

EFFECTS OF THE TEACHER-PROVIDED AND STUDENT-GENERATED  
KEYWORD METHODS ON THE IMMEDIATE AND DELAYED RECALL AND  
RECOGNITION OF VOCABULARY ITEMS UNDER CLASSROOM CONDITIONS  
AT A TURKISH UNIVERSITY

A THESIS PREPARED BY  
BENVA YAVUZ

TO

THE INSTITUTE OF EDUCATION AND SOCIAL SCIENCES  
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS  
FOR THE DEGREE OF MASTER OF ARTS  
IN TEACHING ENGLISH AS A FOREIGN LANGUAGE

BILKENT UNIVERSITY  
SEPTEMBER, 1995

PE  
1068  
.T8  
Y39  
1995

EFFECTS OF THE TEACHER-PROVIDED AND STUDENT-GENERATED  
KEYWORD METHODS ON THE IMMEDIATE AND DELAYED RECALL AND  
RECOGNITION OF VOCABULARY ITEMS UNDER CLASSROOM CONDITIONS  
AT A TURKISH UNIVERSITY

A THESIS PRESENTED BY  
DERYA YAYLI  
TO  
THE INSTITUTE OF ECONOMICS AND SOCIAL SCIENCES  
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS  
FOR THE DEGREE OF MASTER OF ARTS  
IN TEACHING ENGLISH AS A FOREIGN LANGUAGE

BILKENT UNIVERSITY  
SEPTEMBER, 1995

Derya Yılmaz  
tarafından teğışlenmiştir.

PE  
1068  
.T8  
430  
1935

B031342

## ABSTRACT

Title: Effects of the teacher-provided and student-generated keyword methods on the immediate and delayed recall and recognition of vocabulary items under classroom conditions at a Turkish university

Author: Derya Yayli

Thesis Chairperson: Ms. Susan D. Boshier, Bilkent University, MA TEFL Program

Thesis Committee Members: Dr. Phyllis L. Lim, Ms. Bena Gul Peker, Bilkent University, MA TEFL Program

This experimental study aimed at investigating the effects of a teacher-provided keyword method and a student-generated keyword method on recall and recognition of vocabulary items in comparison to a rote rehearsal control group. The study was conducted at Middle East Technical University Department of Basic English. The participants were 47 intermediate-level students in three intact classes.

Research questions focused on the differences, if any, among the three groups in terms of recall and recognition of vocabulary items, as well as the differences, if any, with respect to retention of the learned vocabulary after treatment.

To answer the research questions, three intact classes were chosen from the Middle East Technical University. One class was instructed in the teacher-provided keyword method and a second in the student-generated keyword method. The third group, which served as the control group, was asked to learn the words by rote rehearsal. Each group was allowed 25 minutes to learn the same 20 target vocabulary items.

Before the treatment, each group was given a pretest, and the same test was given as a posttest immediately after treatment to test immediate recall and recognition. To measure long-term retention, that is delayed recall and recognition, the same test was given to the three groups after two weeks. These tests included separate recall and recognition sections.

Pretest and posttest scores of recall and recognition were compared to measure acquisition of vocabulary items. Later, the posttest and retention test scores were compared to measure long-term retention. For the analysis, means and standard deviations were calculated and one-way analyses of co-variance were applied for each comparison of both recall and recognition separately to see if any method was significantly different from the others. Results did not show any statistically significant difference among the groups for the three tests of immediate and delayed recall and recognition of vocabulary. Treatment worked equally very well for all groups and posttest scores for all groups were very high. However, those groups did not differ significantly from each other.

Findings suggest that the keyword strategy, whether provided by the teacher or generated by the students, is not superior to rote rehearsal for either recall or recognition at immediate or delayed testing in university classroom conditions.

BILKENT UNIVERSITY  
INSTITUTE OF ECONOMICS AND SOCIAL SCIENCES  
MA THESIS EXAMINATION RESULT FORM

August 31, 1995

The examining committee appointed by the  
Institute of Economics and Social Sciences for the  
thesis examination of the MA TEFL student

Derya Yayli

---

has read the thesis of the student.  
The committee has decided that the thesis  
of the student is satisfactory.

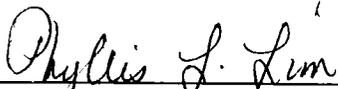
Thesis Title : Effects of the teacher-provided and  
student-generated keyword methods on  
the immediate and delayed recall and  
recognition of vocabulary items  
under classroom conditions at a  
Turkish University

Thesis Advisor : Dr. Phyllis L. Lim  
Bilkent University, MA TEFL Program

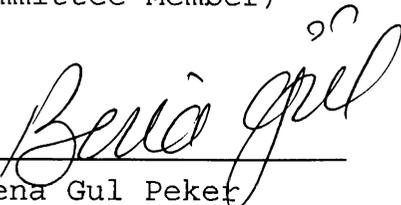
Committee Members : Ms. Susan D. Boshier  
Bilkent University, MA TEFL Program

Ms. Bena Gul Peker  
Bilkent University, MA TEFL Program

We certify that we have read this thesis and that in our combined opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Arts.

  
\_\_\_\_\_  
Phyllis L. Lim  
(Advisor)

  
\_\_\_\_\_  
Susan D. Boshier  
(Committee Member)

  
\_\_\_\_\_  
Bena Gul Peker  
(Committee Member)

Approved for the  
Institute of Economics and Social Sciences

  
\_\_\_\_\_  
Ali Karaosmanoglu  
Director  
Institute of Economics and Social Sciences

## TABLE OF CONTENTS

LIST OF TABLES . . . . .	.viii
CHAPTER 1 INTRODUCTION . . . . .	1
Background of the Problem. . . . .	1
Purpose of the Study . . . . .	6
Significance of the Study. . . . .	6
Research Questions . . . . .	6
CHAPTER 2 LITERATURE REVIEW. . . . .	8
Immediate and Delayed Effects of the Keyword Method . . . . .	8
Effects of the Keyword Method on Recognition Teacher-Provided versus Subject-Generated Keyword Method . . . . .	10
Effects of Teacher-Provided and Student- Generated Keyword Methods on Immediate and Delayed Vocabulary Recall. . . . .	14
Conclusion . . . . .	17
CHAPTER 3 METHODOLOGY. . . . .	19
Introduction . . . . .	20
Research Design. . . . .	20
Subjects . . . . .	20
Instruments/Materials. . . . .	21
Instructional Material. . . . .	22
Testing Material. . . . .	22
Procedure. . . . .	23
Information and Consent . . . . .	24
Pretest . . . . .	24
Practice. . . . .	24
Study . . . . .	25
Immediate and Long-Term Retention Test. . . . .	26
Data Analysis. . . . .	26
CHAPTER 4 RESULTS. . . . .	27
Overview of the Study. . . . .	28
Overview of Analytical Procedures. . . . .	28
Results of the Study . . . . .	29
Vocabulary Recall . . . . .	30
Vocabulary Recognition. . . . .	30

CHAPTER 5 DISCUSSION OF FINDINGS AND CONCLUSION. . . . .	.39
Summary of the Study . . . . .	.39
Discussion of the Results. . . . .	.40
Limitations of the Study . . . . .	.43
Implications for Further Research. . . . .	.45
Pedagogical Implications . . . . .	.46
REFERENCES . . . . .	.48
APPENDICES . . . . .	.52
Appendix A: Informed Consent Form. . . . .	.52
Appendix B: Vocabulary Recall Pretest. . . . .	.53
Appendix C: Vocabulary Recall Posttest . . . . .	.54
Appendix D: Vocabulary Recall Retention Test .55	.55
Appendix E: Vocabulary Recognition Pretest . .56	.56
Appendix F: Vocabulary Recognition Posttest. .58	.58
Appendix G: Vocabulary Recognition Retention Test . . . . .	.60
Appendix H: Study List for the Student- Generated Group. . . . .	.62
Appendix I: Study List for the Teacher- Provided Group . . . . .	.63
Appendix J: Study List for the Rote Rehearsal Group. . . . .	.64
Appendix K: Vocabulary List for the Practice Session. . . . .	.65

## LIST OF TABLES

<u>TABLE</u>	<u>PAGE</u>
1 Means and Standard Deviations of Pretest and Immediate Posttest Scores of All Groups for Vocabulary Recall. . . . .	30
2 Results of ANCOVA: For Immediate Vocabulary Recall .	31
3 Means and Standard Deviations of Immediate Posttest and Long-Term Retention Test of All Groups for Vocabulary Recall. . . . .	32
4 Results of ANCOVA: For Delayed Vocabulary Recall (Long-Term Retention) . . . . .	33
5 Means and Standard Deviations of Pretest and Immediate Posttest Scores of All Groups for Vocabulary Recognition. . . . .	34
6 Results of ANCOVA: For Immediate Vocabulary Recognition . . . . .	35
7 Means and Standard Deviations of Immediate Posttest and Long-Term Retention Test of All Groups for Vocabulary Recognition . . . . .	36
8 Results of ANCOVA: For Delayed Vocabulary Recognition (Long-Term Retention). . . . .	37

## CHAPTER 1 INTRODUCTION

### Background of the Problem

In recent years attention has shifted to the role of the language learner as an active participant in the teaching-learning process. It has been shown that what the learner knows and what the learner thinks about during learning affect the teaching-learning process (Weinstein & Mayer, 1986).

Different learners have been shown to follow different ways in learning a second or a foreign language. Rubin (1987) states that some learners approach the learning task in more successful ways than others. In the classroom, although every student is exposed to the same teaching material, they learn differently and some language learners are more successful than others. Successful learners manipulate their learning processes in many ways, and they are the ones who know how to learn (Rubin, 1987). Successful learners have learning strategies of their own, and they are more autonomous than others. Thus, the aim of teaching learning strategies, according to Wenden (1987), is to create an autonomous learner. According to Weinstein and Mayer (1986), "Good teaching includes teaching students how to learn, how to remember, how to think, and how to motivate themselves" (p. 315). Dickenson (cited in Skehan, 1989) states that strategy training aims at aiding the learner to move toward decision making and autonomy. If we can teach our students how to learn, how to solve problems, and how to

retrieve the learned information, we can be proud of ourselves for creating autonomous learners (Weinstein & Mayer, 1986).

Retrieving learned information calls for teaching the art of memory. Autonomous learners can make optimal use of memory clues to remember the learned information. Thompson (1987) states that we cannot think of any educational goal neglecting the importance of the ability to retain information. Human memory is crucial to learning and, therefore, to learning linguistic skills. Thompson claims that research on the role of memory in second language learning is something relatively new.

One important area that has to do with memory is vocabulary learning in a second language. It is not uncommon for language teachers to hear students say that they cannot learn vocabulary easily or that they forget the new vocabulary items very soon after studying them. There are learning strategies specific to vocabulary and these vocabulary learning strategies are largely based on mnemonic techniques which help individuals learn faster and recall better because they provide learners with useful retrieval cues (Thompson, 1987). Higbee (cited in Thompson, 1987) states that mnemonic means "aiding memory". Mnemonics can be adopted voluntarily and provide long-term retention (Thompson, 1987). Different learners can use different

types of mnemonic techniques, and Levin (cited in Thompson, 1987) states that many learners enjoy using them.

The keyword method is the best known vocabulary learning strategy and it has been researched extensively (Thompson, 1987). The keyword method is discussed as a cognitive strategy by O'Malley and Chamot (cited in Brown, 1994). It is a mnemonic procedure which has three stages. The first stage includes linking a target vocabulary item in the second language with a keyword in the native tongue which has an acoustic or orthographic similarity. In the second stage, the learner forms a stable association between the target word and the keyword. Finally, the third step requires formation of an image that includes referents of the keyword and the target word. For example, to learn the English word "tie", the Turkish word "tay" (pony), which has an acoustic similarity, is chosen as the keyword. The meaning of the word "tie" ("kravat" in Turkish) and Turkish keyword "tay" are combined in a pictorial image like "kravat takan bir tay" (a pony wearing a tie) to aid recall of the meaning of target word "tie". When the learners hear the word "tie", they recall the Turkish keyword "tay" and which is associated to "kravat takan bir tay" in order to link pronunciation of the word "tie" to the image. The target word, which is a part of the image, is thus retrieved easily. The keyword method was developed by Atkinson (1975)

and studies have shown that it is an effective strategy that helps learners develop their vocabulary in language learning (e.g., Pressley, Levin, & Delaney, 1982; Raugh & Atkinson, 1975; Raugh, Schupbach, & Atkinson, 1977; Wang, Thomas, & Ouellette, 1992).

There are four areas to which little research has been devoted or in which findings are mixed. First, long-term retention of the learned vocabulary has been researched in the keyword studies, but findings are mixed. Second, whether the keywords should be supplied by the teacher or should be generated by the learners is still a question which requires more research because the findings are inconsistent (Pressley, Levin, & Delaney, 1982). Third, the effect of the keyword method on the recognition of vocabulary items has been little studied. Finally, little research has been conducted under classroom conditions where language teaching and learning process occurs.

Some studies have shown that the keyword method provides long-term retention as well as short-term retention (e.g., Acikgoz, 1992; Raugh & Atkinson, 1975) whereas some have demonstrated that long-term forgetting is greater in keyword method learners (e.g., Wang, Thomas, & Ouellette, 1992). However, little research has been conducted to study the effects of two types (teacher-provided and student-generated) of keyword method on both short-term (immediately

after training) and long-term retention. Hall, Wilson and Patterson (1981) studied 60 undergraduates with no command of the target language, Spanish. Although the experimenters expected that short-term and long-term retention (one week) would be better accomplished when students generated their own keywords, what they found was just the opposite. Performance of the students in the student-generated keyword group was inferior to both the teacher-provided and the control groups in both short-term and long-term recall.

Most studies on the keyword method have tried to investigate the effect of the method on recall rather than on recognition. The experimenters of most studies were after the translation (recall) of the target vocabulary items. Little research has been carried out to investigate the effects of the keyword method on recognition. However, McDaniel and Pressley (1989) found that the keyword method had positive effects on recognizing the vocabulary items in a reading passage when subjects were immediately tested. Brown and Perry (1991) found that a combined keyword-semantic method was better on recognition than the semantic method or the keyword method alone.

Finally, although vocabulary is considered to be a discrete skill, it is important for both for comprehension and production in the classroom. However, most studies have been conducted in laboratory settings, which provides little

help in understanding the applicability of the method to the classroom setting. Levin, Pressley, McCormick, Miller, & Shriberg (1979) found that the keyword method was effective in classroom setting when necessary conditions (motivational, attentional and instructional) were provided to students.

Thus, more research is needed to investigate which type of keyword method (teacher-provided versus student-generated) provides better immediate and delayed recall and recognition of vocabulary items in a classroom setting.

#### Purpose of the Study

In this experimental study, the effects of the teacher-provided and student-generated keyword methods on immediate and delayed recall and recognition of vocabulary items were investigated in a Turkish university classroom setting.

#### Significance of the Study

Because vocabulary learning is a major part in language learning, it is important to learn which of these mnemonic keyword methods can most effectively facilitate this process. This information would be valuable for teachers so that they can teach their students these techniques.

#### Research Questions

Research questions of this study are:

1. Is one type of the keyword method (student-generated or teacher-provided) superior to the other in immediate

recall of vocabulary compared to the control (rote rehearsal) group?

2. Is one type of the keyword method superior to the other in delayed recall of vocabulary compared to the control group?

3. Is one type of the keyword method superior to the other in immediate recognition of vocabulary items compared to the control group?

4. Is one type of the keyword method superior to the other in delayed recognition of vocabulary compared to the control group?

## CHAPTER 2 LITERATURE REVIEW

Studies on the effectiveness of the keyword method as a vocabulary learning strategy have been conducted since the early 1970s. Most of these studies have shown that the keyword method is an effective strategy of vocabulary learning (e.g. Atkinson, 1975; Pressley, 1977; Pressley et al., 1980) and superior to other vocabulary learning strategies like rote rehearsal and context method (e.g., McDaniel & Pressley, 1984; Pressley, Levin & Delaney, 1982). However, little research has been dedicated to the effects of the experimenter-provided and subject-generated keyword methods on both immediate and delayed recall of vocabulary items in the classroom setting. Moreover, little research has focused on the effects of the keyword method on recognition, as well as recall, of vocabulary items.

## Immediate and Delayed Effects of the Keyword Method

The literature that exists on the immediate and delayed effects of the keyword method shows mixed results. For example, in a study that was conducted in Turkey (Acikgoz, 1992), the experimenter studied the immediate and delayed effects of the keyword method on the acquisition of meaning, spelling, and pronunciation of English words. Subjects were secondary school students (90 sixth-graders and 90 eighth-graders) learning English as a foreign language. To-be-learned words were 50 English words. Vocabulary acquisition

was tested through a recall test in which students were asked to write the Turkish meanings of the English words. In the treatment sessions subjects were informed about the keyword method. Then they were presented with the target words and the keywords. The visual imagery for each word and keyword was presented through pictures. The control groups learned the vocabulary items via rote rehearsal. The treatment lasted three sessions and the recall test was given before and after the treatment and four weeks later to measure long-term retention. Acikgoz found that the keyword method was superior to the rote rehearsal method in both immediate and delayed recall of vocabulary units. In the pronunciation exam, subjects were asked to read English words aloud and each correct pronunciation was counted to obtain the pronunciation score. Differences observed between the control and experimental groups were not statistically significant in the pronunciation exam.

Wang, Thomas and Ouellette (1992) also compared the keyword method and rote rehearsal strategies in four experiments to measure immediate and delayed definition recall with different results. Subjects were students of psychology courses in a university in the United States. In the first experiment, the logic of the keyword method was explained to the subjects in the two keyword groups (immediate and delayed). They were later presented with 22

French words and their keywords and asked to study those vocabulary items. Subjects in the keyword groups were allowed 10 seconds for each item. The two rote learning groups (immediate and delayed) learned how to remember vocabulary items with the help of rote memorization. They were later given the list and 10 seconds to study each word. After the study time, both immediate groups took a recall test in which they were supposed to write the English meaning of each French word. The delay groups took the same exam after one week. Although the keyword group scored better on the immediate test, both delay keyword and rote learning groups scored almost the same in the delayed exam. In other words, information loss for the keyword groups was great in the delayed exam. The other three experiments with similar design and procedure showed that the keyword method did not provide superior long-term retention as expected.

#### Effects of the Keyword Method on Recognition

Most studies of the keyword method have looked at its effect on the recall of target vocabulary items. However, McDaniel and Pressley (1989) studied the effect of the keyword method on both recall and recognition of vocabulary items--recognition as measured by reading comprehension--compared to a no-strategy control group and a semantic context group. Subjects were 75 students enrolled in a psychology course at the University of Notre Dame who were

randomly assigned to three groups. The target vocabulary items were 45 unfamiliar Old English words and their definitions taken from an earlier McDaniel and Pressley (1984) study. Subjects were informed about the vocabulary teaching sessions but not about the tasks following the sessions. The control group subjects were asked to do anything they wanted to remember the words. For the context group, each word was presented in a three-sentence context and subjects were required to guess the meaning via the context clues. The keyword group subjects received keywords with the target words and were asked to create images including the meanings of both the target words and the keywords. Two stories including the target vocabulary items were developed to test and compare vocabulary recognition of the three groups. For each story, embellished (with context clues) and unembellished (without any context clues) formats were used. A 15-question true-false test followed the reading activity to measure reading comprehension of the three groups. McDaniel and Pressley found that the context method did not facilitate subsequent reading comprehension (via vocabulary recognition), and that the keyword method facilitated comprehension more than context and control groups. The context and control groups performed almost the same in the reading comprehension exam. All groups took another exam which aimed to measure the recall of the target

vocabulary items. In the recall exam, subjects were expected to write the definitions of the Old English words. In recall of definitions, the keyword group outperformed the other groups. There was a slight difference between the context group and control group in unembellished story test (the context group was better), but in the embellished story test, the context group did slightly better than the control group. McDaniel and Pressley stated that keyword method was superior for vocabulary recognition as measured by reading comprehension in a written text. For the recall of the same vocabulary items, the keyword method also outperformed the rote rehearsal and context methods.

The effects of the keyword method, semantic processing method and a combined keyword-semantic method were compared to observe their effects on immediate and delayed recall and recognition of vocabulary items by Brown and Perry (1991) in actual classroom settings. (Most keyword studies have been done in laboratory settings.) The main emphasis in the semantic processing method is on the semantic association between the new word and its definition. Anything which ties the meaning of the new word into existing knowledge structures is semantic processing. For the experiment, six intact classes from the English Language Institute at the American University of Cairo were chosen. Three classes were lower-level and three classes were upper-level with

each class at a level receiving one of the three methods: keyword only, semantic only, and combined keyword-semantic. The target vocabulary items were 40 unfamiliar nouns and verbs mostly from textbooks. Both keyword classes at each level were provided with the new word, its definition and an Arabic keyword. Semantic classes were given the new word, its definition, two examples of the word's use in sentences, and a question which they were required to answer by using the new word. The combined keyword-semantic classes received the new word, its definition, the keyword, and example sentences and question. The students in each class received a day of instruction about their methods, and they practiced the methods the next day. The day after the instruction, to measure vocabulary recall, subjects were asked to write the English definition of each word listed on the test paper. For recognition, a four-choice multiple-choice test was constructed. The results of recognition and recall tests indicated that the combined keyword-semantic method provided more retention than the other strategies. That is, retention was aided by the combined semantic-keyword method for both recognition and recall of target vocabulary items. The keyword method alone was better than the semantic method for only lower proficiency students. For the upper level group, the semantic processing method was better than the keyword method alone.

### Teacher-Provided versus Student-Generated Keyword Method

Most studies of the keyword method have neglected to study the effects of teacher-provided versus student-generated keywords. However, Hall (1988) studied the difference between student-generated and teacher-provided cases of keyword training in three experiments. In Experiments 1 and 2, subjects were asked to generate their own keywords whereas subjects were provided with the keywords in Experiment 3. Later the scores of the subjects Experiment 3 were compared to the scores of the same subjects in Experiment 2 to see which case was superior to the other. The first experiment tried to find the best type of presentation of the keywords. In the first experiment, conducted at Northwestern University in the United States, 12 subjects were chosen who were unfamiliar with the German language and with mnemonics. That is to say, these subjects had not been taught any techniques which would help them remember vocabulary items. Four 30-pair lists of concrete vocabulary items with their definitions were chosen from a pool generated by a student of German. There were two conditions under which lists were presented. For one condition, there was only one presentation of items and for the other condition items were presented four times. Keyword training was conducted in biweekly sessions over four weeks. In the first hour, there were explanations and

demonstrations, and then two hours were devoted to practice generating keywords and images. The subjects tried to find concrete keywords for the target vocabulary items. It was found that the second condition (keyword-four repetitions) was superior to the first condition (keyword-one repetition) before treatment; that is four repetitions worked better than one before the keyword method was taught. Subjects in the second condition reported difficulty in using the keyword method. However, in the first condition, learning was better after training.

Whereas most studies on the keyword method use only concrete nouns for which concrete keywords can be easily generated, the second experiment of this study looked at the effects of the keyword method on the recall of "easy keyword" and "typical" lists of vocabulary items as compared to a control group. Subjects were 15 freshmen at Northwestern University in the keyword group and 15 undergraduates in the control group. Two types of vocabulary lists (easy keyword and typical) were provided for both groups. Subjects in the experimental group again generated their own keywords. The easy keyword list included nouns for which obvious concrete keywords could be easily found, but the typical list included unfamiliar words without regard to their concreteness and grammatical class, words for which it might be more difficult to generate

concrete keywords. All subjects were given a pretest and then trained in their groups. The control group subjects studied the two lists via study-test trial after the pretest. The keyword group was informed about the keyword method and instructions focused on identifying the keywords and forming images. There were three 50-minute sessions to study the keyword method. After the treatment, the keyword and control subjects were given the easy and the typical word list which were presented on a computer as a recall examination. The keyword group performed better than the control group in recall of the easy keyword items but the difference was not significant. The control group, on the other hand, outperformed the keyword group in recall of the typical items. It can be concluded that the keyword method might be better for easy keyword words but not typical words, perhaps because typical items do not yield obvious keywords.

The third experiment focused on the different effects of the teacher-provided and the student-generated types of the keyword method. In this experiment, keywords were generated and agreed upon by student committees which did not include the subjects. The subjects were the same 15 students who were trained in Experiment 2. They served both as the student-generated group in Experiment 2 and as the teacher-provided group in Experiment 3. The vocabulary

lists of Experiment 2 were used in this experiment, but this time the keywords were presented in parentheses to the same subjects with the words and definitions and 10 seconds was given for each item. This was followed by a 4-minute recall test and students were asked to write definitions for the target words. The scores of the teacher-provided group were compared to the scores of the same student-generated keyword group in Experiment 2. Hall found that the student-generated group performed slightly better than the teacher-provided group but nonsignificantly. The student-generated and the teacher-provided cases of the keyword method did not display any difference in Hall's study. The effects of student-generated versus teacher-provided keyword methods in typical and easy word lists are not mentioned in the results of Experiment 3.

#### Effects of Teacher-Provided and Student-Generated Keyword Methods on Immediate and Delayed Vocabulary Recall

The only study which investigated the effects of two types of the keyword method (student-generated and teacher-provided) on immediate and delayed recall of vocabulary items was conducted by Hall, Wilson and Patterson (1981). The subjects were 60 Northwestern University undergraduates with no command of the target language, Spanish. In the teacher-provided condition, the keyword method was explained and keywords were presented by the experimenter. The

student-generated group was asked to produce the keywords after an explanation of the keyword method. The control group did not get any information about the keyword method and they were free to use any technique they wanted to use to learn the target words. A list of 30 Spanish vocabulary items and their definitions were selected as target words. The general instructions were the same for the teacher-provided and student-generated groups, but additional instruction was supplied to the student-generated group on discovering keywords. The control group, on the other hand was encouraged to use any technique they wanted to use. All subjects were given 10 minutes to study the target words and student-generated group was also asked to write down the keywords they generated. At the end of the experiment and after one week, all three groups were given 4 minutes to complete the recall test. The student-generated group was significantly inferior to the teacher-provided and control groups in both immediate and delayed vocabulary recall. Teacher-provided and control groups were almost the same in immediate and delayed recall of the words. This study showed that the student-generated keyword method did not work well for the undergraduate university students in the United States. It was even inferior to the control group for both immediate and delayed recall of vocabulary items.

### Conclusion

Little research has integrated the effects of the keyword method on short-term and long-term memory with student-generated and teacher-provided types of the keyword method in EFL classrooms. Moreover, researchers have not studied the effects of the subject-generated and experimenter-provided keyword methods on recognition of vocabulary items. Thus, this study investigated the effects of the teacher-provided and student-generated keyword method on immediate and delayed recognition of vocabulary through a multiple-choice vocabulary exam as well as recall of vocabulary in a Turkish EFL classroom.

## CHAPTER 3 METHODOLOGY

### Introduction

This study was an experimental study on the immediate and delayed effects of two types of keyword vocabulary learning methods (teacher-provided and student-generated) compared to rote rehearsal on recall and recognition of vocabulary items of Turkish college students at a Turkish university. The experiment was conducted in three intact classes at Middle East Technical University. The keyword method was applied to a university classroom condition in Turkey for the first time. Acikgoz's (1992) study had been carried out in a secondary school with 6th- and 9th-grade learners. This study differed from that of Hall, Wilson and Patterson (1981) because it included the effects of the two types of the keyword method and rote learning on vocabulary recognition as well as on recall.

### Research Design

In this study, the experimenter hoped to establish a relationship between the independent variable (type of keyword) and the dependent variables (vocabulary recall and vocabulary recognition) by comparing the scores of two keyword groups and a rote rehearsal control group on the same immediate and delayed vocabulary recall and vocabulary recognition test. The same test was used as a pretest. Pretest scores were used as a co-variate for analysis of the

immediate test, and the immediate test was used as a co-variate for the analysis of the delayed test. One experimental group received treatment in the teacher-provided keyword method; the other experimental group was instructed in the student-generated keyword method; and the control group was given no information on the keyword method but asked to use the rote rehearsal technique. At the end of the treatment, the three groups were tested on recall and recognition of vocabulary items both immediately and after two weeks.

#### Subjects

Subjects were from the Middle East Technical University (METU) Department of Basic English (DBE) School of Foreign Languages which offers a preparatory program for a four-year English-medium education. Three intermediate-level intact classes were used, two as the experimental groups and one as the control group. Subjects were not randomly assigned to the three groups because it was impossible to change the present structure of the classes due to administrative constraints. However, these three classes were randomly assigned as the two experimental groups and one control group. Keeping the classroom setting as is provided more authenticity as in the study of Brown and Berry (1991).

## Instruments/Material

### Instructional Material

The 20 target vocabulary items were selected according to the following criteria as suggested by Atkinson (1975):

1. The words were nouns.
2. The word was not more than two syllables.
3. There was a concrete Turkish keyword available for the English word.
4. The English word was not similar in sound or spelling to its Turkish translation and cognates were avoided.

For the keyword practice sessions, in order to acquaint students with the keyword method, a list consisting of three words that were different from those that were used in the treatment was used (see Appendix K). This list of words was supported with keywords and example pictorial images to show the students how images could be formed using the keyword and the meaning of the word in order to make the connections and the method itself clearer. These words and pictorial images were from the Acikgoz (1992) study. Students in the student-generated group were asked to generate some keywords for the English words they chose. The rote rehearsal control group did not have a practice session.

### Testing Material

One set of testing material was developed and used three times. Words were presented in a different order each time the test was given. The testing material was a two-part test, a recall test and a recognition test. The recall part of the test (see Appendices B, C, D) was a list of the 20 target vocabulary items and students were asked to write the Turkish translation for each word. The recognition part of the test (see Appendices E, F, G) was a multiple-choice exam with a correct translation and three distractors to test the same 20 words that were studied by the students. This test was developed by the experimenter using only other words from the 20-word list as distractors. Using distractors outside the target vocabulary items could have help the students eliminate them easily because they had not studied those items in the study session. To score the test, each correct translation in the recall part and each correct answer in the recognition part was given 1 point. The highest possible score for each part of the test was 20.

The pretest aimed at measuring the existing recall and recognition of the 20 target vocabulary items in order to be able to control for pre-existing differences. The immediate test intended to measure immediate recall and recognition of the vocabulary items directly after the treatment. The delayed (long-term retention) test was administered to

measure two-week retention of the target vocabulary items both for recall and recognition. The testing material and the instructional material were pilot-tested on five students living in a Bilkent University dormitory to see if they could generate the keywords and how much time was needed. They were able to generate keywords and link them to the images in about 30 seconds.

#### Procedure

The procedure of the experiment had six stages for each of the three groups: (a) information about the experiment and solicitation of consent, (b) pretest, (c) practice, (d) study, (e) immediate test, and (f) long-term retention test.

#### Information and Consent

Before giving the pretest, students were informed about the experiment that would be conducted in their classrooms and their participation was asked for. They were asked to participate in an experiment "about vocabulary learning", and the importance of vocabulary in language learning was emphasized by the experimenter. They were given consent forms (see Appendix A) to be read, filled in and signed. Everyone in all three groups agreed to participate.

#### Pretest

A pretest of the 20 target vocabulary items were given to the three groups within the same day by the experimenter.

### Practice

Just after the pretest, on the same day for each group, the two experimental groups (student-generated and teacher-provided) were told that they were going to learn a new technique in learning vocabulary. Both experimental groups were given a practice session by the experimenter just before the study sessions, to teach them about their respective keyword methods and how to set up the necessary associations during the practice sessions. For the teacher-provided keyword group, three examples were first presented. Students were first shown some pictures of possible images for these examples taken from Acikgoz's (1992) study. Those visual images were used only as examples. They then practiced with three other words. They generated images using the provided keywords and Turkish meanings of the target words. For the student-generated group, students were taught how to find possible keywords by using orthographic and acoustic clues for the same practice words. Later, they learned how to form images including the Turkish keyword and meaning of the target word. They were shown the same examples above to teach them how to generate images. Both experimental groups were taught how to generate images including the meanings of the Turkish keyword and the target vocabulary item. The practice session for each group lasted 15 minutes. The control group was given information about

the importance of vocabulary learning and asked to learn the target vocabulary items by rote rehearsal. They did not have a practice session.

### Study

In the study session, the teacher-provided experimental group was given the list of the target words and the keywords (see appendix H), and students were asked to establish the imagery link. The student-generated experimental group received the same list of vocabulary without keywords (see Appendix I). They were asked to generate their own keywords and imagery links. The control group was asked to learn the same vocabulary list with the meanings (see Appendix J) by rote repetition in the same amount of study time with the experimental groups. The study session for each group was a 10-minute session (30 seconds per item).

### Immediate and Long-Term Retention Test

The study session was followed by an immediate vocabulary test of recall and recognition. The same test was again given to the three groups after two weeks to measure long-term retention. All the training and the tests were given in the original classrooms of the students by the experimenter to control for teacher effect.

### Data Analysis

The dependent variables of recall and recognition of each part of the test was analyzed separately. The pretest scores of each part were taken as a co-variate with the immediate posttest scores as the dependent variable in order to investigate the effects of group (the independent variable) on immediate vocabulary learning. To test long-term retention, the immediate posttest scores were taken as the co-variate, the retention test scores as the dependent variables, and group as the independent variable. The means of test scores (immediate and delayed) were compared for the three groups to see the differences, if any, between the three groups. Four one-way ANCOVAs were conducted to test the significance of differences between the groups, if any, for immediate recall and recognition and for long-term recall and recognition.

## CHAPTER 4 RESULTS

## Overview of the Study

This study aimed at investigating the effects of the teacher-provided keyword method, student-generated keyword method and rote rehearsal in three intact classes on recognition and recall of vocabulary items in classroom conditions at a Turkish university. Practice and study periods lasted a total of 25 minutes for each group. Twenty vocabulary items were chosen according to the criteria introduced by Atkinson (1975) and students were tested for both recall and recognition. The recall exam required the students to write the Turkish equivalents of English words and the recognition exam was a multiple-choice exam with three distractors. A pretest before the treatment, a posttest to test for immediate recall and recognition after the treatment, and a long-term retention test two weeks after the treatment to test for delayed recall and recognition were administered. To analyze the difference, if any, among the rote rehearsal, teacher-provided keyword, and student-generated keyword groups, four one-way analyses of co-variance (ANCOVA) using the pretest as co-variate were conducted on posttests (immediate recall and recognition) for immediate vocabulary learning, and on the retention tests using the posttest as a co-variate for long-term vocabulary retention.

### Overview of Analytical Procedures

The statistical analyses of this study were carried out in three stages. The first stage consisted of scoring of the recall and recognition tests. For each correct answer, one point was given for a range of 0-20 for both tests. One student from the rote rehearsal group, four from the student-generated keyword group, and five from the teacher-provided keyword group were dropped from the study due to absence at the long-term retention test, and data from those students were not included in the study.

At the second stage, means and standard deviations were calculated for each group for both kinds of tests as recall and recognition of vocabulary items were analyzed separately in this study. In the third stage, the effect of treatment for each group for immediate recall and recognition was analyzed using the test scores of the pretests and the immediate posttests. Then, to compare the effects of group on long-term retention, the test scores of the immediate posttests and long-term retention tests administered two weeks after the pretests were utilized.

Four one-way ANCOVAs were conducted with the pretest recall and recognition scores used as the co-variate for the immediate recall and recognition of vocabulary and the recall and recognition posttest scores used as the co-variates for long-term vocabulary recall and recognition

retention, to test whether the mean scores of the three groups were significantly different for each comparison mentioned above.  $F$  values and  $p$  values were calculated through these ANCOVAs.

#### Results of the Study

The recall and recognition parts of the vocabulary tests were analyzed separately in this study. The independent variable--group--was treated as the categorical data and means and standard deviations for each test of each group were calculated. The pretests for both recall and recognition were used as co-variates when the effects of group on immediate vocabulary recall and recognition, taking the posttest scores of both tests as the dependent variable, were tested. For the analysis of long-term retention of recall and recognition of vocabulary, posttest scores were used as co-variates whereas long-term retention test scores were taken as the dependent variable.

#### Vocabulary Recall

Table 1 presents the mean scores and standard deviations for the pretest and the posttest scores for each group.

Table 1

Means and Standard Deviations of Pretest and Immediate  
Posttest Scores of All Groups for Vocabulary Recall

Group	<u>n</u>	Pretest		Posttest	
		<u>M</u>	<u>SD</u>	<u>M</u>	<u>SD</u>
Rote Reh.	14	1.07	0.92	18.29	2.75
T-Prov.KW	17	1.35	1.11	18.76	1.09
S-Gen.KW	16	2.00	1.26	17.63	1.86
All Groups	<u>N</u> = 47	1.47	1.16	18.23	1.98

Note. Rote Reh. = Rote rehearsal group; T-Prov.KW = Teacher-provided keyword group; S-Gen.KW = Student-generated keyword group.

Mean scores indicated that treatment made a big difference for each group. Each method had very similar results. To test the differences between the groups (methods), a one-way ANCOVA was administered with the pretest as a co-variate. This analysis aimed at investigating whether or not there were statistically significant differences between the three groups at  $p < .05$ . The summary of all effects of the ANCOVA for the immediate recall of vocabulary items is presented in Table 2.

Table 2

Results of ANCOVA: For Immediate Vocabulary Recall

(N = 47)

Source of variance	Effect		Error		F	p
	df	MS	df	MS		
Between groups	2	.27	43	.48	.56	.58

\*p < .05.

According to the results of one-way ANCOVA, there was no statistically significant difference ( $F(2, 43) = 0.56, p = .58$ ) among the three groups' recall of the vocabulary items immediately after the treatment. Treatment caused a striking difference in recall of the vocabulary items for all groups. A follow-up post hoc test was not needed since differences between the groups were not significant at  $p < .05$ .

A second analysis was carried out to observe the long-term recall retention of the three different groups to see whether or not the three groups of three different methods were able to retain the learned vocabulary items after a 2-week period. Means and standard deviations of the posttest and the retention test of the three groups were first calculated to compare the differences. Table 3 presents the means and the standard deviations for the posttest and the

long-term retention test for the three groups.

Table 3

Means and Standard Deviations of Immediate Posttest and Long-Term Retention Test of All Groups for Vocabulary Recall

Group	n	Posttest		Retention test	
		<u>M</u>	<u>SD</u>	<u>M</u>	<u>SD</u>
Rote Reh.	14	18.29	2.75	9.07	4.18
T-Prov.KW	17	18.76	1.09	11.47	5.35
S-Gen.KW	16	17.63	1.86	11.31	3.74
All Groups	<u>N</u> = 47	18.23	1.98	10.62	4.54

Note. Rote Reh. = Rote rehearsal group; T-Prov.KW = Teacher-provided keyword group; S-Gen.KW = Student-generated keyword group.

According to Table 3, there is a forgetting in recall of vocabulary items for all groups. Rote rehearsal group scored slightly worse than the keyword groups in long-term retention test whereas the keyword groups were almost the same. The three groups were then analyzed through a one-way ANCOVA using the posttest as a co-variate to investigate if the three groups showed statistically significant differences. The summary of all effects of ANCOVA for the long-term retention of vocabulary recall is presented in Table 4.

Table 4

Results of ANCOVA: For Delayed Vocabulary Recall (Long-Term Retention)

(N = 47)

Source of variance	Effect		Error		F	p
	df	MS	df	MS		
Between groups	2	6.99	43	6.37	1.09	.34

\*p< .05.

The results of the one-way ANCOVA showed no statistically significant difference ( $F(2, 43) = 1.09$ ,  $p = .34$ ) between the groups for delayed recall (long-term retention). Because there was not a significant difference between the groups, a post hoc test was not carried out.

Vocabulary Recognition

Test scores for vocabulary recognition were analyzed in the same way as the vocabulary recall results were. Test scores for vocabulary recognition, like vocabulary recall, ranged between 0-20. Means and standard deviations for each group for both the posttest and the retention test were calculated. Table 5 presents the mean scores and standard deviations for the pretest and the posttest scores of vocabulary recognition for the three groups.

Table 5

Means and Standard Deviations of Pretest and Immediate  
Posttest Scores of All Groups for Vocabulary Recognition

Group	n	Pretest		Posttest	
		<u>M</u>	<u>SD</u>	<u>M</u>	<u>SD</u>
Rote Reh.	14	4.29	2.95	19.57	.94
T-Prov.KW	17	3.76	1.92	19.82	.53
S-Gen.KW	16	4.06	1.77	19.75	.58
All Groups	<u>N</u> = 47	4.04	2.19	19.71	.68

Note. Rote Reh. = Rote rehearsal group; T-Prov.KW = Teacher-provided keyword group; S-Gen.KW = Student-generated keyword group.

For immediate vocabulary recognition, the three methods were almost the same with regard to the posttest scores. Subjects in all groups scored very high with very small standard deviations. This means subjects did not appear to be different between or within the groups. Treatment caused a big change between the scores of the pretest and immediate posttest for all groups. Any possible differences among the groups were analyzed through one-way ANCOVA using the posttest as a co-variate. Table 6 presents the summary of all effects of ANCOVA for the immediate recognition of the vocabulary items after the treatment.

Table 6

Results of ANCOVA: For Immediate Vocabulary Recognition

(N = 47)

Source of variance	Effect		Error		F	p
	df	MS	df	MS		
Between groups	2	5.60	43	3.93	1.42	.25

\*p &lt; .05.

Again, the results of ANCOVA did not show a statistically significant difference ( $F(2, 43) = 1.42, p = .25$ ) between the groups. According to Table 6, there is not a statistically significant difference between the groups at  $p < .05$ . Therefore, a further post hoc test was not needed to see where the difference was.

Long-term retention of the recognition of the vocabulary items was analyzed by considering the scores of the posttest and retention test. Means and standard deviations of the posttest and the long-term retention test of the three groups were calculated. Table 7 presents the means and standard deviations for the posttest and the retention test for the three groups for vocabulary recognition.

Table 7

Means and Standard Deviations of Immediate Posttest and Retention Test of All Groups for Vocabulary Recognition

Group	n	Posttest		Retention test	
		<u>M</u>	<u>SD</u>	<u>M</u>	<u>SD</u>
Rote Reh.	14	19.57	.94	17.86	2.90
T-Prov.KW	17	19.82	.53	17.82	3.00
S-Gen.KW	16	19.75	.58	19.00	1.93
All Groups	<u>N</u> = 47	19.71	.68	18.23	2.66

Note. Rote Reh. = Rote rehearsal group; T-Prov.KW = Teacher-provided keyword group; S-Gen.KW = Student-generated keyword group.

Forgetting for the recognition of the vocabulary items was not great over time in contrast to the recall results. Most students were able to easily recognize the target vocabulary items in the multiple-choice exam two weeks after the treatment. The student-generated keyword group appeared to perform slightly better than the rote rehearsal and teacher-provided keyword groups. One-way ANCOVA was again used to test if there were significant differences between the three methods in long-term retention of recognition of vocabulary items using the posttest as a co-variate. Table 8 presents the summary of all effects of ANCOVA for the long-term retention of vocabulary recognition two weeks after the treatment.

Table 8

Results of ANCOVA: For Delayed Vocabulary Recognition (Long-Term Retention)

(N = 47)

Source of variance	Effect		Error		F	p
	df	MS	df	MS		
Between groups	2	29.60	43	18.58	1.59	.21

\*p< .05.

The results of the ANCOVA showed no statistically significant difference ( $F(2, 43) = 1.59, p = .21$ ) between the groups for the long-term retention of vocabulary recognition. Hence, this analysis was not followed by a post hoc test.

All the analyses showed no statistically significant difference between the groups after treatment at  $p < .05$ . None of the three methods proved to be different from the other in recall and recognition of vocabulary items at both immediate and delayed testing. The well-known keyword strategy was not superior to traditional rote rehearsal in the conditions of the present study. No difference was found between teacher-provided and student-generated keyword groups in immediate and long-term retention of vocabulary recall and recognition.

## CHAPTER 5 DISCUSSION OF FINDINGS AND CONCLUSION

## Summary of the Study

This study investigated the effects of two types of keyword method--teacher-provided and student-generated--on the immediate and delayed vocabulary recall and recognition compared to traditional rote rehearsal as vocabulary learning strategies. In this study, 47 Middle East Technical University preparatory school students at the intermediate level of proficiency participated as subjects in three intact classes. Each class was randomly assigned to the three methods. At the beginning of the study, a pretest including the 20 target vocabulary items was administered. Immediately after the pretest each experimental group was trained in the specific vocabulary learning method they had been assigned to. The control group was only instructed to use rote rehearsal. Each group then had 10 minutes to learn the target vocabulary. A posttest--the same as the pretest--followed the study session to measure immediate recall and recognition. Two weeks later, the same test was given as the long-term retention test. The three groups' immediate posttest scores were compared for both recall and recognition in two separate one-way ancovas using the pretest as a co-variate. Similar analysis were conducted on long-term retention scores with the posttest as a co-variate. The results of

the analyses showed no statistically significant difference at  $p < .05$  between the three groups (methods) at immediate and delayed (two weeks) recall and recognition of vocabulary items. None of the three methods seemed to be superior to the other(s). These findings may be the results of some limitations of the study which will be assessed separately.

#### Discussion of the Results

In the literature, the findings about the delayed effects of the keyword method are mixed. Acikgoz (1992) found that the keyword method was superior to rote rehearsal for the acquisition of vocabulary items. In a study carried out at a Turkish secondary school with EFL learners who were randomly assigned to the groups. Vocabulary acquisition was measured through a recall test after a 3-session treatment and the same exam was given four weeks later. Subjects in the keyword groups were able to retain more items than the rote rehearsal groups four weeks after the end of the treatment. However, a study of students of psychology at a U.S. university, Wang, Thomas and Ouellette (1992) found that forgetting was greater in keyword groups although those groups scored better than rote rehearsal group in immediate testing. The test type was a recall test administered immediately after and one week after, the treatment. This present study found no significant differences between the three groups and did not support either of the studies

mentioned above.

The effects of the keyword method on recognition of vocabulary items have been little studied. That is, most researchers have looked at the recall of vocabulary to see those effects. However, McDaniel and Pressley (1989) studied the effects of the keyword method on recognition through reading comprehension compared to no-strategy and the semantic context method with two types of texts developed by the experimenters. Seventy-five students in a psychology course at the University of Notre Dame participated in the study. Reading comprehension was measured through a true-false test. It was found that the keyword method provided more comprehension, that is, greater recognition of vocabulary items than the other groups. However, the present study measured vocabulary recognition through a multiple-choice test, and no differences between the three groups were found. Brown and Perry (1991) compared three different groups--keyword, semantic processing, and combined keyword and semantic processing--to see the effects of those different methods on recall and recognition of vocabulary items. The experimenters found that the combined keyword-semantic method provided more long-term retention for recall and recognition of vocabulary than the other methods. In the present study, the three groups were not significantly different in recognition of

vocabulary items, which again does not support any of the studies mentioned above.

One of the areas little research has been devoted to in keyword studies is the difference between teacher-provided and student-generated types of the keyword method. Hall (1988) compared those two cases in three experiments. He asked the 15 subjects to generate their own keywords in Experiments 1 and 2, and he provided them with keywords in Experiment 3. Later, results from both cases were compared to see the differences between the two types of the keyword method. Hall found that teacher-provided and student-generated types of the keyword method were not significantly different, a finding similar to this present study.

The effects of the teacher-provided and student-generated keyword method on the immediate and delayed vocabulary recall were studied by Hall, Wilson and Patterson (1981). Subjects were 60 Northwestern University undergraduates with no command of the target language, Spanish. They found that the student-generated group was inferior to the teacher-provided and control groups in immediate and delayed vocabulary recall. However, the present study showed no significant difference between the three groups.

The findings of the present study seem not to support any positive findings of the studies in the literature.

These results may be due to some shortcomings of the study which should be taken into consideration when evaluating the findings.

#### Limitations of the Study

Most of the keyword studies in the literature were conducted in laboratory conditions where it is easier to control the extraneous factors such as different environmental conditions in the classrooms that can influence the results of an experiment. Also using intact classes, which is usually necessitated in classroom experiments, precludes randomization of subjects. Although ANCOVA was used to control for initial differences, true experimental design calls for random selection of subjects.

The major problem of the study, however, may be that ceiling effect--an undesired high scoring for all groups--made the groups incomparable--in the immediate recall and recognition tests and delayed recognition test. Only the long-term recall test does not show the ceiling effect. Because many students scored 100% in those tests, possible differences in the immediate vocabulary recall and recognition and long-term retention of recognition of the three groups are not detectable. The mean score of all students in the immediate recall test is 18.23 out of 20. This shows that the difficulty level of the test was low. Also the mean scores of the immediate recognition test ( $M =$

19.71) and delayed recognition test ( $M = 18.23$ ) also indicate the difficulty level of tasks was too low. To solve such a problem, the number and difficulty level of the items should have been increased. The criteria of Atkinson (1975) were perhaps not appropriate for the present study considering the level of the students. The target vocabulary items selected according to those criteria were 2-syllable concrete nouns for which easy keywords are available. Thus, those vocabulary items were easy to remember especially at immediate testing (posttest) for all students. Perhaps, there should have been more complex and abstract words of different parts of speech to make the three methods more comparable.

In the study sessions for each group, a 30-second time period was allotted for each word. This period may have been too long for the students and students in the keyword groups may have used different mnemonic devices and rote rehearsal in addition to the keyword method.

Although the students were told to use the specific method they had learned, they may have used different strategies, especially rote rehearsal. Turkish students are generally memorization-oriented and they may automatically turn to rote memorization when confronted with new vocabulary items. Some measures should have been taken to make sure that they had used the specific method they had

been assigned to. The student-generated group could have been asked to write each keyword they found and the teacher-provided group could have listed the keywords for each item they were given, as well as answer the test questions.

#### Implications for Further Research

The issues touched on in the limitations of the study should be taken into consideration for further keyword studies. Researchers should continue to experiment with the classroom potential of the keyword method since a laboratory condition does not reflect what really happens in the classroom. Learners with different levels of language proficiency in Turkey can be the subjects of future studies, and the sample size of the subjects should be larger so that results can be more generalizable. In addition, keyword studies might address individual differences in order to discover those who can best benefit from it. Group experiments only reflect the applicability of the methods in terms of group averages.

There have been many studies on the effects of the keyword method on vocabulary learning for more than 20 years often showing positive results. Further studies should continue to investigate vocabulary learning, and other vocabulary learning strategies as well as the keyword method should also be studied to determine the best strategies for learners in order to make them more autonomous in vocabulary

learning.

#### Pedagogical Implications

This study aimed at showing teachers the usefulness of the keyword method as a vocabulary learning strategy in EFL classrooms in Turkey. It also aimed to demonstrate that learning strategies in language learning are important to create autonomous learners who can manipulate their own learning processes. Although the findings of the present study are discouraging, researchers have shown that learning strategies are crucial in language learning and many studies support the usefulness of learning strategies (e.g., O'Malley, 1987; Oxford, Lavine, & Crookall, 1989; Wenden, 1987).

The researcher observed that students at Middle East Technical University really enjoyed learning a vocabulary learning strategy--the keyword method. Students stated that they liked the experiment very much, and they were able to find different keywords for the target vocabulary items. They did not want the experimenter to leave, and they were eager to learn more about vocabulary learning strategies. Some teachers of other classes not involved in the study stated that their students also wanted somebody to go into their classrooms to teach them about vocabulary learning strategies. This situation shows that language learners in Turkey are open to learning strategies and they are curious

about new ones. Thus, teachers and curriculum designers should encourage strategy training and incorporate different strategies into their syllabi and curricula.

As mentioned above, in a group experiment, individual differences are not detected. Some individuals may benefit from different strategies. Instead of focusing on a specific strategy, students should be provided with a repertoire of learning strategies so that they can make best use of the strategies by adopting them selectively according to their individual characteristics.

## References

- Acikgoz, K. (1992). Effects of the mnemonic keyword method on teaching English vocabulary to Turkish students. The proceedings of the First International Conference on ELT Held at Hacettepe University, 1, 71-82.
- Atkinson, R. C. (1975). Mnemotechnics in second-language learning. American Psychologist, 30, 821-828.
- Brown, D. H. (1994). Principles of language learning and teaching. Englewood Cliffs: Prentice Hall.
- Brown, T. S., & Perry, F. L., Jr. (1991). A comparison of three learning strategies for ESL vocabulary acquisition. TESOL Quarterly, 4, 655-670.
- Hall, J. W. (1988). On the utility of the keyword mnemonic for vocabulary learning. Journal of Educational Psychology, 80, 554-562.
- Hall, J. W., Wilson, K. P., & Patterson, R. J. (1981). Mnemotechnics: Some limitations of the mnemonic keyword method for the study of foreign language vocabulary. Journal of Educational Psychology, 73, 345-357.
- Levin, J. R., Pressley, M., McCormick, C. B., Miller, G. E., & Shriberg, L. K. (1979). Assessing the classroom potential of the keyword method. Journal of Educational Psychology, 71, 583-594.

- McDaniel, M. A., & Pressley, M. (1984). Putting the keyword method in context. Journal of Educational Psychology, 76, 598-609.
- McDaniel, M. A., & Pressley, M. (1989). Keyword and context instruction of new vocabulary meanings: Effects on text comprehension and memory. Journal of Educational Psychology, 81, 204-213.
- O'Malley, J. M. (1987). The effects of training in the use of learning strategies on learning English as a second language. In A. L. Wenden & J. Rubin (Eds.), Learner strategies in language learning (pp. 133-144). Englewood Cliffs: Prentice Hall.
- Oxford, R. L., Lavine, R. Z., & Crookall, D. (1989). Language learning strategies, the communicative approach, and their classroom implications. Foreign Language Annals, 1, 29-39.
- Pressley, M. (1977). Children's use of the keyword to learn simple Spanish vocabulary words. Journal of Educational Psychology, 69, 465-472.
- Pressley, M., Levin, J. R., & Delaney, H. D. (1982). The mnemonic keyword method. Review of Educational Research, 52, 61-91.

- Pressley, M., Levin, J. R., Hall, J. W., Miller, G. E., & Berry, J. K. (1980). The keyword method and foreign word acquisition. Journal of Experimental Psychology: Human Learning and Memory, 6, 163-173.
- Raugh, M. R., & Atkinson R. C. (1975). A mnemonic method for learning a second language vocabulary. Journal of Educational Psychology, 67, 1-16.
- Raugh, M. R., Schupbach, R. D., & Atkinson R. C. (1977). Teaching a large Russian language vocabulary by the mnemonic keyword method. Instructional Science, 6, 199-221.
- Rubin, J. (1987). Learner strategies: Theoretical assumptions, research history and typology. In A. L. Wenden & J. Rubin (Eds.), Learner strategies in language learning (pp. 15-30). Englewood Cliffs: Prentice Hall.
- Skehan, P. (1989). Individual differences in second-language learning. London: Edward Arnold.
- Stern, H. H. (1983). Fundamental concepts of language teaching. London: Oxford University.
- Thompson, I. (1987). Memory in language learning. In A. L. Wenden & J. Rubin (Eds.), Learner strategies in language learning (pp. 43-56). Englewood Cliffs: Prentice Hall.

- Wang, A. Y, Thomas, M. H., & Ouellette, J. A. (1992).  
Keyword mnemonic and retention of second language  
vocabulary words. Journal of Educational Psychology,  
84, 520-528.
- Weinstein, C. E., & Mayer, R. E. (1986). The teaching of  
learning strategies. In M. C. Wittrock (Ed.), Handbook  
of research on teaching (pp. 315-327). New York:  
Macmillan.
- Wenden A. L. (1987). Conceptual background and utility.  
In A. L. Wenden & J. Rubin (Eds.), Learner strategies in  
language learning (pp. 3-13). Englewood Cliffs: Prentice  
Hall.

Appendix A  
Consent Form

I am conducting an experiment on a vocabulary learning strategy and planning to make it useful for all colleagues. Your scores on the tests will not affect your grades and will not be given to anyone. Your names will not be announced anywhere. Participation is voluntary and you can resign anytime you want.

I appreciate your help.

DERYA YAYLI

Bilkent University

MA TEFL

I would like to participate in this experiment and I know that I can withdraw anytime I want.

Name : \_\_\_\_\_

Date : \_\_\_\_\_

Signature : \_\_\_\_\_

If you have any questions about the study, you may contact the study advisor:

Dr. Phyllis L. Lim, Director

MA TEFL Program

Bilkent University

## Appendix B

Vocabulary Recall Test #1

Name:

May 2, 1995

Write the Turkish equivalents (one word or more) of the following words.

(Aşağıdaki kelimelerin Türkçe karşılıklarını bir veya daha fazla kelimeyle yazınız.)

PARADE	_____	JASPER	_____
DUNGEON	_____	MARTIN	_____
CULPRIT	_____	SHANTY	_____
CASKET	_____	THIMBLE	_____
ALLEY	_____	PATROL	_____
COLLAR	_____	ANCHOR	_____
CHIMNEY	_____	SHRUB	_____
MIDDEN	_____	CRADLE	_____
BANQUET	_____	MIRAGE	_____
SAUCER	_____	MANGER	_____

## Appendix C

Vocabulary Recall Test #2

Name:

May 2, 1995

Write the Turkish equivalents (one word or more) of the following words.

(Aşağıdaki kelimelerin Türkçe karşılıklarını bir veya daha fazla kelimeyle yazınız.)

MIDDEN	_____	CRADLE	_____
DUNGEON	_____	CASKET	_____
CULPRIT	_____	MANGER	_____
SAUCER	_____	SHANTY	_____
MARTIN	_____	BANQUET	_____
SHRUB	_____	JASPER	_____
ALLEY	_____	COLLAR	_____
PATROL	_____	ANCHOR	_____
CHIMNEY	_____	PARADE	_____
THIMBLE	_____	MIRAGE	_____

## Appendix D

Vocabulary Recall Test #3

Name:

May 16, 1995

Write the Turkish equivalents (one word or more) of the following words.

(Aşağıdaki kelimelerin Türkçe karşılıklarını bir veya daha fazla kelimeyle yazınız.)

PATROL \_\_\_\_\_

MARTIN \_\_\_\_\_

SAUCER \_\_\_\_\_

CULPRIT \_\_\_\_\_

THIMBLE \_\_\_\_\_

PARADE \_\_\_\_\_

BANQUET \_\_\_\_\_

CHIMNEY \_\_\_\_\_

JASPER \_\_\_\_\_

ANCHOR \_\_\_\_\_

COLLAR \_\_\_\_\_

MANGER \_\_\_\_\_

SHRUB \_\_\_\_\_

MIDDEN \_\_\_\_\_

DUNGEON \_\_\_\_\_

SHANTY \_\_\_\_\_

CASKET \_\_\_\_\_

ALLEY \_\_\_\_\_

CRADLE \_\_\_\_\_

MIRAGE \_\_\_\_\_

## Appendix E

Vocabulary Recognition Test #1

Name :

May 2, 1995

Circle the Turkish equivalents of the given English words.

(İngilizce kelimelerin Türkçe karşılıklarını seçiniz.)

Cevapları daire içine alınız.)

- |            |  |            |   |
|------------|--|------------|---|
| 1. DUNGEON | a) çalı<br>b) yemlik<br>c) zindan<br>d) akbaba               | 2. ALLEY   | a) baca<br>b) suçlu<br>c) yeşim<br>d) sokak               |
| 3. SAUCER  | a) çapa<br>b) gecekondu<br>c) çöp yığını<br>d) fincan tabağı | 4. BANQUET | a) yaka<br>b) tabut<br>c) ziyafet<br>d) dikiş yüksüğü     |
| 5. PATROL  | a) beşik<br>b) devriye<br>c) kırlangıç<br>d) geçit töreni    | 6. COLLAR  | a) yaka<br>b) ziyafet<br>c) kırlangıç<br>d) fincan tabağı |
| 7. MIRAGE  | a) çapa<br>b) serap<br>c) zindan<br>d) çöp yığını            | 8. THIMBLE | a) çalı<br>b) yemlik<br>c) gecekondu<br>d) dikiş yüksüğü  |
| 9. SHRUB   | a) çalı<br>b) yeşim<br>c) akbaba<br>d) devriye               | 10. MARTIN | a) baca<br>b) beşik<br>c) kırlangıç<br>d) geçit töreni    |
| 11. SHANTY | a) sokak<br>b) ziyafet<br>c) gecekondu<br>d) çöp yığını      | 12. CASKET | a) suçlu<br>b) tabut<br>c) devriye<br>d) fincan tabağı    |

13. JASPER a) yeşim  
b) beşik  
c) zindan  
d) yemlik
14. CULPRIT a) baca  
b) serap  
c) suçlu  
d) dikiş yüksüğü
15. CHIMNEY a) baca  
b) beşik  
c) zindan  
d) kırlangıç
16. ANCHOR a) çapa  
b) suçlu  
c) ziyafet  
d) gecekondur
17. MANGER a) sokak  
b) yemlik  
c) çöp yığını  
d) fincan tabağı
18. CRADLE a) tabut  
b) beşik  
c) sokak  
d) devriye
19. PARADE a) serap  
b) yeşim  
c) geçit töreni  
d) dikiş yüksüğü
20. MIDDEN a) baca  
b) yaka  
c) kırlangıç  
d) çöp yığını

## Appendix F

Vocabulary Recognition Test #2

Name:

May 2, 1995

Circle the Turkish equivalents of the given English words.

(İngilizce kelimelerin Türkçe karşılıklarını seçiniz.)

Cevapları daire içine alınız.)

- |            |  |             |   |
|------------|--|-------------|---|
| 1. SHRUB   | a) çalı<br>b) yeşim<br>c) akbaba<br>d) devriye               | 2. MARTIN   | a) baca<br>b) beşik<br>c) kırlangıç<br>d) geçit töreni    |
| 3. SAUCER  | a) çapa<br>b) gecekondü<br>c) çöp yığını<br>d) fincan tabağı | 4. BANQUET  | a) yaka<br>b) tabut<br>c) ziyafet<br>d) dikiş yüksüğü     |
| 5. MANGER  | a) sokak<br>b) yemlik<br>c) çöp yığını<br>d) fincan tabağı   | 6. CRADLE   | a) tabut<br>b) beşik<br>c) sokak<br>d) devriye            |
| 7. PATROL  | a) beşik<br>b) devriye<br>c) kırlangıç<br>d) geçit töreni    | 8. COLLAR   | a) yaka<br>b) ziyafet<br>c) kırlangıç<br>d) fincan tabağı |
| 9. DUNGEON | a) çalı<br>b) yemlik<br>c) zindan<br>d) akbaba               | 10. ALLEY   | a) baca<br>b) suçlu<br>c) yeşim<br>d) sokak               |
| 11. MIRAGE | a) çapa<br>b) serap<br>c) zindan<br>d) çöp yığını            | 12. THIMBLE | a) çalı<br>b) yemlik<br>c) gecekondü<br>d) dikiş yüksüğü  |

13. JASPER a) yeşim  
b) beşik  
c) zindan  
d) yemlik
14. CULPRIT a) baca  
b) serap  
c) suçlu  
d) dikiş yüksüğü
15. PARADE a) serap  
b) yeşim  
c) geçit töreni  
d) dikiş yüksüğü
16. MIDDEN a) baca  
b) yaka  
c) kırlangıç  
d) çöp yığını
17. SHANTY a) sokak  
b) ziyafet  
c) gecekondur  
d) çöp yığını
18. CASKET a) suçlu  
b) tabut  
c) devriye  
d) fincan tabağı
19. CHIMNEY a) baca  
b) beşik  
c) zindan  
d) kırlangıç
20. ANCHOR a) çapa  
b) suçlu  
c) ziyafet  
d) gecekondur

## Appendix G

Vocabulary Recognition Test #3

Name :

May 16, 1995

Circle the Turkish equivalents of the given English words.

(İngilizce kelimelerin Türkçe karşılıklarını seçiniz.)

Cevapları daire içine alınız.)

- |             |   |            |   |
|-------------|---|------------|---|
| 1. CHIMNEY  | a) baca<br>b) beşik<br>c) zindan<br>d) kırlangıç        | 2. MIRAGE  | a) çapa<br>b) serap<br>c) zindan<br>d) çöp yığını           |
| 3. SHANTY   | a) sokak<br>b) ziyafet<br>c) gecekondü<br>d) çöp yığını | 4. CASKET  | a) suçlu<br>b) tabut<br>c) devriye<br>d) fincan tabağı      |
| 5. MIDDEN   | a) baca<br>b) yaka<br>c) kırlangıç<br>d) çöp yığını     | 6. THIMBLE | a) çalı<br>b) yemlik<br>c) gecekondü<br>d) dikiş yüksüğü    |
| 7. SHRUB    | a) çalı<br>b) yeşim<br>c) suçlu<br>d) devriye           | 8. JASPER  | a) yeşim<br>b) beşik<br>c) zindan<br>d) yemlik              |
| 9. CRADLE   | a) tabut<br>b) beşik<br>c) sokak<br>d) devriye          | 10. PARADE | a) serap<br>b) yeşim<br>c) geçit töreni<br>d) dikiş yüksüğü |
| 11. DUNGEON | a) çalı<br>b) yemlik<br>c) zindan<br>d) geçit töreni    | 12. MANGER | a) sokak<br>b) yemlik<br>c) çöp yığını<br>d) fincan tabağı  |

13. ANCHOR a) çapa  
b) suçlu  
c) ziyafet  
d) gecekondu
14. ALLEY a) baca  
b) suçlu  
c) yeşim  
d) sokak
15. PATROL a) beşik  
b) devriye  
c) kırlangıç  
d) geçit töreni
16. SAUCER a) çapa  
b) gecekondu  
c) çöp yığını  
d) fincan tabağı
17. CULPRIT a) baca  
b) serap  
c) suçlu  
d) dikiş yüksüğü
18. MARTIN a) baca  
b) beşik  
c) kırlangıç  
d) geçit töreni
19. COLLAR a) yaka  
b) ziyafet  
c) kırlangıç  
d) fincan tabağı
20. BANQUET a) yaka  
b) tabut  
c) ziyafet  
d) dikiş yüksüğü

## Appendix H

Study List for the Teacher-Provided Group

<u>WORD</u>	<u>KEYWORD</u>	<u>MEANING</u>
MIRAGE	garaj	serap
CRADLE	kredi	beşik
ALLEY	ali	sokak
CASKET	kasket	tabut
SHANTY	şantiye	gecekondu
DUNGEON	sancın	zindan
MIDDEN	miden	çöp yığını
SHRUB	şarap	çalı
MANGER	mangır	yemlik
COLLAR	kollar	yaka
ANCHOR	ankara	çapa
JASPER	jest	yeşim
CHIMNEY	çim	baca
BANQUET	banket	ziyafet
PARADE	para	geçit töreni
THIMBLE	tembel	dikiş yüksüğü
CULPRIT	kalp	suçlu
SAUCER	sos	fincan tabağı
MARTIN	mart	kırlangıç
PATROL	petrol	devriye

## Appendix I

Study List for the Student-Generated Group

You may write the keywords you found in the blanks provided.

(Boşluklara bulduğunuz anahtar-kelimeleri yazabilirsiniz.)

<u>WORD</u>	<u>KEYWORD</u>	<u>MEANING</u>
MIRAGE	_____	serap
CRADLE	_____	beşik
ALLEY	_____	sokak
CASKET	_____	tabut
SHANTY	_____	gecekondu
DUNGEON	_____	zindan
MIDDEN	_____	çöp yığını
SHRUB	_____	çalı
MANGER	_____	yemlik
COLLAR	_____	yaka
ANCHOR	_____	çapa
JASPER	_____	yeşim
CHIMNEY	_____	baca
BANQUET	_____	ziyafet
PARADE	_____	geçit töreni
THIMBLE	_____	dikiş yüksüğü
CULPRIT	_____	suçlu
SAUCER	_____	fincan tabağı
MARTIN	_____	kırlangıç
PATROL	_____	devriye

## Appendix J

Study List for the Rote Rehearsal Group

<u>WORD</u>	<u>MEANING</u>
MIRAGE	serap
CRADLE	beşik
ALLEY	sokak
CASKET	tabut
SHANTY	gecekondu
DUNGEON	zindan
MIDDEN	çöp yığını
SHRUB	çalı
MANGER	yemlik
COLLAR	yaka
ANCHOR	çapa
JASPER	yeşim
CHIMNEY	baca
BANQUET	ziyafet
PARADE	geçit töreni
THIMBLE	dikiş yüksüğü
CULPRIT	suçlu
SAUCER	fincan tabağı
MARTIN	kırlangıç
PATROL	devriye

Appendix K

Sample Words for the Practice Session

TIE

PILLOW

FISH