The Effect of Keyword Method on Vocabulary Retention of Senior High School EFL Learners in Iran

Mohammad Davoudi1* Dina Yousefi2
1. Assistant Professor of TEFL, Department of English Language and Literature, Hakim Sabzevari University, Sabzevar, Iran
2. M.A holder of Applied Linguistics, Office of Education, Ferdows, Iran

Abstract
This study aimed at investigating the effect of keyword method, as one of the mnemonic strategies, on vocabulary retention of Iranian senior high school EFL learners. Following a quasi-experimental design, the study used thirty eight (n=38) female senior high school students in grade four from two intact classes at a public high school. The students were randomly assigned to experimental and control groups. The experimental group was instructed through the keyword method and the control group learned vocabulary through the traditional method. To analyze the data, paired-samples t-test and independent samples t-tests were run. It was found that students in the experimental group significantly outperformed the students in the control group in vocabulary retention by keyword method. Also, a significant difference was found between the performance of the keyword group and traditional group in delay recall posttest. Overall, this study illustrated that the use of keyword method can largely reduce learners' problems in the acquisition and retention of L2 words. The findings of this research may have pedagogical implications for teachers and learners.

Keywords: Keyword method, mnemonic strategies, vocabulary retention, EFL learners

1. Introduction
Learning vocabulary plays an important role in language learning. As Paivio (1986) asserts, vocabulary acquisition is an important end in itself, even more significant and complicated than is, by and large, perceived by teachers of foreign languages. A problem which frustrates the language learners is the fact that they cannot recall the vocabulary items they have memorized. Waring (2002) believes that one of the most important problems with vocabulary learning which the second language learners are faced with is that what they learn today will be forgotten tomorrow. The explanation offered by Waring is that the learners may do enough for the fast comprehension but not enough for retention over time. Retention over time requires making conscious effort on the part of the learner to link the word with meaning. According to Waring (2002), our brain is designed to forget something, not to remember. It seems that we should devise methods and techniques which empower our learners to recall the words for a longer period of time and help them to permanently acquire the vocabularies.

According to Atay and Ozbulgan (2007), the learners need to be given explicit instruction of vocabulary learning strategy to facilitate their awareness of vocabulary learning strategies that they can use to learn on their own outside the classroom. The question is whether there is an effective strategy which can help the EFL learners to enhance their vocabulary learning and retention. What should our teachers do to foster vocabulary acquisition in our learners? What strategy or strategies should we implement to promote the process of language learning specifically to enhance retention of a vast number of necessary vocabulary items in EFL learning in Iran?

Among various vocabulary learning strategies, which have been supported by ample research studies such as Wyra, Lawson, and Hongi (2007), keyword method is one of the well-documented strategies. According to Henson and Eller (1999), mnemonics is one of the most efficient strategies to assist learners to remember new vocabulary. As Higbee and Kanihara (1985) point out; students mainly acquire new words by associating them with something memorable, such as a word or an object.

The first research question of the present study is to find out if using mnemonic strategy facilitates vocabulary learning and retention more than traditional ways of vocabulary learning. Another research question is related to the long terms effects of keyword method on vocabulary retention. And the last research question addresses the learners' preference with regard to traditional method or mnemonic approach of vocabulary learning. Thus, the following research questions are posed in this study.

1. Does the use of keyword method as a mnemonic strategy have a significant impact on enhancing vocabulary retention of senior high school students?
2. Does keyword method have better effects on senior high school students' vocabulary retention over a longer period of time?

2. Literature Review
Memory plays an important role in learning a new language. The difference between a successful and unsuccessful learner might be due to the workings of their memory. Therefore, the kind of remembering has a
The opinions are stored in mind on short term or long term basis. The purpose of language learning is to move things into long term memory. Unfortunately, forgetting is part of the process of trying to retrieve items from memory (Lewis, 1999).

Schmitt (1997) also states that memory strategies, traditionally known as mnemonics, are one type of consolidation strategies. They usually involve relating the word to some previous knowledge. For example, using the pictures of the word instead of definitions or linking it to some L2 words already familiar to the learner is one example of consolidating strategy. Using unrelated words or grouping the words according to some categories such as synonyms or common themes are also examples of memory strategies. In addition, words' phonological or orthography form can be used as a mnemonic strategy.

The term mnemonics is derived from the name of the goddess of memory Mnemosyne. Mnemonic is the Greek word that dates back to 500 B.C, and can be defined as any procedure that facilitates recall of pieces of new information. A mnemonic device is a learning technique that assists in memorization of information. Mnemonics can be verbal, visual, kinesthetic or auditory (Yates, 1966, cited in Amiryousefi & Ketabi, 2011).

Mnemonic strategies were first utilized in general education settings by college undergraduates who were trying to learn foreign language vocabulary (Atkinson, 1975). Mnemonic strategies promote learning and memory capacity and involve connecting newly learned items with the knowledge stored through visual and acoustic cues. To implement mnemonic strategies we do not need a wealth of extra materials or planning. According to Levin (1983), mnemonic instruction is advantageous for learners of all age groups. However, elementary learners might not be able to learn or recall as efficiently as older learners do.

Several scholars have classified mnemonics strategies differently. For example, Thomson (1978, cited in Saeedi, Ketabi, & Dastjerdi, 2011) categorized mnemonics into five groups such as, spatial, linguistic, visual, verbal, and physical response method. Baddley (1999) believes that mnemonic strategies are classified into visual and verbal method. Oxford (1990), presents four main strategies, including creating mental image, applying images and sound, employing action, and reviewing.

Mnemonic instruction is a well-documented and researched area in language teaching. However, most of the studies that have been conducted so far have investigated the effect of mnemonic instruction of students with learning disabilities. There have been researchers who tried to investigate the effect of mnemonic instruction on ordinary college students. Recently, Carney and Levin (2000) carried out a study in which college students utilized mnemonic strategy to learn and recall painting-to-artist matches. The outcome showed that those learners who utilized mnemonics considerably outperformed the learners who did not utilize mnemonics on tests that required remembering artists and their works.

In another study conducted by Levin, Shriberg, Miller, McCormick, and Levin (1980), keyword method was used by fifth graders and fourth graders to remember the US states and their capitals. The learners who used keyword methods outperformed the control group learners who did not use keyword method.

Saeedi and Mohajerian (2012) examined the comparative effectiveness of Keyword and context methods on vocabulary retention and the rate of forgetting of EFL learners. The participants of this study were 40 learners from two intact classes. Two classes were randomly assigned to the keyword and context group. The result revealed that learners recalled more vocabulary immediately after treatment, and also the rate of forgetting was more in the context group than the keyword group.

In a recent study, Marzban and Amoli (2012) studied the effect of utilizing mnemonic techniques on immediate and delayed recall of high intermediate college students. The experiment showed that mnemonic technique was significantly effective in improving recall and retention of vocabulary.

In another study, Zarei, Hasani & Keysan (2013) investigated the differences among the effects of mnemonic and mapping techniques involving the keyword method, peg word method, the loci method, argument mapping, concept mapping, and mind mapping on L2 vocabulary comprehension and production. 151 Iranian Female students from a public pre-university school in Islam Shahr participated in this study. The finding of this study indicated that the keyword method group had the third highest mean after the peg word and loci on vocabulary comprehension, but a low mean on vocabulary production test.

In a different context, Wyra, Lawson, and Hungi, (2007) implemented keyword method to compare the experimental performance group in learning new Spanish words and their English definitions to that of a control group which used standard keyword procedure. The findings showed that retrieval training can be a significant indicator of both backward and forward recall performance.

Chen (2008) studied the effect of keyword method on ESP vocabulary instruction with the use of quasi-experimental and open-ended questionnaire. Forty participants from two intact classes in a university in Taiwan were randomly assigned as the keyword and traditional method group. The keyword group received the keyword strategy while the traditional group concentrated on learning words by presenting definitions. The results showed that the keyword group recalls more target vocabulary than traditional group.

Ashoori (2012) also compared the effects of keyword, context, and word list instructional strategies on long-term vocabulary recall of 65 female students from Kish Institution in Tehran. After one week treatment a
posttest in two steps; cued recall and word recall were applied to evaluate the effectiveness of the techniques on retention of target vocabulary. The findings of this study showed that keyword group was better than the other groups in both cued recall test and word recall test, and there was no difference between the context and word list groups.

In Anjomanafrooz and Tajallis' (2012) paper, the effects of using mnemonic association on vocabulary recall of Iranian EFL learners were discussed in two separate experiments with adults and adolescents. In each experiment, the participants were divided into two groups of experimental (mnemonic) and control (rote). The data indicated that adult learners in mnemonic association group had better performance than the rote group, while results in adolescent group in mnemonic association had not been effective in comparison with rote group.

3. Method

3.1 Participants

The participants of this study initially consisted of thirty eight (n = 38) female senior high school students in grade four from two intact classes at AL Zahra public high school and Enghelab high school in Sabzevar, ranging in age from 17 to 18. All of them were native speakers who had learned English for six years; three years in guidance school and three years in high school. In addition, they had the same English teacher. The reason why grade four students were selected was that they were preparing to pass the university entrance examination. Therefore, a vocabulary class was planned either for the purpose of this study or to help them in expanding their lexical knowledge. The students were randomly assigned to experimental and control groups. Twenty students were placed in the experimental keyword group and eighteen students were put in the control group. The experimental group was instructed through the keyword method and the control group learned vocabulary through the traditional method.

3.2 Instrumentation

The materials and data collection instruments that were used in this study are as follows:

3.2.1 Word knowledge pre-test

Word knowledge pre-test used in this study consisted of 60 vocabulary items selected from the Oxford elementary dictionary and coursebook that were contextualized in 60 sentences. The target words were contextualized in sentences to form a word knowledge pre-test. The target words were bolded in each sentence and the students were asked to write the definition of the words in Persian. Longman dictionary and teacher-made sentences were used to get samples of the sentences containing these words. For the test performance, scores were tabulated out of 30. One point was assigned for each correct answer. The aim of this study was to elicit unknown words and to make sure that students were unfamiliar with these words (see appendix B).

3.2.2 Post-test

At the end of the experimental period, open-ended format test including 30 items was used as a comprehension post-test to assess the effect of the selected presentation technique on vocabulary retention. In this test, the students were asked to write the definition of each word in Persian and the participants in the experimental group were asked to write the keywords in addition the meaning in Persian (Appendix C).

3.2.3 Vocabulary delayed recall posttest

Three weeks later a surprise recall test that was same with posttest was given to two groups so that the researcher could examine the effect of both methods on vocabulary retention and long term memory (Appendix D).

3.3 Research Design

The design of the study was of a quasi-experimental type. The study was carried out within a 2 month period during which the students attended the classes for 8 sessions. The first session was allocated to briefing the learners about how the classes would go through to the end. A pretest was given to the participants to see if they belonged to the same level of vocabulary proficiency. Then, they were divided into two experimental and control groups each of which were taught the same lessons with the difference that the experimental group received the mnemonic instruction while the control group were taught through the traditional learning method. In this study the independent variable was keyword method, the learners' vocabulary scores was the dependent variable.

3.4 Procedure

The keyword method technique was used to teach vocabulary items presented in grade four senior high school books and Oxford Dictionary. Initially, 38 students with the above mentioned characteristics participated in the present study. Before introducing the treatment, a standard 30 minutes pre-test including 30 items in multiple-choice format was administered to check the homogeneity of the participants and their language proficiency. Then, the word knowledge pre-test (checklist) was given to ensure that the students did not have any prior knowledge of the target words. The checklist included 60 bolded vocabulary items which were contextualized in
60 sentences. The words were chosen from Oxford dictionary and Pre-university English coursebook. The time for pretest was 25 minutes. After the pretest the words that were familiar for the students were deleted and unknown words were selected for the post test.

After word knowledge pre-test, the students were assigned to two groups and each group was randomly assigned to either experimental or control group. In the first session, the researcher presented a full explanation of the selected technique (keyword method) to the participants in the experimental group. The researcher prepared some keywords that correlate with the words available in the textbook. The treatment lasted for eight sessions; each session was held once a week and lasted for 25 minutes. After the learners were introduced to the new words and their corresponding definitions, they were given the keyword method strategy study sheet to complete in their own ways. That is, learners themselves were required to help to create a corresponding keyword for each vocabulary item. Each session seven to eight words were taught using this strategy. In the control group, the materials were taught in the traditional method using translation and definition and memorization strategies.

At the end of the program, two groups were tested on a vocabulary test which incorporated the vocabulary item covered during the period. In addition, students were given weekly quizzes to check their progress. Also, three weeks later a vocabulary delayed recall posttest was conducted to check the students' retention of the target words and to see the effect of keyword method on long-term memory, and finally in sixteenth session they were given a questionnaire to express their opinion on the effectiveness of the techniques.

4. Results and Findings

4.1 The results of the English proficiency between two classes

Before conducting the study, it was essential to compare the English proficiency level of the students in two different classes to be sure two classes were at the same level of proficiency. Therefore, the researcher took a proficiency test.

Table 4.1
Descriptive Statistics for English proficiency test

<table>
<thead>
<tr>
<th>N</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Std. Error mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>7.00</td>
<td>27.00</td>
<td>15.600</td>
<td>6.56466</td>
<td>1.46790</td>
</tr>
<tr>
<td>18</td>
<td>9.00</td>
<td>27.00</td>
<td>15.889</td>
<td>5.01631</td>
<td>1.18236</td>
</tr>
</tbody>
</table>

Table 4.1 makes it clear that the mean score of the class A reached 15.60 and mean score of the class B was 15.88. To see whether this slight difference between two groups was statistically significant or not, a t-test was run.

Table 4.2
Independent Samples T- Test results for English proficiency between two classes

<table>
<thead>
<tr>
<th>Group</th>
<th>Levine's Test for Equality of Variances</th>
<th>t-test for Equality of Means</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F</td>
<td>Sig.</td>
</tr>
<tr>
<td>Class A Vs. Class B</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Equal variances assumed</td>
<td>2.738</td>
<td>.107</td>
</tr>
<tr>
<td>Equal variances not assumed</td>
<td>-.153</td>
<td>35.126</td>
</tr>
</tbody>
</table>

As Table 4.2 shows, the estimated level of significance (α) is more than probability value (α= .881>.05). The assumption of homogeneity of variances was proved. Therefore, the results of t-test ( t= .151, sig= .881>.05) indicated that there is no significance differences between two classes from overall English proficiency point of view.

4.3 The results of word knowledge pretest

To make sure that participants in each group had the same prior knowledge about the target words, the researcher
administered the pretest. Therefore, the data was analyzed by the independent sample t-test. The findings are shown in Tables 4.3 and 4.4.

**Table 4.3**
**Group Statistics for word knowledge pretest**

<table>
<thead>
<tr>
<th>Test</th>
<th>Group</th>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Std. Error Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word knowledge pretest</td>
<td>Exp</td>
<td>20</td>
<td>3.100</td>
<td>2.65370</td>
<td>.59338</td>
</tr>
<tr>
<td></td>
<td>cont</td>
<td>18</td>
<td>3.000</td>
<td>2.16930</td>
<td>.51131</td>
</tr>
</tbody>
</table>

**Table 4.4**
**Independent Samples Test for Word knowledge pretest**

<table>
<thead>
<tr>
<th>Word knowledge pretest</th>
<th>Levine's Test for Equality of Variances</th>
<th>t-test for Equality of Means</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F</td>
<td>Sig.</td>
</tr>
<tr>
<td>Equal variances not assumed</td>
<td>.128</td>
<td>35.695</td>
</tr>
</tbody>
</table>

As shown in Table 4.4, the computed F was 1.523 and the level of significance was Sig= 0.225 > 0.05. The assumption of equal variances was proved. The results of t-test demonstrated that there was no significance difference between two groups on the word knowledge pretest (t= 0.126, Sig= 0.90 > 0.05). Furthermore, from the two mean scores in Table 4.3 (M=3.10, M= 3.00) the researcher concluded that those to be learned words were almost new to all the participants.

**4.4 The results of the posttest of two groups**

**Table 4.5**
**Descriptive statistics for posttest results for both group Statistics**

<table>
<thead>
<tr>
<th>Test</th>
<th>group</th>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Std. Error Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Posttest</td>
<td>experimental</td>
<td>20</td>
<td>22.850</td>
<td>4.71587</td>
<td>1.05450</td>
</tr>
<tr>
<td></td>
<td>control</td>
<td>18</td>
<td>15.944</td>
<td>7.22310</td>
<td>1.70250</td>
</tr>
</tbody>
</table>

Table 4.5 indicates that there is a noticeable difference between two groups in posttest. The mean of the keyword group (M= 22.85) was higher as compared to the traditional group (M= 15.94). In order to find out whether this difference was significant or not, a T-test was run.

**Table 4.6**
**Independent Samples t-test for Posttest**

<table>
<thead>
<tr>
<th>Group</th>
<th>Levine's Test for Equality of Variances</th>
<th>t-test for Equality of Means</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F</td>
<td>Sig.</td>
</tr>
<tr>
<td>Keyword Vs. Traditional Equal variances assumed</td>
<td>7.434</td>
<td>.010</td>
</tr>
</tbody>
</table>

The results presented in Table 4.6 shows that (t= 3.52, Sig= 0.001), the training of mnemonic keyword
method was effective, that is to say, there was statistically significant difference between the performances of the two groups.

4.5 The results of delayed recall posttest of two groups

Table 4.7
Descriptive Statistics results of Delayed recall Posttest of two groups

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Std. Error Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Keyword Group</td>
<td>20</td>
<td>7.00</td>
<td>28.00</td>
<td>17.1500</td>
<td>6.05479</td>
<td>1.35389</td>
</tr>
<tr>
<td>Traditional Group</td>
<td>18</td>
<td>1.00</td>
<td>22.00</td>
<td>8.4444</td>
<td>6.22351</td>
<td>1.46689</td>
</tr>
</tbody>
</table>

Table 4.8
Independent Samples t-test of two groups

<table>
<thead>
<tr>
<th>Delayed recall posttest</th>
<th>Levine's Test for Equality of Variances</th>
<th>t-test for Equality of Means</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F</td>
<td>Sig.</td>
</tr>
<tr>
<td>Kwg Vs. Trg</td>
<td>Equal variances Assumed</td>
<td>4.368</td>
</tr>
<tr>
<td></td>
<td>Equal variances not assumed</td>
<td>4.361</td>
</tr>
</tbody>
</table>

Kwg= Keyword group, Trg= traditional group, std= standard deviation.

The results of t-test in Table 4.8 indicated that there was a significant difference between two groups in delayed recall posttest (t= 4.36, Sig= 0.000). Also, the mean score of the keyword group as shown in Table 4.7 (M= 17.1500) was higher than that of traditional group which was (M= 8.44). Hence, it is obvious that the keyword group performed significantly better than traditional group in the delayed recall test. Since the keyword group outperformed participants of the traditional group in this test, it was concluded that keyword method had long-term effects on vocabulary retention.

4.6 Investigation of Research Questions

The research questions posed in the beginning of this study are restated here for the sake of ease of interpretation and analysis.

1. Does keyword method as a mnemonic strategy have a significant impact on enhancing vocabulary retention of senior high school students?
2. Does keyword method have better teaching effects on senior high school students’ vocabulary retention over a longer period of time?

4.6.1 Investigation of the first research question

The first research question aimed to discuss the effect of mnemonic keyword method on L2 vocabulary retention and as compared to the traditional method. Looking at the obtained data of the study, it was found that the difference between keyword group and the traditional group reached a significant level (Sig=0.001). Furthermore, the mean score of the keyword group was (M=22.85) higher than that of traditional group which was (M=15.94). Hence, the keyword group performed better than the traditional group. That is to say, the keyword method has a significant impact on students’ vocabulary retention than traditional method. Thus, the first null hypothesis is rejected.

4.6.2 Investigation of the second research question

The aim of the second research question was to see long-term effects of using keyword method on memory. A close look at Tables 4.6 and 4.7 shows that there was statistically significant difference between the experimental group and control group in their delayed recall posttest (Sig=0.000). On the other hand, the observed mean score of the experimental group also revealed that the difference between the two groups was meaningful, indicating that the keyword group functioned better than the traditional group. In short, the keyword method as mnemonic strategy had better instructional effects on senior high school students’ vocabulary retention than traditional method over a longer time period. Thus, the second null hypotheses is also rejected.
4.7 Discussion
The major finding of the analysis as shown in Tables 4.1-4.8 indicate a positive answer to the main questions of the study. It was found that the keyword mnemonic method had a positive effect on students' vocabulary retention. This was proved through higher mean scores that the experimental group obtained in the posttests. Specifically, the experimental group performance was more differentiated than that of the control group in posttests. Furthermore, the pretest results for both groups did not show any significant difference between the two groups. This means that before the application of the experiment they had nearly similar proficiency level. That is to say they had the same language background.

The findings of the current study as mentioned before, showed a significant difference between keyword group and traditional group. These findings of this study are in line with many previous studies which compared the keyword method with other vocabulary learning strategies (such as Zarei & Salimi, 2012) with translation method (such as Avila & Sadoski, 1996), with other mnemonic techniques, including peg word, loci, visual, etc. (such as Richmond, et. al, 2008). In addition, these results are in line with the study applied by Baleghizadeh & Ashori (2010) which indicated the powerful impact of the keyword method on learners' memory in recalling word definition.

As it was mentioned, the result of the present study illustrated the beneficial effect of the keyword method on the subjects' retention of the words that made a significant difference between keyword group and traditional group. This difference between two groups could be attributed to many reasons. The first reason was providing visual imagery. According to Shapiro and Water (2005, cited in Salimi & Zarei, 2012), by providing interactive images, the keyword method provides visual stimuli that lead to the better retention than other kinds of stimuli.

The second reason for the beneficial value of the keyword method as a mnemonic technique was that the students in experimental group received keyword method and learned vocabulary by the creation of links between new information and subjects' schemata. According to Lawson & Hogben (1996), this reason is one of the most significant reasons for the success of the keyword method in vocabulary acquisition. Positive attitudes and beliefs of the learners with respect to the value of keyword method in vocabulary learning was another reason for the better performance of the participants in the experimental group.

5. Conclusion
The current study investigated the effect of keyword method as a mnemonic strategy on vocabulary retention of EFL learners. The data from this study indicated that students in the experimental group significantly outperformed the students in the control group in vocabulary retention by keyword method. The finding of this study is in line with some previous research findings (e.g., Win, 2012) concerning the impact of keyword technique on the students' vocabulary retention ability. In other words, the treatment offered to the experimental group affected this group to some extent. Therefore, the first null hypotheses stating that mnemonic strategy (keyword) has no impact on enhancing vocabulary retention was rejected.

The statistical analysis of the second research question showed that there is a significant difference between the performance of the keyword group and traditional group in delay recall posttest. Thus, the second null hypotheses stating that keyword method has no effect on vocabulary retention over a longer time period was rejected as well. Overall, this study illustrated that the use of keyword method can largely reduce learners' problems in the acquisition and retention of L2 words.

The findings of this research may have pedagogical implications for teachers and learners. Since the keyword method is more beneficial to students' vocabulary learning, teachers should stimulate students' motivation for this method and various vocabulary learning strategies with the understanding that there is a need to improve self-instruction in the educational system. From an educational perspective, it is also beneficial for educators to apply vocabulary training techniques such as mnemonic techniques specially keyword method to provide ground for logical understanding of specific contexts, different materials, and issues.

REFERENCES
828.