actually hinder learning because of a limited working memory capacity (Klemm, 2007). When the words sharing common meanings are taught at the same time, the process of discriminating subtle differences between the words’ meaning requires some of the working memory trace (Erten & Tekin, 2008). As a result, a reduced amount of memory capacity is available for associating the unfamiliar L2 form and the known concept. This relatively diminished chance for consolidation of the meaning could result in a weaker connection, and therefore less effective learning (Erten & Tekin, 2008).

Researchers have argued that semantic relatedness impedes vocabulary learning based on the findings of a body of empirical studies (Bolger & Zapata, 2011; Erten & Tekin, 2008; Finkbeinter & Nicol, 2003; Tinkham, 1993; Warning, 1997). Tinkham compared the learning rate of the participants in a semantically related treatment condition and an unrelated treatment condition. All participants, who were native English speakers, were asked to remember artificial words that were paired with English words, and match the
pairs in as few trials as possible. Participants in the unrelated group completed the task faster than those in the related group. Later, Waring replicated this study with Japanese students. The only difference in the experiment design was the fact that English words were translated into and replaced with Japanese. The result of this study paralleled that of the original study (Tinkham, 1993). Finkbeinter and Nicol also studied how semantic relatedness influences vocabulary learning. Two groups of English-speaking participants received vocabulary training individually through computers. Thirty-two artificial words which were grouped into 4 categories, 2 related-sets and 2 unrelated-sets, were presented with their pronunciations and corresponding pictures to show meanings. The average required time for the translation tasks in both directions (L1-L2, L2-L1) were shorter in the unrelated group.

Other researchers have conducted similar studies focusing on young learners (Erten & Tekin, 2008). All the participants, who were fourth-grade Turkish students, learned 2 sets of 40 semantically related words and the same number of unrelated words. Oral presentation of the target words with corresponding pictures and the written form were followed by practice that consisted of repetition, gestures, and mimes to aid learning. Word and picture matching tasks were conducted three times for each category as a pre-, post-, and delayed test. Erten and Tekin reported that learners achieved higher scores on the test of unrelated words.

Even more recently, Bolger and Zapata (2011) investigated the combined effect of context and lexical-sets. The researchers borrowed the 4 sets of 32 pseudo words and its corresponding pictures from the study undertaken by Finkbeinter and Nicol (2003) and developed 4 stories containing a set of words each. Participants read two stories individually on the computer screen according to their treatment groups and completed two tests: a semantic categorization and a stimulus-matching verification. Bolger and Zapata reported that the unrelated group still surpassed the related group but the gap between scores was neutralized compared to Finkbeinter and Nicol’s study. The reduced effect of using semantic sets in Bolger and Zapata’s study was attributed to the contextual support.

As described above, a number of studies have investigated the effect of using lexical-sets in L2 vocabulary learning. Notwithstanding the somewhat discrepant findings that previous research has yielded, the idea of using lexical-sets seems to be widely applied in some language learning course books. Moreover, it is often regarded as one of the more effective methods. Although some recent studies (Baleghizadeh & Naeim, 2011; Hashemi & Gowdasiaei, 2005; Jullian, 2000) reported the positive role of lexical-sets in vocabulary learning, advantages of this method may still be arguable.

Additionally, in order to apply this approach to young Korean learners of English, further inquiry seems to be essential. Thus, this study weighs the following three items:
young beginner, real words, and classroom dynamics. First, very few studies have been implemented on this issue in Korea with young learners. The focus of the present study is on primary school students who are at the beginning level. Second, real L2 words are used as the target vocabulary instead of pseudo words. The need for a study applying real L2 words was pointed out in Erten and Tekin’s (2008) study. Unlike many studies (Boler & Zapata, 2011; Finkneiner & Nicol, 2003; Tinkham, 1993, 1997; Waring, 1997), that used artificial words paired with L1 words, this study employs real English words paired with Korean equivalents to explore the L2 learning process. Third, the current study is conducted in four classes of public primary schools to better account for classroom dynamics. One recent investigation (Baleghizadeh & Naeim, 2011) which reported on the effectiveness of lexical-sets was a single subject study; therefore, a study conducted in a classroom seems a valuable way to ascertain the results. Additionally, Waring (1997) noted that experimental designs which several studies have followed may have been stressful to participants, thus resulting in the possibility of different outcomes being produced in other circumstances.

Thus, this study aims to examine to what extent semantic relatedness affects L2 vocabulary recall and retention for primary school students in Korea. For this purpose, the effects of the two types of vocabulary presentation, SR and SU are compared. The specific research questions are as follows:

1. To what extent do different types of vocabulary presentations (SU and SR) affect L2 vocabulary recall?
2. To what extent do different types of vocabulary presentations (SU and SR) affect L2 vocabulary retention?
3. To what extent does presentation of semantically related words affect learners’ perceptions toward its effectiveness?

III. METHOD

1. Participants

The participants of this study were 116 fifth-grade students in four classes from a public primary school in Gwangju Metropolitan City. However, nine students were excluded from the final sample because five of them transferred to another schools and a further four were absent for one of the vocabulary tests. Thus, the final sample of 107 students, 54 male students and 53 female students, completed the entire treatment and tests. All the participants received regular English classes three times per week from their own
homeroom teacher and each class period was 40 minutes in length. Four classes were divided into two groups and then one group (Class A and Class B) was randomly assigned as the semantically unrelated group (Group U) and the other one (Class C and Class D) was assigned to the semantically related group (Group R).

### TABLE 1

<table>
<thead>
<tr>
<th></th>
<th>Group U</th>
<th></th>
<th>Group R</th>
<th></th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Class A</td>
<td>Class B</td>
<td>Class C</td>
<td>Class D</td>
<td></td>
</tr>
<tr>
<td>Boy</td>
<td>11</td>
<td>15</td>
<td>15</td>
<td>13</td>
<td>54</td>
</tr>
<tr>
<td>Girl</td>
<td>15</td>
<td>12</td>
<td>14</td>
<td>12</td>
<td>53</td>
</tr>
<tr>
<td>Total</td>
<td>26</td>
<td>27</td>
<td>29</td>
<td>25</td>
<td>107</td>
</tr>
</tbody>
</table>

2. Materials

1) Vocabulary List

While selecting the target vocabulary, students’ interest, semantic relatedness, and learnability were taken into consideration. First, 30 main concepts were drawn from the dictionaries: *Collins junior thesaurus* (1989), *Longman children’s picture dictionary* (2003), *The basic Oxford picture dictionary* (2003), *The Heinle picture dictionary* (2005), *The Oxford picture dictionary* (2008). Second, 29 non-participating students attending the same primary school took a survey which asked them to choose two interesting topics. According to the survey results, 10 topics were selected in total. Third, six vocabulary items were drawn on each topic from the dictionaries previously listed, and 60 words were compiled in total. Fourth, five primary school teachers examined whether learners could understand the meaning of these 60 words in their L1 because slightly difficult words were selected to exclude known words as much as possible. Through this examination, 13 words were excluded and 47 words remained. Fifth, a pilot test was conducted in order to include as many unknown items as possible. The same students who participated in the survey completed the pilot test, and 10 words which were relatively well known to students were excluded. Lastly, three new words were added in order to distribute the same number of vocabulary items to each category. The 40 selected words were organized into two types: semantically related word sets and unrelated word sets. In other words, the participants learned identical vocabulary items but in a different order according to their group (Appendix A).
2) Vocabulary Tests

Because of the students' level, the vocabulary test focused mainly on the receptive aspect - recognition of the written form. According to this purpose, in the present study a test of 32 modified items following the format of the Vocabulary Level Test (Nation, 2001) was administered. Forty target words were randomly divided into six groups. Four to six Korean equivalents were listed on the right side and eight to 12 target words were given on the left side. The number of target words was twice as many as that of equivalents so half of the target words were not able to be matched to their counterparts and thus served as distracters. One point was allocated for each correct answer; therefore, the maximum score was 32 points for every vocabulary test.

This test was implemented three times in the present study as a pretest, post test, and delayed test. To prevent participants from remembering the tested items, 32 words were randomly chosen out of the 40. None of the answers or scores on the tests were revealed to students for this exact same purpose.

3) Treatment Materials

Two types of treatment materials were provided during the four weeks of treatment: PowerPoint materials for the whole class, and worksheets for individual self-study. The PowerPoint material was designed to minimize the teacher as a variable during the presentation phase. Four classes were taught by their homeroom teachers who vary in their teaching styles and English proficiency levels; these differences could affect the quality of presentation and eventually the test scores. Hence, the target words were presented through PowerPoint material with pictures that showed the meaning of the words, the spoken form, the written form, and the Korean equivalent. After the initial presentation, participants also listened and repeated the pronunciation of each word eight times as a whole class activity. Worksheets for self-study provided identical information on the new words to the PowerPoint material except for the spoken form. The worksheets included some tasks for indentifying the word form and matching this form to meaning. All material was distributed only during the allocated time to control the time spent on vocabulary learning.

4) Questionnaires and Interviews

A pre-questionnaire was administered before the treatment for two reasons. First, seven items revised from Kim (2009) and Lim (2009) were used for gathering information about students’ English learning background. Second, three items were used to investigate learners’ attitudes to vocabulary learning by a four-point Likert scale (1: strongly disagree;
The Effects of Semantic Relatedness on EFL Vocabulary Recall and Retention

4: strongly agree). In addition to these seven items, one additional item about the perceived effectiveness of lexical-sets was included in the survey of only Group R. After the treatment this one item was asked Group R again as a post-questionnaire to examine whether the treatment would affect their perception or not. Lastly, semi-structured interviews were conducted by the researcher to uncover the participants' perceptions. Four homeroom teachers and four students from each treatment group, eight students in total, were individually interviewed for about five to 10 minutes.

3. Procedures

The procedure was divided into four phases: before, during, immediately after, and four weeks following the treatment. Before the treatment, a pre-questionnaire and vocabulary pretest was conducted. Participants completed the pre-questionnaire two weeks before and the vocabulary pretest one week before the treatment had begun. In the pre-questionnaire, Group R had one more question, related to perceived effectiveness toward SR, on their questionnaire than Group U. Students were also informed that the extra vocabulary learning session would start the next week.

During the treatment, 20-minutes extra time was assigned four times a week before the regular classes began, usually from 8:40 a.m. to 9:00 a.m. The treatment for the study lasted four weeks from October 17 to November 11 in 2011. Each learning session consisted of a whole-class presentation and a self-study phase. In the presentation phase, the target words are shown about four times through the PowerPoint material to the whole-class with pictures, spelling, pronunciation, and a Korean translation. The rest of the given time was allocated for individual study through the tasks in the worksheet. During all learning sessions, homeroom teachers were present in the classroom for the purpose of classroom management, not instruction. Although participants completed a quiz at the end of every week, the scores were not recorded because the quiz was intended to keep participants interested in the learning procedure.

Right after the treatment, a vocabulary posttest was conducted from November 14 to the 16 in order to explore the treatment effect on vocabulary meaning recall. During this same period, Group R received a post-questionnaire. Additionally, semi-structured interviews were conducted on November 16 to gather more qualitative information. Approximately one month later, the delayed test of vocabulary was administrated without any review of the target words in order to inspect the treatment effects on vocabulary retention.

4. Data Analysis

The test scores from three L2 vocabulary tests and the data from the questionnaire were
analyzed to ascertain the effects of semantic relatedness on L2 vocabulary learning and perception. First, an analysis of variance (ANOVA) with repeated measures was employed to analyze the vocabulary test scores. The within-subjects variable was Time (pre, post, delayed) and the between-subjects variable was Group (unrelated, related). The dependent variable was test scores of each vocabulary test. A significance level of .05 was applied.

Second, the data from the questionnaires were analyzed with a paired t-test to investigate the perception change in Group R. The independent variable was Time (pre, post) and the dependent variable was scores of each questionnaire item measured by a four-point Likert scale. A significance level of .05 was applied.

IV. RESULTS

1. Vocabulary Test Results

Data in the form of the number of correct answers out of 32 questions was analyzed using ANOVA with repeated measures. Table 2 displays the descriptive statistics of the test scores. In the pretest, Group U achieved a mean of 1.83 and the average of the Group R was 3.30. The average of the post-test was 25.91 for Group U and 20.61 for Group R. Lastly, Group U achieved a mean of 23.38 in the delayed test, and that of Group R was 17.80. The complete source Table (Table 3) illustrates that there was a significant main effect for Group (Group U vs. Group R), and Time (pre vs. post vs. delayed).

<table>
<thead>
<tr>
<th>Time</th>
<th>Group</th>
<th>N</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre</td>
<td>U</td>
<td>53</td>
<td>1.83</td>
<td>2.19</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>54</td>
<td>3.30</td>
<td>3.24</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>107</td>
<td>2.57</td>
<td>2.88</td>
</tr>
<tr>
<td>Post</td>
<td>U</td>
<td>53</td>
<td>25.91</td>
<td>8.77</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>54</td>
<td>20.61</td>
<td>9.40</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>107</td>
<td>23.23</td>
<td>9.43</td>
</tr>
<tr>
<td>Delayed</td>
<td>U</td>
<td>53</td>
<td>23.38</td>
<td>9.12</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>54</td>
<td>17.80</td>
<td>9.68</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>107</td>
<td>20.56</td>
<td>9.78</td>
</tr>
</tbody>
</table>
The Effects of Semantic Relatedness on EFL Vocabulary Recall and Retention

TABLE 3
Source Table: ANOVA with Repeated Measures by Group and Time

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>Sig</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Between Subjects</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>76810.18</td>
<td>1</td>
<td>76810.18</td>
<td>629.16</td>
<td>.00*</td>
</tr>
<tr>
<td>Group</td>
<td>789.40</td>
<td>1</td>
<td>789.40</td>
<td>6.47</td>
<td>.01*</td>
</tr>
<tr>
<td>Error</td>
<td>12818.86</td>
<td>105</td>
<td>122.08</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Within Subjects</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td>17377.94</td>
<td>1</td>
<td>17377.93</td>
<td>471.33</td>
<td>.00*</td>
</tr>
<tr>
<td>Time x Group</td>
<td>664.18</td>
<td>1</td>
<td>664.18</td>
<td>18.01</td>
<td>.00*</td>
</tr>
<tr>
<td>Error</td>
<td>3871.31</td>
<td>105</td>
<td>36.87</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* p < .05

The significant main effect for Group means that Group U significantly outperformed Group R as shown in Table 4. Because the significant main effect for Time was found in Table 3, a post-hoc test for Time was performed to identify which mean was significantly different from others. Table 5 demonstrates that there were significant mean differences between 3 pairs: the pretest and posttest; the posttest and delayed test; and the pretest and delayed test. To elaborate, the scores of the test significantly increased in the posttest compared to the pretest and then significantly decreased in the delayed test. Despite this decline, participants achieved significantly higher scores in the delayed test compared to the posttest.

TABLE 4
Post Hoc Test: Pairwise Comparison by Group

<table>
<thead>
<tr>
<th>(I) Group</th>
<th>(J) Group</th>
<th>MD (I-J)</th>
<th>Std. Error</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>R</td>
<td>U</td>
<td>-3.14</td>
<td>1.23</td>
<td>.01*</td>
</tr>
</tbody>
</table>

* p < .05

TABLE 5
Post Hoc Test: Pairwise Comparisons by Time

<table>
<thead>
<tr>
<th>(I) Time</th>
<th>(J) Time</th>
<th>MD (I-J)</th>
<th>Std. Error</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre</td>
<td>Post</td>
<td>-20.70</td>
<td>.85</td>
<td>.00*</td>
</tr>
<tr>
<td>Pre</td>
<td>Delayed</td>
<td>-18.02</td>
<td>.83</td>
<td>.00*</td>
</tr>
<tr>
<td>Post</td>
<td>Delayed</td>
<td>2.67</td>
<td>.44</td>
<td>.00*</td>
</tr>
</tbody>
</table>

* p < .05

一하지 않고 KYOBO Book Centre
At the same time, the significant interaction between Group and Time was found (Table 3). To determine which of the means were significantly different from others, a post-hoc test was performed. The results of the post-hoc test are displayed in Table 6 and Table 7.

### TABLE 6
Post Hoc Test: Pairwise Comparisons for Interaction Time x Group

<table>
<thead>
<tr>
<th>Time</th>
<th>Group (I)</th>
<th>Group (J)</th>
<th>MD (I-J)</th>
<th>Std. Error</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre</td>
<td>U</td>
<td>R</td>
<td>-1.47</td>
<td>.54</td>
<td>.008</td>
</tr>
<tr>
<td>Post</td>
<td>U</td>
<td>R</td>
<td>5.30</td>
<td>1.76</td>
<td>.003*</td>
</tr>
<tr>
<td>Delayed</td>
<td>U</td>
<td>R</td>
<td>5.58</td>
<td>1.81</td>
<td>.003*</td>
</tr>
</tbody>
</table>

* * p < .0055

### TABLE 7
Post Hoc Test: Pairwise Comparisons for Interaction Group x Time

<table>
<thead>
<tr>
<th>Group</th>
<th>Time (I)</th>
<th>Time (J)</th>
<th>MD (I-J)</th>
<th>Std. Error</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>U</td>
<td>Pre</td>
<td>Post</td>
<td>-24.08</td>
<td>1.20</td>
<td>.000*</td>
</tr>
<tr>
<td></td>
<td>Pre</td>
<td>Delayed</td>
<td>-21.55</td>
<td>1.18</td>
<td>.000*</td>
</tr>
<tr>
<td></td>
<td>Post</td>
<td>Delayed</td>
<td>2.53</td>
<td>.63</td>
<td>.000*</td>
</tr>
<tr>
<td>R</td>
<td>Pre</td>
<td>Post</td>
<td>-17.32</td>
<td>1.19</td>
<td>.000*</td>
</tr>
<tr>
<td></td>
<td>Pre</td>
<td>Delayed</td>
<td>-14.50</td>
<td>1.17</td>
<td>.000*</td>
</tr>
<tr>
<td></td>
<td>Post</td>
<td>Delayed</td>
<td>2.82</td>
<td>.62</td>
<td>.000*</td>
</tr>
</tbody>
</table>

* * p < .0055

All significance values, except for between groups at the pretest time, are smaller than the adjusted alpha level according the Bonferroni Adjustment\(^1\) (Table 6; Table 7). No difference between groups at pretest indicates that the two groups are at the same level before the treatment. Scores of both groups increased significantly in the post test and then significantly decreased in the delayed test, while Group U noticeably outperformed Group R in both tests. These results seem to indicate that SU is more effective than SR.

2. Findings from Questionnaires and Interviews

Participants in Group R were asked how much they agreed with the effectiveness of SR

---

\(^1\) The Bonferroni Adjustment is a “procedure to control type I error rates. In this procedure, the original alpha level chosen for significance is divided by the number of planned comparisons. Each comparison must be significant at this level in order to be declared significant” (Hays, 1994, p. 452). According to the Bonferroni Adjustment, significance levels for this analysis should be adjusted to 0.0055 (=0.05/9) because there were nine pairwise comparisons in total.
before and after the treatment. The results of the statistical analysis with a paired t-test are shown in Table 8. The mean of the pre-questionnaire was 3.42 and that of the post-questionnaire was 3.22. Although about 85% of learners still agreed with the effectiveness of SR in the post-questionnaire (Table 9), the result of the t-test shows that the change of learners’ perception was statistically significant (Table 8). These results can be interpreted as signifying that the SR treatment obviously changed students’ perception of the method from a more effective to less effective one. In other words, some learners may have noticed the ineffective feature of studying related words together resulting in changes to their opinion in the post-questionnaire.

**TABLE 8**

<table>
<thead>
<tr>
<th>Time</th>
<th>N</th>
<th>M</th>
<th>SD</th>
<th>t</th>
<th>df</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre</td>
<td>51</td>
<td>3.42</td>
<td>0.09</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post</td>
<td>51</td>
<td>3.22</td>
<td>0.09</td>
<td>2.11</td>
<td>50</td>
<td>0.04*</td>
</tr>
</tbody>
</table>

* p < .05

**TABLE 9**

<table>
<thead>
<tr>
<th>Time</th>
<th>Disagree</th>
<th>Agree</th>
<th>Total (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Pre</td>
<td>0.0</td>
<td>77.7</td>
<td>40.4</td>
</tr>
<tr>
<td>Post</td>
<td>0.0</td>
<td>15.1</td>
<td>50.9</td>
</tr>
</tbody>
</table>

1 = strongly disagree, 4 = strongly agree

In addition to the findings described above, the results of the second part of the pre-questionnaire about the learners’ attitudes to vocabulary learning are provided in Table 10 and Table 11. The means of all three survey questions were more than 2.88 (Table 10). These high mean scores seem to indicate that learners generally have positive attitudes toward vocabulary learning. As illustrated in Table 11, slightly more than 90% of learners agreed that vocabulary knowledge could contribute to improving their general English proficiency. To the other two questions, about three out of four students responded that they are interested in vocabulary learning and eager to learn more words in school.
TABLE 10
Descriptive Statistics of Learners’ Attitudes to Vocabulary Learning

<table>
<thead>
<tr>
<th>Questions</th>
<th>N</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vocabulary learning is useful to improve general English proficiency.</td>
<td>107</td>
<td>3.27</td>
<td>0.65</td>
</tr>
<tr>
<td>Learning vocabulary is interesting.</td>
<td>107</td>
<td>2.88</td>
<td>0.72</td>
</tr>
<tr>
<td>I’d like to learn more vocabulary in school.</td>
<td>107</td>
<td>2.98</td>
<td>0.77</td>
</tr>
</tbody>
</table>

1 = strongly disagree, 4 = strongly agree

TABLE 11
Frequency for Learners’ Attitudes to Vocabulary Learning

<table>
<thead>
<tr>
<th>Questions</th>
<th>Disagree</th>
<th>Agree</th>
<th>Total (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vocabulary knowledge contributes to general English proficiency.</td>
<td>1.9</td>
<td>56.1</td>
<td>100</td>
</tr>
<tr>
<td>Learning vocabulary is interesting.</td>
<td>3.7</td>
<td>57.9</td>
<td>100</td>
</tr>
<tr>
<td>I’d like to learn more vocabulary in school.</td>
<td>4.7</td>
<td>56.1</td>
<td>100</td>
</tr>
</tbody>
</table>

1 = strongly disagree, 4 = strongly agree

In addition to the surveys, semi-structured interviews were conducted to investigate the learners’ perceptions toward semantic relatedness in vocabulary learning. Learners in Group R tended to more frequently report that they confused words with each other: “I can say these two words (pointing out the two words, *thigh* and *calf*) are names of body parts. But I’m still confused which one is which.” Some words were pointed out as the most troublesome pairs by two homeroom teachers of Group R: “I think many of my students had some difficulties to learn the words *thigh* and *calf*, and *mustache* and *beard*. Frankly, even I was confused, too.” These responses seem to parallel those of Tinkham’s (1997) study.

V. DISCUSSION

This study aimed to investigate the effects of semantic relatedness on vocabulary recall and retention of young Korean learners of English. To reflect real L2 learning process especially of Korean young learners, four classes of primary school students who are at the beginning level participated in the study. For the same purpose, this study employed real English words and natural classroom environment instead of the artificial words and
The Effects of Semantic Relatedness on EFL Vocabulary Recall and Retention

The first and second research questions attempted to examine to what extent semantic relatedness influences on L2 vocabulary recall and retention respectively. Participants in both groups produced significant progress in their posttest and then appreciably diminished their scores in the delayed test. Despite the loss, the scores of both groups in the delayed test were still significantly higher compared to those of the pretest. These results symbolize that both vocabulary presentation methods have a positive effect on L2 vocabulary recall and retention. Among the two groups, Group R yielded significantly higher scores than Group U in post- and delayed tests of vocabulary. The divergent post-treatment scores of the two groups may suggest that SU demonstrates a greater influence than SR does on L2 vocabulary recall and retention.

Overall, SU yields better results on vocabulary recall and retention. These positive influences can be explained by various reasons. One possible reason could be drawn from interference theory. That is to say, similarities between the words provided for Group R might have promoted interference. Though a natural level of interference probably occurred under the unrelated words treatment condition (Santrock, 2004; Slavin, 2003), a stronger effect seemed to appear under the related words treatment condition. The level of similarity and the time gap between learning have generally been accepted as contributing to the strength of the interference (Bower, 2000; Nation, 2000, Tinkham, 1993), hence, learners in Group U might have experienced less interference which in turn may have led to less difficulty in learning. Moreover, there seems to be a possibility that the interference was linked to the cross-association of individual items as revealed in the interview. One learner said that she was unable to pair the target words and their Korean meaning despite the fact that she confidently recall two Korean translations of the words. Even teachers of Group R agreed with the learner’s explanation. This feedback seems to correspond with what Waring (1997) reported. In his study learners noted that associating the pseudo words with their corresponding meaning caused more difficulty than remembering the artificial words (Waring, 1997). Furthermore, Kroll and Stewart’s (1994) concept of the asymmetrical storage model could provide another possible reason (Kroll & Stewart, 1994 as cited in Randall, 2007). According to this model, at the initial stage of learning, the new L2 form is connected more firmly to the L1 form and only weakly to the concept. Due to the weak connection, the L2 form tends to be processed mostly through the L1 form which has a firm connection to concept. This mediated process would be substituted with a direct process after learners build stronger associations between the new form and its concept.

In the present study, receiving sets of related words together might have increased the process of discriminating between subtle meaning differences separating the words. As a result of this cognitive load requiring an adequate amount of the working memory (Waring,
1997), less of the working memory would be available for associating an unfamiliar L2 form with a known concept. Thus, learners in Group R who had relatively fewer chances for building links between form and meaning may have experienced an increased potential for interference or cross-association which eventually resulted in less effective learning. On the other hand, learners in Group U might have been able to focus on building connections between form and meaning without concerning themselves with other possible connections such as meaning between related words. Additionally, Group U was likely able to benefit from the distinctiveness among words when unrelated words were presented. Hunt and Mitchell (1982) stressed that non-similarity among words could promote discrimination, and thus improve the learning (as cited in Papathanasiou, 2009b). Learners in Group U who received five words from different categories may have learned the words better by using the words' notable features.

Another reason for the considerable advantage of SU might be attributed to the proportion of familiar words. In the present study, only one out of 10 words was known to learners at the initial stage; therefore, students were likely affected greatly by interference. For this reason, semantic mapping is suggested as a useful activity for "clarifying and enriching the meaning" when learners already know the words (Graves, 2006, p. 79). Nation (2000) similarly pointed out that teaching similar words may have its value after the meanings of the words are "well established" (p. 9). Furthermore, Folse (2004) highlighted the fact that using semantic relatedness may be useful for reviewing rather than introducing the words. To sum up, the process of connecting a form to a meaning differentiates L2 vocabulary learning from L1 word recall. Therefore, new L2 forms may need to be firmly connected to the concept before they are compared with others. To encourage the connecting process, SU seems more expedient.

The last research question was designed to explore whether receiving SR affects learners' perceptions toward its effectiveness. To investigate this, the data from the questionnaires were analyzed using a paired t-test. The results seem to demonstrate that the treatment noticeably affected the perceived effectiveness of the treatment method. This result was supported by the findings from the interviews. The interviews revealed that several of the students may have recognized the interference or cross-association between words, especially those in the same lexical-set.

Though a significant change in perceptions was found (Table 8), about 85% of participants in Group R still seemed in favor of presentation of related words over that of unrelated words in general (Table 9). This phenomenon could be explained by two possible reasons. First, the treatment period of four weeks might not have been long enough for participants to detect any disadvantages of SR which was strongly believed by them to be more effective. Second, there is a possibility that organization of the textbook may affect students' perceptions. Even after learners recognized the fact that learning
similar words is less effective for vocabulary learning, they may accept SR as productive because it has been, and continues to be, one of the more common techniques that they have learned words through in their text books.

VI. CONCLUSION

The present study aims at understanding the role of semantic relatedness in L2 vocabulary recall and retention for young learners of English in Korea. The results showed that SU generates significantly greater gains in L2 vocabulary recall and retention. Additionally, learners’ perceptions toward SR changed from more effective to less effective by receiving vocabulary learning sessions through SR.

From these findings of the present study, several implications for L2 vocabulary teaching are listed. One of the more important findings from this study is that extra-vocabulary learning sessions, of any kind, could contribute to vocabulary growth. This finding is likely to support the need for extra vocabulary-focused learning in elementary schools. Many researchers have recommended explicit or direct vocabulary teaching at initial stages of learning to ensure learners acquire the 2,000-3,000 most frequent words, often referred to as core vocabulary (Nation, 2008; Schmitt, 2000; Thornbury, 2002). Because this core vocabulary could cover around 80% of the running words in most written text and around 90% of basic conversation (Nation, 2008), it is anticipated that learners will have more chance to learn vocabulary incidentally from comprehensible input (Coady, 1997; Schmitt, 2000). Therefore, scholars have insisted that core vocabulary should be learned as soon as possible at the initial stage of learning (Meara, 1995; Thornbury, 2002). Nation (2008) stated that high-frequency words deserve to be taught explicitly despite the limited class time available and could be learned in three to five years in school. In Korea, it is reasonable to conjecture that elementary schools, where public English education starts, are responsible for teaching the core vocabulary to enable learners to prepare for further learning. Jeong (2009) observed that the importance of vocabulary learning seems to be reflected in the revised English curriculum (Ministry of Education, Science and Technology [MEST], 2008). However, the number of recommended words in the current English curriculum (MEST, 2011), about 500 words, is still less than a quarter of the core vocabulary size. Hence, there remains the distinct possibility that additional vocabulary-focused learning will contribute to optimal English education.

Second, material developers and teachers could emerge as responsible for being aware of the strengths and weaknesses of vocabulary presentation methods. None of the available presentation methods should be excluded (Waring, 1997), but rather should be applied
appropriately in order to maximize the strengths and minimize the weaknesses of each respective method. For example, presentation of semantically related words could deepen and widen the understanding of known words by connecting and contrasting them (Folse, 2004; Jullian, 2000; Papathanasiou, 2009b). On the other hand, presentation of semantically unrelated words could significantly assist L2 vocabulary learning, especially when the majority of the words are new to the learners (Baleghizadeh & Naeim, 2011; Nation, 2001). When material developers and teachers have knowledge of practical findings from empirical studies, their presentation methods could then be diversified according to the level of target learners and type of the activity in the course books and the classroom. What has been commonly suggested as an alternative for SR is thematic clustering (Bolger & Zapata, 2011; Folse, 2004; Tinkam, 1997; Waring, 1997). When SR is used, confusion could be reduced by making words non-interchangeable (Thornbury, 2002). For example, words can be introduced with commonly associated collocates, in different contexts, or through different sensory modes (Folse, 2004, Nation, 2001).

Third, students need to be instructed in the differing effectiveness of using these two approaches: SR and SU. Nation (2000) stated that learners often express a preference for SR and this tendency was revealed in the questionnaire data of the current study; many young Korean learners stated that they prefer to learn semantically related words together. Though significant numbers of learners changed their thoughts, about 85% of learners still tended to favor learning related words; therefore, learners need to be instructed not to study related words together from the first. Moreover, learners may benefit from knowing the possible advantages and disadvantages of using these as learning strategies (Schmitt, 2000). They could use SU when they study new words for the first time, and then move to SR to review known words.

Some limitations of the present study should be addressed. First, this study focused on only several restricted aspects of word knowledge: receptive aspects or recognition of the written form. Second, the period of the treatment may not have been long enough to examine consequential perception changes, as noted in the discussion. Additionally, there is a possibility that other variables may have contributed to the recognition of this negative effect. From these limitations, suggestions for further studies can be drawn. A longer treatment is needed to ascertain the findings of present study. Moreover, other aspects of vocabulary learning and other possible variables should be accounted for before the results are generalized.
REFERENCES


341-361.


Education.


## APPENDIX
### Organization of Vocabulary Items

#### 1) Group R

<table>
<thead>
<tr>
<th>Week</th>
<th>Target vocabulary items</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>instep calf shin sole thigh&lt;br&gt;mole mustache freckle beard sideburn&lt;br&gt;2 overweight skinny slender lean chubby&lt;br&gt;cooperative stubborn timid reliable moody&lt;br&gt;3 dust rinse polish scrub vacuum&lt;br&gt;strain dice grate stir-fry whisk&lt;br&gt;4 sneak wander strut trudge toddler&lt;br&gt;paw claw fin hoof feather</td>
</tr>
</tbody>
</table>

#### 2) Group U

<table>
<thead>
<tr>
<th>Week</th>
<th>Target vocabulary items</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>instep thigh cooperative moody overweight&lt;br&gt;paw mole sneak dust strain&lt;br&gt;2 feather claw toddler wander mustache&lt;br&gt;rince stubborn skinny calf dice&lt;br&gt;3 polish vacuum sideburn freckle slender&lt;br&gt;fin timid strut shin grate&lt;br&gt;4 lean chubby whisk stir-fry sole&lt;br&gt;hoof reliable scrub beard trudge</td>
</tr>
</tbody>
</table>

Applicable levels: primary education

Key words: vocabulary learning, semantically related words, interference theory, cross-association

Shin young Bak
Department of English Language Teaching
International Graduate School of English
17 Yangjae-daero 81-gil, Gangdong-gu,
Seoul 134-847, Korea

Email: totoro@igse.ac.kr

Received in June, 2012
Reviewed in July, 2012
Revised version received in August, 2012