THE EFFECT OF THE KEYWORD METHOD ON ESP VOCABULARY LEARNING AND RETENTION

NERGİS KOPARAN

JANUARY 2018

THE EFFECT OF THE KEYWORD METHOD ON ESP VOCABULARY LEARNING AND RETENTION

A THESIS SUBMITTED TO THE
GRADUATE SCHOOL OF EDUCATIONAL SCIENCES
OF
BAHÇEŞEHİR UNIVERSITY

BY

Nergis KOPARAN

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF ARTS
IN THE DEPARTMENT OF ENGLISH LANGUAGE EDUCATION

JANUARY 2018

Approval of the Graduate School of Educational Sciences

Assist. Prof. Dr. Enisa MEDE

Director

I certify that this thesis meets all the requirements as a thesis for the degree of Master of Arts.

Assist. Prof. Dr. Hatime ÇİFTÇİ

Coordinator

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

Prof. Dr. Derin ATAY

Supervisor

Examining Committee Members

Prof. Dr. Derin ATAY	(BAU, ELT)
Assist. Prof. Dr. Enisa MEDE	(BAU, ELT)
Dr. Pınar ERSİN	(MU, ELT)

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last Name: Nergis KOPARAN

Signature: 

ABSTRACT

THE EFFECT OF THE KEYWORD METHOD ON ESP VOCABULARY LEARNING AND RETENTION

Koparan, Nergis

Master's Thesis, Master's Program in English Language Education

Supervisor: Prof. Dr. Derin ATAY

January 2018, 80 pages

The purpose of this study is to investigate the effect of the keyword method on ESP vocabulary learning and retention. The study also attempts to explore the perceptions of students toward the application of the keyword method in Aviation English learning. To achieve this purpose, a quasi-experimental research design was adopted. Thirty-two Turkish EFL students from two intact classes in Civil Aviation Cabin Services Program at a private university in Istanbul were assigned as the experimental and control groups. The experimental group was instructed through the keyword method while the control group was taught target vocabulary items by means of definitions and synonyms. The data was collected through vocabulary knowledge scales and structured interviews. The findings revealed that the keyword method had a positive impact on students' ESP vocabulary learning and retention along with their perceptions of the keyword strategy. The students in the experimental group recalled more target vocabulary than the control group did. On a basis of these findings, the study provides practical implications to integrate the keyword method into Aviation English classes.

Keywords: Vocabulary Learning, the Keyword Method, ESP, Vocabulary Retention

ÖZ

ANAHTAR SÖZCÜK YÖNTEMİNİN ÖZEL AMAÇLI İNGİLİZCEDE KELİME ÖĞRENİMİ VE AKILDA TUTMA ÜZERİNE ETKİSİNİN İNCELENMESİ

Koparan, Nergis

Yüksek Lisans Tezi, İngiliz Dili Eğitimi Yüksek Lisans Programı

Tez Yöneticisi: Prof. Dr. Derin ATAY

Ocak 2018, 80 sayfa

Bu çalışmanın amacı anahtar sözcük yönetiminin özel amaçlı İngilizcede kelime öğrenimi ve akılda tutma üzerine etkisini araştırmaktadır. Çalışma, ayrıca öğrencilerin anahtar sözcük yönteminin Havacılık İngilizcesi kelime öğreniminde kullanılmasına yönelik algılarını incelemeyi hedeflemektedir. Bu amaçlar doğrultusunda, yarı deneysel bir araştırma dizaynı benimsenmiştir. İstanbul'da özel bir üniversitenin Sivil Havacılık Kabin Hizmetleri Programının iki sınıfından toplam otuz iki öğrenci deney grubu ve kontrol grubu olarak atanmıştır. Kontrol grubuna hedef kelimeler tanımları ve eş anlamları aracılığıyla öğretilirken, deney grubuna anahtar sözcük yöntemiyle öğretilmiştir. Veriler; sözcük bilgisi ölçekleri ve yarı yapılandırılmış görüşmelerden toplanmıştır. Bulgular; anahtar sözcük yönteminin öğrencilerin anahtar sözcük stratejisine yönelik algılarının yanı sıra, özel amaçlı İngilizce kelime öğrenimi ve akılda tutabilmeleri üzerine olumlu bir etkisi olduğunu göstermiştir. Deney grubundaki öğrencilerin kontrol grubundaki öğrencilere göre daha fazla sayıda hedef kelime hatırladığı ortaya çıkmıştır. Bu araştırma, anahtar sözcük yönteminin Havacılık İngilizcesi derslerinde kullanımı açısından pratik uygulamalar sağlamaktadır.

Anahtar Kelimeler: Kelime Öğrenimi, Anahtar Sözcük Yöntemi, Özel Amaçlı İngilizce, Kelime Akılda Tutma

To my beloved mother

ACKNOWLEDGEMENTS

First of all, I would like to express my sincere gratitude to my advisor Prof. Dr. Derin ATAY for her constructive attitude, invaluable support, prompt and regular feedback every time I needed and all her contributions in each phase of this study. This thesis could not have been completed without her strong motivation and guidance.

Besides my advisor, I owe special thanks to my examining committee members Assist. Prof. Enisa MEDE and Dr. Pınar ERSİN for their valuable and insightful comments which helped me look at my research from different perspectives.

I would like to express my heartfelt gratitude to all the students who participated in this study. Their voluntary participation and invaluable help contributed a lot during this process.

Last but not least, my deepest gratitude goes to my beloved mother and brother, Fatma Aytaçoğlu and Orhun Koparan, who love me unconditionally, for their endless support and patience, positive attitude, priceless understanding in each step of my thesis.

TABLE OF CONTENTS

LIST OF TABLES

TABLES

Chapter 1

Introduction

1.1 Theoretical Framework

The growing use of intentional vocabulary learning strategies in EFL classrooms has made it possible for English language teachers and instructors to employ numerous vocabulary teaching and learning methods and strategies. This development has also enabled language teachers to incorporate intentional learning strategies into their teaching process for vocabulary knowledge acquisition so that their students get more motivated and willing to gain a greater deal of vocabulary knowledge. It becomes more common that language learners are likely to use explicit vocabulary learning strategies while studying on their own within and out of the language classes. It is simple not only for students but also for teachers to benefit from intentional strategies with the intention of acquiring and increasing vocabulary knowledge when necessary.

Researchers have been examining the impacts and advantages of the keyword method, a mnemonic method among one of the intentional vocabulary learning strategies, on vocabulary acquisition in English as second language. A wide number of vocabulary learning strategies have been confirmed by rich research studies in the related field and one of the well-documented strategies is the keyword method (Davoudi & Yousefi, 2016). The keyword method can be considered to be an effective method in promoting L2 vocabulary knowledge since it offers EFL teachers, instructors and learners a more effective, more permanent in long-term and entertaining learning environment and process (Koksal & Cekic, 2014). The research studies which were conducted in this field showed that the keyword method is a beneficial and an effective instructional way to teach EFL learners the target vocabulary. The changes which intentional vocabulary learning strategies has led to in second language vocabulary acquisition has made it possible for EFL learners to study and practice vocabulary items on their own too whenever they need.

The common use of these explicit vocabulary learning strategies like the keyword method has also enabled language educators, teachers and instructors to change from incidental vocabulary learning methods to intentional vocabulary learning that is believed to be more effective and catchier by researchers.

1.2 Statement of the Problem

One of the main problems that learners of English for Specific Purposes come across is to be able to gain and improve vocabulary knowledge and store it in short and long-term memory. Traditional vocabulary learning methods remain insufficient to enable learners of English language to learn target vocabulary word and phrases. In addition, these traditional methods can help learners neither retain the target vocabulary items receptively nor use them productively in a sentence within an appropriate context. On the other hand, incidental vocabulary learning which has been used by teachers and instructors so far has failed to produce adequate positive results in EFL learners' immediate and delayed retention of the target vocabulary knowledge. Intentional vocabulary learning methods and strategies not only enable learners to acquire new words and phrases in the ESP field in an explicit and deliberate way but also to be able to recall them more readily at receptive and productive level when they need. The keyword method is one of these intentional vocabulary learning methods that allows students to study new lexical items in-class and out of class on their own and that also provides better retention for them even in long-term.

Another problem is that most of the incidental and intentional vocabulary learning methods have limited properties for vocabulary learning. In other words, these vocabulary learning methods mostly remain unsuccessful since they are lacking in sufficient reminders to offer learners. However, thanks to the opportunities provided by the mnemonic keyword method for EFL learners to study and gain new words and phrases, students can learn the target vocabulary items and have a chance to practice them in a short period of time. With this method, students are asked to associate the target word or phrase to be taught with a key word which could come to mind easily in their native language. Also, they learn each lexical item within a story which has the meaning of the target word or phrase at the end and which is accompanied by meaningful visual images describing the content of the story. Thus, the teacher can create a classroom atmosphere where students can improve their vocabulary knowledge in a short period of time with the aid of the reminders the keyword method presents.

1.3 Purpose of the Study

So many studies carried out so far have investigated the effect of a variety of vocabulary learning strategies on the learning and recalling of vocabulary items for English for specific purposes. However, the effectiveness of using a mnemonic method which brings

verbal and imaginal processes together and the impact of such a mnemonic method on learners' receptive and productive performance in ESP vocabulary learning have not been popular issues in the research field of language teaching and learning till recently. Hence, there is not much evidence that has explored the advantages and disadvantages or positive and negative effects of the keyword method on learners' achievement in vocabulary acquisition. Likewise, there is no enough evidence revealing the positive and negative effects of the keyword method on students' perceptions of using this vocabulary teaching and learning method. The present study; therefore, attempts to provide some insights into ESP vocabulary learning through the keyword method instruction, and the effectiveness of applying the keyword method on retrieving the previously taught vocabulary items in short-term and long-term memory. To make it clear, this study, which was conducted at a private university in İstanbul, Turkey, aims to find out whether integrating such a mnemonic method having a meaningful visual image has an impact on intermediate level Turkish Civil Aviation Cabin Services students' performance in ESP vocabulary learning and recalling.

1.4 Research Questions

To investigate the effect of the keyword method on vocabulary learning in English for specific purposes, the following research questions are addressed in this study:

1. Will there be a significant difference between the students who are instructed in the keyword method and those who are instructed through the traditional method in the retrieval of ESP vocabulary knowledge?

   1.1 Will there be a significant difference in the retrieval of ESP vocabulary knowledge at receptive and productive levels?
2. What are the EFL learners' perceptions toward learning ESP vocabulary through the keyword method?
3. Is there a difference between the experimental and control groups in the use of target vocabulary in oral communication?

1.5 Significance of the Study

It is commonly accepted that intentional vocabulary learning strategies are useful and effective ways of presenting vocabulary items explicitly. Up to now, plenty of research has been done with the intention to analyzing the use of the mnemonic keyword method with vocabulary component of language learning. And all this research has supported the concept

that the implementation of the keyword method can enable learners to enhance their vocabulary knowledge and acquisition. Using the keyword method can help language teachers to provide a richer and more permanent learning environment for students in terms of vocabulary acquisition. The findings obtained from many of the studies in this field prove that vocabulary learning via the keyword instruction is more effective than learning through the traditional vocabulary learning methods and strategies, even though there are rarely some exceptional results.

There are various studies conducted in so many countries which reveal the positive effects of vocabulary learning through the keyword method and of employing the keyword method on learners' vocabulary acquisition. However, in the literature, there is a lack of adequate studies on the impact of the keyword method in ESP vocabulary learning. Furthermore, this lack can be seen more obviously in Turkey, in particular, at Civil Aviation Cabin Services Programs of universities in Turkey. In parallel with this, it is almost not possible to gain findings and make inferences with respect to its practical efficiency and usability within aviation context. Therefore, the present study aims to investigate short-term and long-term effectiveness of using the mnemonic keyword method on the students' acquisition of the target ESP vocabulary items in two intact Civil Aviation classes at a private university in Turkey. It will provide insights into the related subject for further research too.

1.6 Definitions

ESP: Refers to English for specific purposes

EFL: Refers to English as a Foreign Language (Mayo & Lecumberri, 2003).

ELT: Refers to English Language Teaching (Harmer, 2007).

SLA: Refers to Second Language Acquisition

VKS: Refers to Vocabulary Knowledge Scale

VLT: Refers to Vocabulary Level Test

KWM: Refers to the Keyword Method

EG: Refers to Experimental Group

CG: Refers to Control Group

Chapter 2

Literature Review

2.1 Introduction

This chapter includes the significance of vocabulary and vocabulary size in second language learning. Incidental vocabulary learning and intentional vocabulary learning along with related studies are also discussed. Various vocabulary memorization strategies are also described; and a variety of studies on vocabulary learning through these strategies are reviewed in the chapter.

2.2 The Importance of Vocabulary in L2

Schmitt (2000) highlights that lexical knowledge is must for reaching communicative competence and acquiring a second language. Cahyono and Widiati (2008) argue that vocabulary teaching intends to make it possible for learners to understand the concepts of unfamiliar words, to obtain a higher number of words and to use words successfully when they have communicative purposes.

Nation (2001) adds that vocabulary knowledge and language use have a complementary relationship; that is, the more vocabulary items you have, the more you can make use of language and the more language is used, the more increase occurs in vocabulary knowledge. Besides, Rivers and Nunan (1991) give an explanation of the essentialness of adequate vocabulary: We would not be able to use the structural and functional language that we might have been taught for comprehensible communication if we did not have extensive vocabulary knowledge.

Many researchers agree that vocabulary is one of the most important elements which need to be reflected to the curriculum. Krashen (1989) stated that when students travel, they don’t carry their grammar books, they carry dictionaries and revealed second language acquirers most of the time reported that a major problem was their lack of vocabulary. Wilkins (1972) explained that there does not occur much value so long as anyone does not have a good knowledge of vocabulary. This is essential to convey the meaning which individual wishes to express. Maybe we can convey very little without grammar but we can convey nothing without vocabulary. As a conclusion, the significance of vocabulary learning is mostly recognized by language teachers and applied linguists now and also more ways of promoting vocabulary knowledge are being explored day by day.

Nation and Meara (2010) state that English vocabulary is of complexness in that it comprises of three main aspects with regard to form, meaning, and use, along with layers of meaning which are directly linked to the roots of individual words.

2.3 Vocabulary Size in L2

Schmitt (2000) argued that communicative competence and the acquisition of a second language fundamentally call for lexical knowledge. In this sense, through vocabulary teaching, we intend to enable learners to catch the concepts of unfamiliar words, gain a larger number of words and the most importantly make use of words both accurately and fluently for communicative purposes. Underlining the size of lexical items, Schmitt and McCarthy (1997) suggested that language performance is enhanced by vocabulary knowledge and the increase of vocabulary knowledge is enhanced by the use of language, which allows us to show regard to estimates of vocabulary size and their importance for second language learners.

It is essential to acquire vocabulary knowledge to be fluent in second language (L2) use. To communicate successfully in L2, learners need a great deal of vocabulary. According to the studies, to comprehend a wide range of written and spoken discourse, learners must respectively know around 6000 to 7000 and 8000 to 9000 word families (Nation, 2006). Furthermore, Nation (2006) states that we can decide how many words a learner of English as a second or foreign language needs to know in a several ways. One way is to determine how many words there are in English and to consider that as a learning goal. Although studies having done this reveal figures of 114,000 word-families, one objection is that native speakers cannot know all the words in their native language and these figures are too many for second language learners. A second way of deciding vocabulary learning goals is to look at how many words a native speaker knows. History of research in this area indicates that well-educated native speakers are of around 20,000 word-families. A third way is to find how much vocabulary you need to learn so as to use English in a conversation or while reading a novel and watching a movie.

Norbert and Diane (2014) argued that high-frequency English vocabulary should comprise the most frequent 3,000 word families and low-frequency English vocabulary should be restricted to the 9,000 level on the ground that 8-9,000 word families are adequate to be able to read a vast number of authentic texts (Nation, 2006). It turned out in one of Nation's recent studies (2006) in which he adopted a frequency-based approach instead of the

four-part categorization that there is a need for readers to have the knowledge of 8-9,000 word families to be able to understand a broad range of English texts.

From the point of Norbert and Diane (2014), we need to carefully examine exactly what the extent and boundaries of high frequency and low frequency vocabulary are. Traditionally, high frequency vocabulary has been believed to account for around the first 2,000 most frequent word families; in other words, the most frequent 2,000 word families build the traditionally-accepted numerical value for high frequency vocabulary. Schonell, Meddleton and Shaw (1956) concluded that it was adequate to gain 2,000 word families to be involved in daily conversation and provided with evidence that nearly 99% of the discourse in the speech of Australian workers is covered by 2,000 word families. As it has been testified above, the figure of 2,000 originates from frequency counts and research done over 50 years old.

Further evidence was presented with a small frequency study. Cobb (2007) explored if at different frequency levels occurred enough to be acquired solely from being exposed incidentally. He determined 30 target words - 10 selected words from each of the 1,000, 2,000 and 3,000 levels – to look at how often they appeared in an extract with 517,000 words taken from the Brown written English corpus. He concluded that at least eight out of ten target words belonging to the first 1,000 and seven words from the second 1,000 frequency level appeared six or more times. As for the third 1,000 frequency level, this figure decreased to three-five words. This result showed that the 3,000 figure is the lowest frequency which one can call as 'high-frequency' in parallel with learning opportunities from reading.

Conversely, low-frequency vocabulary, the other extreme of the frequency continuity, has been represented in several ways, varying from anything beyond 2,000 word families to the whole word families beyond the 10,000 frequency level. Low-frequency vocabulary is included at a level where vocabulary is very infrequent and thus has a very restricted usefulness (Norbert & Diane, 2014). One corpus of contemporary written Dutch was analyzed so as to decide on how much vocabulary achieving university study called for (Hazenberg & Hulstijn, 1996). It was concluded that a minimum of 10,000 base words were needed in order to reach a sufficient coverage of these corpora.

Underscoring strong correlations between receptive vocabulary size tests and reading comprehension tests were found in accordance with empirical studies, Stenius (2008) states vocabulary size is helpful to predict as much variance as 72% in reading, which means, to

many researchers, receptive vocabulary size is the determinant factor to be successful in L2 reading. In this particular, how large a receptive vocabulary is needed to be able to comprehend an authentic text is a question for teachers and learners. Within the scope of determining how many words have to be known so as to gain comprehension of a text adequately, Laufer (1989) pointed out that it would be sufficient to gain a coverage of 95% for comprehending a reading adequately. However, it has been suggested by recent research that approximately 98% of the word tokens in a text have to be knowable for learners (Hu & Nation, 2000). Bond (2000) also supported that learners who know less than 80% of the target words come across in the passages showed weak comprehension nearly all the time.

Schmitt (2008) indicated that most of the learners cannot succeed to reach vocabulary learning goals even moderately. Additionally, he asserted it is no longer considered valid that exposure to language tasks centering upon other linguistic aspects or communication will simply result in a sufficient lexis. Instead, he supported that a more proactive and principled approach needs to be adopted to enhance vocabulary learning.

2.3.1 Vocabulary size and L2 skills. Having good vocabulary mastery is a must to the mastery of every language skill, including both receptive (listening & reading) and productive (speaking & writing) skills (Cahyono & Widiati, 2008). In second language acquisition, vocabulary is considered as an essential element in all four skills of reading, writing, speaking and listening (Piribabadi & Rahmany, 2014).

Vocabulary size has a strong correlation with reading skill as it has with listening skill. In his more recent empirical study, Nation (2006) tested a diverse of authentic texts such as newspapers and novels in English language. He figured out that one can obtain the 98% coverage, which is believed to be necessary for an effective reading, by means of the most frequently knowable 8-9,000 word families. Consequently, it has been suggested that high-frequency vocabulary in English can go up to practically 3,000 word families; and that low-frequency vocabulary starts at about the 9,000 frequency level.

In a study conducted with 115 advanced EFL learners by Stenius (2007) a considerable correlation between a listening comprehension test (Cambridge Certificate of Proficiency in English) and a receptive vocabulary size test (the Vocabulary Level Test; VLT) was found (0.70). In parallel with the result above, Stenius suggested (2008) that vocabulary size has a crucial role in listening comprehension. Kelly (1991) also claimed that the main

obstacle for EFL learners to gain successful listening comprehension is vocabulary knowledge.

Nation (2001) put forward that vocabulary plays an undeniable role in assessing to what extent a written work is quality. Daller and Phelan (2007) carried out a study and found that measuring how rich the texts were in terms of lexis could provide a prediction about teachers' ratings of EFL essays. It was also proposed in a study by Astika (1993), for which 210 writing samples were analyzed by the consideration of organization, language use, vocabulary, content and mechanics; that the vocabulary component comprised of as much as 84% of the variance.

Stenius (2008) found it reasonable to assume that there exist a close relationship between the ratings of the compositions written and measures of vocabulary size. This assumption was supported by another study (Laufer & Nation, 1995). They correlated the scores of learners on the productive version of the VLT with vocabulary type they preferred in their compositions. The consequences of this analysis demonstrated that there were significant correlations between learners' productive vocabulary and their lexical frequency profiles. Those learners who had large vocabularies preferred fewer high-frequency words and more low-frequency and academic words in their writings than the learners who had smaller vocabularies.

2.4 Incidental Vocabulary Learning

According to Laufer and Hulstijn (2001), incidental vocabulary learning is characterized in the way that learners are not informed that they are going to be tested later on when they are given a task, by which new lexical items are acquired. The aim of incidental vocabulary learning, which enables learners to gain new words from the context with having no intention to do so, is to get an understanding of the message in the text (De Ridder, 2002). Loewen and Ellis (2009) defined incidental learning as an absence of intentionality to learn in which we might still consciously pay attention to some characteristics of the L2.

Pigada and Schmitt (2006) conducted a case study to investigate the efficiency of incidental vocabulary learning through extensive reading. In this study, where the main focus was learning spelling, meaning and structural features, they found that learners were able to improve 65% of the target words. Similarly, spelling was highly improved via a couple of

exposures whilst it turned out that meaning and grammatical knowledge was advanced to a smaller extent.

Wang (2013) collected data from 45 Taiwanese learners of English as a foreign language who were all freshmen studying at a technological college. The study was an attempt to examine if incidental vocabulary learning through extensive reading could open a road of increasing word knowledge. The participants, who had lower level proficiency English with a percentage below 50% on the placement test delivered in the summer time, were involved in EFL extensive reading program lasting 15 weeks. These participants were averagely at the age of 18. They were assigned to read 30 texts in English language with no dictionary and teacher support. To find out learners' incidental vocabulary learning gains in the 50 target words encountered within the program, vocabulary pre- and post-test were given. In the Vocabulary Knowledge Scale (VKS) pre-test, participants were asked to put a check mark in an appropriate column to sign if they had encountered or knew these 50 target words. Also they were required to write down English synonyms, Chinese translation of the word meanings, or to produce sentences. In the VKS post-test administered unexpectedly four months after the pre-test, the same 50 target words in the pre-test were measured, listed in a different order this time. In accordance with the results, a great deal of vocabulary was gained by the participants after they complete the EFL extensive reading program. Their mean score was found 59.20 in the pre-test and 64.98 in the post-test. Thus, the EFL extensive reading treatment had an impact upon incidental vocabulary learning gains of the participants.

Pellicer-Sanchez (2016) reached similar findings in her eye-tracking study on the efficiency of reading for incidental vocabulary acquisition. Asserting that reading is a significant source of incidental vocabulary learning in second language (L2), Pellicer (2016) examined incidental L2 vocabulary acquisition from online reading of unknown lexical items besides reading. Data was gathered from 23 L2 and 25 L1 participants. Ranging in age from 19 to 21 years, L1 participants consisted of undergraduate students of an institution in U.K. L2 participants, advanced learners who had experienced the requirement of English proficiency for university entry, comprise postgraduate students and postdoctoral researchers at a university of U.K. Their ages were between 22 and 42 years old. As methodology, she used a combination of online and offline measures, which mean in turn, eye-tracking and vocabulary tests. The online measures were obtained via monitoring participants' eye movements by means of a head-mounted SMI Eye link I eye-tracker in four categories: first fixation duration, gaze duration, number of fixations and total reading time. As for the offline

measures, three vocabulary tests (form recognition, meaning recall and meaning recognition) were used. As a consequence, this study demonstrated how effective reading is for acquiring a few components of vocabulary knowledge incidentally.

Zeeland and Schmitt (2013) provided a thorough analysis of vocabulary knowledge gains from listening. Postgraduate students at a British university, 30 participants in total were involved in this study. Because of minimum acceptance level of TOEFL iBT or IELTS (100/6.0), it can be said with assurance that they already had high-intermediate or advanced level. The participants, who came from 17 various first languages, were divided into two groups: After receiving the listening input, 20 participants were administered a post-test immediately while 10 of them were given a two-week delayed post-test. In the study for which 4 different listening passages and 24 target words were selected, Zeeland and Schmitt (2013) measured participants' knowledge of the three vocabulary dimensions with form recognition, grammatical knowledge and meaning recall tests in order. In the form and grammar tests which were multiple-choice participants were requested to meticulously read the options before marking one and to prefer 'I don't know/remember' option if they had no idea about the correct answer. The results of this study indicated that one source of incidental vocabulary learning is listening in second language; and that learners could enhance knowledge of 29% of the non-words urgently after listening and retain 19% of the words until after two weeks.

One of the studies revealed that incidental learning may not have positive results. In the study conducted by Waring and Takaki (2003), an immediate multiple-choice test was administered to Japanese participants. The findings demonstrated that among 25 words the meaning of 10.6 was recognized. Nonetheless, three months later, the score of meaning recognition dropped to 6.1, which points out that through incidental vocabulary learning from reading, it seems possible for learners to recognize words to a certain extent, not fully.

Another study carried out (Pellicer-Sanchez & Schmitt, 2010) was about the effect of incidental vocabulary acquisition from an authentic novel on L2 learners' knowledge on the spelling, word class and meaning of words in the text. 20 Spanish learners of EFL, ranging in age from 23 to 26 years old, took part in this study and 34 target words were selected. To evaluate the role of frequency, target words were grouped considering how many times they appeared in the text. As measurement instruments, a recognition test of spelling, a recall test of word class, and both recognition and recall measures of meaning knowledge were created.

Multiple-choice test and structured one-to-one interview constituted the form of the test. First, the target words' spelling was assessed with a multiple-choice test. Second, word class was measured in the way that the participants were asked to write the part of speech. The assessment of knowledge of meaning was involved. Cards, each of which had one of the 34 target words, were shown to the participants and they were asked to say everything they recalled about the meaning of the target word as the interviewer was writing down notes. The results revealed that incidental vocabulary learning may occur from an authentic reading only with 43% meaning recognition.

2.5 Intentional Vocabulary Learning

Intentional vocabulary learning plays a significant role in L2 vocabulary acquisition since incidental vocabulary learning may not adequately have positive results. Intentional learning is a way of vocabulary learning with an effort consciously and purposely to place lexical items into memory (Barcroft, 2009).

According to Schmitt (2008), intentional vocabulary learning, on condition that we aim to teach vocabulary specifically on an explicit focus, brings about better and faster way of retention and gaining productive levels of mastery. Elgort and Nation (2010) claim that deliberate techniques such as word-lists, words cards and learning from vocabulary notebooks, are beneficial approaches to achieve second language vocabulary acquisition effectually. Likewise, it was asserted by Elgort (2011) that incidental learning is not as efficient as intentional learning of vocabulary in that it necessitates long-term and extensive exposure to linguistic input and this type of natural language learning condition is nearly impossible to experience in English. In opposition to incidental learning, intentional learning of vocabulary forces the pace of the process of lexical development on account of repetitions and memorization strategies, which can be accomplished by learners on their own in a short period time.

It has been suggested that vocabulary teaching not only refers to teaching specific words but also providing needful strategies for learners to improve their vocabulary knowledge (Hulstjin, 1992). Laufer (2005) explained the necessity of an explicit approach by listing reasonable causes. For one thing, the exact meanings of individual words do not even arouse learners' attention when they get the overall message. Secondly, it is most of the time unreliable to guess the meaning from context, in especial, provided that 98% of the words in the discourse are unknown by the learner. The third reason why vocabulary requires an

explicit approach is that words might not produce enough engagement to be acquired and recalled if they could be readily guessed from context. Lastly, it is a must to meet new words again and again after we have met in discourse in order to avoid forgetting; and this would be possible only when learners have to read 1-2 graded readers every week. Unfortunately a typical learner, as might be expected, would not read that much (Laufer, 2005).

It was also asserted that learners need to be taught vocabulary strategy instruction explicitly so that they can benefit to learn by themselves outside the classroom (Atay & Ozbulgan, 2007). The necessity and effectiveness of vocabulary learning methods has been accepted in both L1 and L2 instruction (Chen & Hsiao, 2010). Alqahtani (2015) believed that if English teachers want to introduce a new vocabulary or lexical item to their students, it would mean that they actually want their students to recall new vocabulary. In that case, the new vocabulary has to be learnt, practiced and revised explicitly so that students can be prevented from forgetting.

Schmitt and McCarthy (1997) suggested several strategies under three subtitles as follows: (a) guessing from context, (b) benefiting from word cards and mnemonic techniques to remember words, and (c) using vocabulary cards to recall the word pairs of foreign language and native language. Proposing three strategies to teach vocabulary just as Schmitt and McCarthy, Murcia (2001) recommended that an appropriate strategy among mnemonic devices is the keyword technique in which the keyword is reminded to the learner when he or she sees or hears the target word.

So as to deepen their receptive and productive mastery of vocabulary items, it is necessary for learners first to remember the words well enough to recognize them (Yang & Dai, 2012). Sagarra and Alba (2006) claimed that if vocabulary memorization strategies involve deeper processing based on the meaning, they may lead to better retention of words when compared to shallow processing in which surface features of words to be recalled; such as sounds, orthography or the number of vertical lines in a word, are emphasized (Shapiro & Waters, 2005). From this point of view, one problem which needs to be solved and which students, most of the time, have to face is that they can easily forget the newly learned words. For the purpose of solving this problem, a variety of vocabulary learning strategies have been examined by researchers (Amiryousefi & Ketabi, 2011).

2.5.1 Vocabulary teaching strategies

2.5.1.1 Peg word method. According to peg word method, one can relate target vocabulary items to the items that can be easily memorized and function as pegs or hooks. Thus, target vocabulary items can be recalled easily. Peg word method functions in two stages: Firstly, the teacher asks students to remember 10 number-rhyme pairs such as *one is john, two is shoe, three is tree,* etc. Secondly, she or he asks them to visualize the word and try to associate it to rhyming words. For example if the teacher wants to teach '*exploration'*, the peg word for it will be '*John'*. After the meaning of the word is introduced to students, they can visualize that *John is doing oil explorations* (Thompson, 1987).

2.5.1.2 The loci method. Through the loci method, which is actually the oldest mnemonic method, a familiar place (e.g. a room or a house) is imagined and then every word is associated with a part of it to be recalled. For instance, suppose that the target words to be taught are *era, mission, density, disturb, distant, artificial, etc*. In this case, the location we are familiar with could be '*the moon'* and the mental picture may be made as follows: "In the robot era, some robots having artificial hands and legs live. They have a mission on the moon. No one can disturb them because they live in a distant area" (Thompson, 1987).

2.5.1.3 Spatial grouping. According to Holden (1999), spatial grouping works in that students learn target words by forming patterns like a triangle. In other words, they are not asked to write the words in a column; but asked to write in the form of patterns. This enables them to remember words better.

2.5.1.4 The finger method. The finger method is particularly beneficial when language teachers intend to teach children some items such as numbers, days of the week and months of the year. In this method, students are asked to link each word with a finger (Holden, 1999).

2.5.1.5 Visualization or imagery. In this method, students do not use real pictures. They imagine a picture or a scene that is linked with the target vocabulary item. To illustrate, if the new word to be learned is 'exploration', the students can come up with a mental image that '*A scientist is using special drills for oil exploration*' and they can relate it to a scientist's picture. This method seems similar to the loci method. The difference is that learners here imagine a picture or image for each target word while those using the loci method associate

the words with a familiar place and have an imaginary walk along that place (Holden, 1999; Thompson, 1987).

2.5.1.6 Use of pictures. Thompson (1987) argues that new words can be taught better if they are paired with pictures. Through this method, a picture can be benefited in order to make the meaning of the target word clear. Along with a picture, a definition of the word can be presented too. The method of pictures is more useful to learn concrete words and works better with learners at elementary or pre-intermediate levels.

2.5.1.7 Physical response method. According to physical response method, the student moves his body or just some parts of his body in a specific way so as to clarify the word meaning. For instance, the student can get up on his tiptoe and walk around the classroom if the target word to be learned is '*tiptoe'* (Thompson 1987).

2.5.1.8 Physical sensation method. This strategy was proposed by Oxford and Scarcella (1994). By help of physical sensation method, the new word is learned by associating it to a physical sensation. To exemplify, when you would like to learn the word '*frigid'*, you may feel cold.

2.5.1.9 Story-telling or the narrative chain. In this method, the target words are linked together by the learner within a story. The learner first associates the target vocabulary items with some topics; and then makes up a story which consists of the words and connects all of them (Thompson, 1987; Holden, 1999).

2.5.1.10 The keyword method. Foreign language teachers often apply a variety of strategies to make students memorize vocabulary words. One of these strategies used for vocabulary learning is the Keyword Method (KWM), a mnemonic method which was designed to enable students to learn vocabulary knowledge in a foreign language. It is one of the most commonly used methods in vocabulary learning.

Featuring in verbal linkage and visual imagery in the memory process, the key word method has been one of the most popular and effective vocabulary learning strategies in that it fosters both immediate and delayed retention of second language vocabulary (Atkinson & Raugh, 1975; Rodriguez & Sadoski, 2000). It enables learners to learn better and faster and have longer retention besides changing teacher-centered classrooms to student-centered ones (Ashoori, 2012).

The keyword method, originally proposed by Atkinson and his colleagues, involves two stages: First it makes a connection, on the basis of acoustic similarities, between a new word and a familiar word namely keyword; second establishes an imaginable link through the correlation of the target word with the keyword (Baleghizadeh & Ashoori, 2010). Tabatabaei and Hossainzadeh Hejazi (2011) pointed out that keyword method is a beneficial way to enhance vocabulary learning since a meaningful visual image for the meaning of a new word is provided through the instrument of the keyword method.

Shapiro and Waters (2005) clarified why the KWM works and can be effective with two motives; visual imagery and cognitive effort. In first place, it draws advantage from visual memory strength and it is well proved in cognitive literature that being visually stimulated builds very robust memories. In the light of strong nature of visual memory, it is not nonsense that the effectualness of the KWM may be coupled with the shape of a visual image which links the target word to its meaning. In the second place, the method may be successful because it calls for learners to exert a higher amount of cognitive effort. Given plenty of evidence that making greater effort during study brought about fostered memory retention, it is reasonable to deem the KWM works owing to the fact that it requires learners to be actively involved in attaching meaning to the target words.

Rodriguez and Sadoski (2000) stated that in Paivio's Dual Coding Theory (DCT), one of several theoretical rationales evolved to strengthen the use of mnemonic-based methods like the keyword, it was propounded that the keyword method precisely employs verbal and imaginal processes all together (as cited in Paivio, 1971, 1986, 1991). In addition, he argued that, with the help of mental images, appropriate referential interconnections are created between L2 verbal representations and the imagery system.

Atkinson and Raugh (1975) described the keyword method that it divides the learning process of a vocabulary item into two stages: In the first stage, students are requested to link the spoken foreign word to the keyword (an English word) which sounds like some part of the foreign word. In the second stage, students are asked to create a mental image of the keyword in an interaction with the English translation. Hence forth, the keyword method is known to be a chain of two links making connections between a foreign word and its English translation: the foreign word is associated to a keyword in the way that it sounds similar, which is acoustic link; and then the keyword is associated to the English translation by means of mental imagery, which is mnemonic or imagery link.

Shapiro and Waters (2005) explained the application of the keyword method this way: In the application of this mnemonic method formed by two steps, we firstly make up a keyword to be learned. The keyword to be learned must be acquainted to the students and acoustically resemble some part of the target word (e.g. the keyword *steel* in English to learn the word *stylo* in French). Then, we visualize an interaction between the keyword and the target word by imagining a big pen made of a steel girder. Students need to vividly play out the interaction in their minds. To exemplify, students should be able to picture a reputed movie star who is endeavoring to autograph using a large steel girder which shines in the sunlight. This interaction may be seen outrageous, but it should not be forgotten that the sillier the interaction is, the more efficient the method is (Shapiro & Waters, 2005).

2.5.2 Studies on intentional vocabulary learning. Barcroft (2009) conducted a study, by identifying a number of strategies frequently used by the participants, in order to assess participants' vocabulary learning performance through intentional L2 learning. The rest 93 participants, 69 first-semester and 24 second-semester students, provided data to be analyzed. The experimental words selected from 24 nouns and of which lengths could be 2-, 3-, 4- and 5- syllable words were presented on separate slides in a computer programme with a picture referring to the target word above each word. These pictures were drawn in 24 simple lines. As procedure, a pre-test, in which they were asked to write English translations next to the 24 Spanish words, was given to the participants. The participants had a view of word-picture pairs on a screen for each of which the view duration was six seconds. After the administration of two post-tests (a picture to L2 recall test and a L2-L1 recall test), they were requested to describe the strategies that they used while endeavoring to learn the 24 Spanish words. According to the results of the first correlational analysis, there was a significant positive correlation between the number of strategies used and picture-to-L2 recall score. With the second correlational analysis, it was revealed that number of strategies and L2-L1 recall score showed a considerable positive correlation, too.

Another study, conducted by Hung (2015) at a private university in Taiwan, was an attempt to investigate whether digital flashcards could be incorporated into a university course or not so that English language learners can promote their vocabulary knowledge through intentional vocabulary learning. All the seventy-five participants ranging from 18 to 20 years old were selected from three intact classes. *Study Snack*, a free website by which students can make their own flashcards related to any topic, was applied as an intentional vocabulary learning tool in this study. Three varying formats of flashcard use were involved in the study:

In a self-practice format, five flashcards were created by the participants individually and they were unaccompanied when the material was practiced. A pair-exchange format required the participants to exchange and study the flashcards with different pairs following making five flashcards individually first. In a group-based format, the participants again created five flashcards by themselves; thereafter collaboratively determined ten useful words to appear in their group flashcards to be practiced and reviewed later on. The overall findings of this study emphasize that the participants' vocabulary learning performance came out positive along with their attitudes toward the use of flashcards. The consequences of the study also revealed that this deliberate vocabulary learning technique with digital flashcards can be applied not only to provide lexical development in and out of class but also to offer learning opportunities in group.

Demir (2013) carried out a study in Turkey based on intentional vocabulary learning and examined whether the vocabulary instruction by the aid of in-class vocabulary strategies that he developed served the purpose of English vocabulary retention of Turkish 8th grade EFL students or not. Having a fair distribution in terms of success and gender, the participants were comprised of 66 students for experimental group and 63 students for control group. The thirty target words taught to the groups were newspaper cuttings taken from a local English newspaper. With the intent of measuring vocabulary retention of experimental and control groups at two different times and determining the differences between them in retention, the researcher employed independent samples t-tests and gathered the data from post-test and retention test. In the analysis of the data, SPSS 16.0 package was used. In the procedure adopted for EG, the first step was the teacher had students circle the unknown words in the heading, make guesses about the meanings by using the picture of the text and make comments on what the story told. The second step was the teacher asked students cut out 30 target words from the newspaper clipping and use them each in a sentence. At the third step, the teacher had EG students discover a new word in the body of every target word by consulting a dictionary. Students won 1 point for each discovery and the first three students getting 5 points were awarded with reinforcement. Next, they were asked to write the words they discovered on the board.

Example: Target word: Opportunity Discovered word: Port (1 point)

At the last step, the teacher invited EG students to invent new English words or names by acrostically placing the initials of the thirty target words in an order. Example: Chin +

Adult + Thrill = CAT or ACT (1 point). The main conclusion from this study was that making students involved in inventive and exploratory in-class strategies accompanied by hands-on activities was a working method of teaching vocabulary. Moreover, the scores of the post-test and retention-test demonstrated that CG students (control group) could not retain as much vocabulary as EG students did in the wake of getting instructed with in-class vocabulary strategies.

In another study carried out by Ashoori Tootkaboni (2012), the effects of three strategies on long-term retention of English vocabulary items in English as a foreign language classroom were compared. Having elementary level of proficiency, 65 female students from a language institute took part in the present study conducted in Tehran, Iran. They were at the age of 13-17. As the target vocabulary items produced from words in Oxford Elementary Dictionary, twenty English words which were carrying the criteria of being unfamiliar and concrete to the subjects were employed. For one thing, they were given a pre-test comprising of twenty- five words of which five were omitted from the research before the learning period. Then, the researcher in the keyword group comprehensively explained the keyword technique with examples. Participants had enough practice of the method and asked questions when necessary. In the same way, in the context group, the researcher gave an explanation of the strategy and a variety of factors such as contextual and word factors that would help them find out the meaning of an unknown word. In wordlist group, the researcher gave the learners a list of words as well as their translations in Farsi and asked them to memorize the list. They were given 10 minutes to learn 10 unknown words. In order to measure how capable the students were of recalling the meaning of the target words, the participants took a cued-recall test a week after the treatment. In this test, the participants were required to translate 20 English words into Persian. Just after that, word-recall test including 20 sentences was administered to them to fill in the missing vocabulary word on the basis of the context. The same testing procedure was applied to all the treatment groups. The results of the study showed that learners in the keyword group recalled more definitions (18.10) context (13.78) and wordlist (16.00) subjects did. The wordlist subjects remembered more definitions (16.00) than the context subjects (13.78). As for the group differences in word-recall tests, the keyword group had the highest mean (17.10) compared to the context and wordlist groups. The mean of the wordlist group was the lowest (11.14).

Kılıçkaya and Krajka (2010) did an experimental research with the purpose of finding out whether online vocabulary teaching would be more efficient than the traditional

instruction. How useful online vocabulary teaching and the traditional methods followed in an Academic English class with upper-intermediate level were compared. In a private university in Ankara, Turkey, 38 students studying English so as to pass the proficiency exam held by their own university participated in the research. Their ages ranged from 17 to 19. Data collection instruments were pre- and post-tests which contained 10 academic reading passages and 5 multiple-choice questions assessing the vocabulary items in those passages. After the participants were shortly informed about the purpose of the study and signed a Consent Form, the pre-test was administered by the instructors. The findings from the pre-test proved that there were no significant differences between control and experimental groups. Throughout five weeks, control group practiced the target words in ten reading passages through traditional method with the aid of vocabulary notebooks, cards and paper dictionary. On the other hand, experimental group practiced the same target words in the passages through *WordChamp*. During the five week-lasting training, all the vocabulary items were regularly reviewed with both groups. On the last day of the class, both groups were given the post-test. Data obtained from the post and follow-up tests scores were analyzed with the SPSS software package. The participants of the experimental group who used *WordChamp* while learning vocabulary items performed considerably better than the control group participants who studied the items by way of traditional vocabulary learning.

2.5.3 Studies on vocabulary learning through the keyword method. The present study is regarding teaching vocabulary items through the keyword method instruction. In this sense, a group of studies with regard to the effect of the keyword method in vocabulary learning are mentioned below.

Baleghizadeh and Ashoori (2010) carried out a study to compare the effect of keyword and word list methods on immediate retention of English vocabulary in a classroom setting. Averagely at the age of 13.5 and being students at Ansari Junior High School in Astara, Iran, 44 female participants took part in this study. The participants were randomly selected from 2 intact classes that came together twice a week and each of which lasted 90 minutes. All the participants had elementary level in English language proficiency and had no idea of the keyword method before. For the experiment, a list of randomly-decided-20 English words along with corresponding keywords and their Persian equivalents were made. It was compulsory that the words were concrete and the participants had not come across any of the target words previous to the experiment. The two classes were randomly assigned as the experimental group (EG) and the control group (CG). The CG participants were given the list

of 20 words in total (each session 10 words) containing the target words and their Persian equivalents; and they were asked to study these words by using rote learning method for two sessions. The EG participants received completely the same words; but with a difference that there was a keyword (sounding similar to the English word) between every English word and the Persian equivalent. After the participants in the experimental group were explained the keyword method, they were asked to practice the method, and then to learn the vocabulary list by using the provided keywords. At the end of the second session, both groups were administered the same examination sheet including the whole English words which they were taught for two sessions; and they were given 20 minutes to write the definitions of all the target words in Persian. The data was analyzed through totalizing the correct answers. According to the analyses, the mean score of the experimental group is highly greater than the mean score of the control group. Thus, the findings released that the experimental group, which received vocabulary instruction through the keyword method, considerably outperformed the control group, which received vocabulary instruction with the use of wordlist method.

Campos, Pinal and Fabello (2014) also conducted a study to assess the efficacy of keyword mnemonics on bilingual students' receptive and productive vocabulary recall in the learning of their non-dominant language. In total 237 bilingual students, 116 of whom are women and 121 of whom are men, participated in this study. The third and fourth year Compulsory Secondary Education students from Galicia in Northwest Spain, the participants ranged from 14 to 17 in age. All of them, along with their parents, spoke Spanish not Galician language. They spoke Galician only in some classes where the instruction language was Galician. Thirty concrete Galician words were chosen for the learners who did not speak Galician more than 10% of the time. Selected from the entire sample, a group of 35 participants were assigned to produce the keywords for the study. This group was provided with 30 Galician words as well as their equivalents in Spanish; and then was asked to generate a keyword for every Galician word. The keyword needed to be a concrete Spanish word that is acoustically similar to the Galician word. Furthermore, the participants needed to write a sentence regarding the Galician word to its Spanish meaning. An example was Galician word *boxo*, keyword *bollo* (loaf in English), equivalent Spanish word *barriga* (belly in English). And the interactive sentence was "*imagine a belly full of loaves*". Randomly selected, only 20 of the 30 Galician words were used in the study. The remaining 202 students were divided into two groups in the way that 102 applied keyword mnemonics and 100 used the rote

method so as to learn the meanings of Galician words in Spanish. The researchers both showed and read aloud twice to the rote method group every Galician word and the equivalent Spanish word. The keyword method participants were demonstrated the Galician word, the keyword, the equivalent Spanish word and the interactive sentence (all of them were read aloud twice by the researchers later on) containing the keyword and the Spanish word. Then, the participants were given a worksheet of randomly-listed Galician words which they were asked to match with their corresponding in Spanish in a maximum of 2 minutes (receptive recall). After that, the participants were given randomly-listed Spanish words and asked to write the corresponding word in Galician in a maximum of 2 minutes again (productive recall). The whole participants were provided with 4 trial words before the presentation of the real list. Lastly, the participants were applied the Spanish version (Perez-Fabello and Campos 2004) of the Gordon test of Visual Imagery Control (TVIC, Richardson 1969) and were not given any time limit. The analyses released that participants using keyword mnemonics recalled the meaning of more Galician words than those who used the rote method. It also turned out that image control had an influence on recalling the meaning of Galician words. Participants with higher image control ability performed better recall of the meaning of Galician words compared to those who had lower image control ability.

Sagarra and Alba (2006) did an experimental study to investigate how effective three vocabulary learning methods; rote memorization, semantic mapping and the keyword method. The participants consisted of 778 third-semester learners of Spanish as a second language from a university in U.S. Totally 24 experimental Spanish words were included in the materials of the present study. These words were supposed to be highly concrete and readily imageable. Prior to the treatment phase, students were asked to complete a pre-test in order to assess if they knew the target words previously. In the pre-test they were requested to write English translations of Spanish experimental words. The researchers informed the students that they would be shown 24 new Spanish words with their equivalents in English; this would appear in three sets of eight words each and they would need to complete a recognition test in the wake of every set. Firstly, in the rote memorization set, students were instructed to write the Spanish words with their translations and not to establish any other links. Secondly, in the keyword method set, the researchers asked participants to link the Spanish word to an English word which sounded similar to some part of the Spanish word, and then write down the connection. In the final treatment, the semantic mapping set required students to learn eight words by diagraming semantic associations belonging to the target word. The participants all

got exposed to the same target words, vocabulary learning methods, instructions and examples. The only difference was that the groups were provided with different presentation order of the learning methods besides the eight words in each set. After all above, participants were administered six posttests with 8 items each; three for the immediate posttest and three for the 3-week delayed posttest. In every posttest given, participants were instructed to match pictures with eight Spanish nouns. For the data analysis, the researchers scanned the bubble sheets and assigned one point per correct answer. Also, the researchers excluded the incomplete bubble sheets and the ones including more than one answer per item. As a result, mean percentages revealed that the keyword method was the treatment providing more long-term effects compared to the other two. This meant that vocabulary learning methods calling for deeper processing through form and meaning associations, like the keyword method, yielded the most retention.

However, another study conducted by Rodriguez and Sadoski (2000) examined the impacts of rote, context, keyword and context/keyword methods on immediate and long-term retention of English vocabulary as a foreign language. A total of 160 ninth grade students, members of eight intact EFL classes from two different schools in Venezuela, participated in the study carried out in a natural classroom setting. Their ages varied from 13 to 18 years. The students had been studying English a foreign language for more than 2 years. Also, two experienced female instructors participated in the study and applied two different methods to teach target words in four intact classes. Before learning started, the eight intact classes were provided with a training session where they were shortly told the importance of vocabulary in foreign language learning. Participants were asked to learn 15 new English words, in groups, in the rote rehearsal condition, in the keyword condition, in context condition and in the context/keyword instruction. Participants of the immediate condition were administered a 2.min multiplication task; this was followed by a test of cued recall. On the other hand, the multiplication task was not given to participants of the delayed condition. Instead, they were administered the cued-recall test after one week. On the cued-recall test there was a list of the 15 vocabulary words along with spaces to write the Spanish translations. The results of this study revealed that participants performed highly better on immediate cued recall than on delayed cued recall. Moreover, the context/keyword method proved the superiority over the other methods particularly in the delayed condition. The context/keyword group turned out to be superior to the keyword in both the immediate and delayed recall.

Atay and Ozbulgan (2007) suggested that explicit instruction on vocabulary learning strategies may enable students to become more competent in the process vocabulary learning. For this purpose, they (2007) carried out a study on the effect of memory strategy instruction besides contextual learning on the ESP vocabulary recall of Turkish EFL students; and they discovered if there was any difference in the use and preferences of learners' memory strategies as a consequence of such instruction. The participants consist of fifty male Army Aviation pilots whose ages ranged from 23 to 35. Also, being graduates of the ELT departments in İstanbul, Turkey, and native speakers of Turkish, two teachers participated in this study. Two intact classes were randomly assigned to be control and experimental group. To collect the data, a multiple-choice vocabulary knowledge test (VKT) prepared by the researchers was administered. The test was given to both groups on the first and last day of the course. In the VKT, each item of which was worth two points, there were a total of 50 items. In addition, the frequently used memory strategies questionnaire (FUMSQ) was given to the participants in the experimental group as a pre- and post-test. The participants here were told they could mark more than one strategy. As for the procedure, both groups took six hours of English classes a day. While the students of control group completed one unit in six hours by spending more time on role-plays and listening tasks; the experimental group students finished all these activities in five hours and got one hour-memory strategy instruction. On the first day of the program, the teacher gave an introductory lesson to the students in the experimental group on vocabulary learning; and talked about the necessity of vocabulary knowledge in language learning. Then, he continued with vocabulary strategies focusing on all the memory strategies given. For instance, he explained a strategy and presented examples of it. On the following day, in the wake of five hour-regular instruction, a handout of memory strategies was distributed to students by the teacher. For the rest of the study, the experimental group students looked at the target words and tried to apply the strategies they believed to be the most efficient and suitable for both learning and retaining. In the last hour of each week, the students did an analysis of which strategies worked and which did not, and why. Also, they reflected their difficulties and successes they came across in using the strategy out of the class. The results of this study revealed that there was a significant difference between the vocabulary knowledge gain scores of the control and experimental groups. Moreover, there appeared an increase in the percentage of use and variety of strategies in the post-test. And it was seen that the most commonly used strategies

by students were “connecting the new word to a previous personal experience” and “use of semantic maps”.

Another study by Tavakoli and Gerami (2012) investigated in what way the two different mnemonic approaches of vocabulary teaching; which were the Keyword method and Pictorial method, affect learning and retention of lexical items. The selected participants were all female EFL learners whose native language was Persian. Ranging from 18 to 32, a total of 60 learners comprised the participants of the study. They were categorized into three equal groups: The first experimental group (EG1), which was expected to receive the keyword method treatment; the second experimental group (EG2), which was to receive the pictorial method instruction; and the control group (CG), which did not receive any effective instruction. For data collection, four instruments were used. These were, in order, the Nelson test, pre-test, immediate post-tests and delayed post-test. Before the experiment, the tests were piloted with some students who resembled the participants of the study. For one thing, the Nelson test with fifty items was given to make sure of homogeneity of the learners; and the final participants of the study were selected according to the results of the test. Second, the participants were administered the pre-test and asked to write the translation or the English definition of each word in Persian; or put a tick if they were familiar with the word but could not remember. In the EG1, the researcher introduced the keyword method in the first treatment session. In the EG2, the pictorial method was introduced in the first session. In the CG, the researcher handed out papers including same words as the EGs and asked the learners to memorize these words through reading the Persian corresponding of every word written in front of it. Just after the treatment, the immediate posttests were applied to measure participants’ short-term memory regarding the words they were taught on the same day. At the end of the study, the delayed posttest was administered for the measurement of the students’ long-term memory regarding the words instructed. They were expected to choose the word they thought to be the most appropriate among other choices; and to complete the meaning of the presented sentence per item by marking it in the answer sheets. According to the results, the EG1 performed much better than the other two on the posttest. Additionally, the EG2 did better on the test than the control group which had the lowest score. The keyword method proved to affect the participants’ retention of lexical items positively not only in short-term memory but also in long-term memory.

Piribabadi and Rahmany (2014) made an investigation to compare the effect of the instruction through the keyword method and word-list method on ESP vocabulary learning.

The treatment was done at a university in Iran and 120 participants were randomly selected from industrial engineering students with intermediate level. In both classes, the participants, who were not different from one another in terms of educational and language background, were grouped to be 30 lower intermediate and the other 30 upper intermediate students as a means of Oxford placement test. Thirty-two vocabulary items were chosen from the ESP course book of Industrial Engineering. A pre-test was administered to both groups so as to distinguish between students' mastery of English vocabulary knowledge. After that, the entire population of every class received two different treatments lasting for four weeks. The students of the first class were taught ESP vocabulary items by means of the keyword method instruction; and the students of the second class were taught through the word-list method instruction. Lastly a multiple-choice post-test was given to all the participants in order to examine the effects of the treatments in both groups at the end of the study. To analyze the data, the researchers applied a two-way ANOVA. According to the results, the upper-intermediate students who were given the keyword method instruction obtained higher mean scores when compared to the upper-intermediate students who were given the word-list method instruction. In other words, the upper-intermediate students in the keyword method group performed better than those in the word-list method group. The same results appeared for the lower-intermediate students too. The ones in the keyword method instruction had better performance than those in the word-list method class. In summary, the application of the keyword method instruction to both upper-intermediate and lower-intermediate level became more effective than the application of the word-list method instruction.

Chen and Hsiao (2010) examined the efficiency of the keyword method training in ESP vocabulary instruction. Randomly assigned as the keyword and traditional groups, forty students from two intact classes participated in the study which was conducted at a technology university in Taiwan. All the participants used the same course book called English for Business because of the fact that they were all taking Business English. As the keyword training material, a set of 15 Chinese-English pairs were selected and the participants were asked to complete a word recognition list with the intent of ensuring that all the selected ESP vocabulary items were unfamiliar to them. During the treatment, the experimental group was given the keyword method training while the control group was taught through traditional vocabulary presentation. On the first day of the experiment, the logic of the keyword method was explained; and the students were introduced the new vocabulary word. This was followed by a mnemonic picture and keyword linkage for this new

word. After that, the Chinese meaning of this new word was shown to the students. On the second day, participants were demonstrated the 15 target words in a random order. They were presented each picture and then the Chinese meaning of every new vocabulary word. On the third day of the experiment, the instructor had students review the 15 target words with brief repetitions of the second session. Finally, all the participants were instructed to recall their meanings by studying the words on their own. On the other hand, the control group received traditional vocabulary instruction via repetition, translation and example sentence explanation along with definitions and synonyms. After the experiments, the instructor administered an immediate vocabulary knowledge recall posttest to all the participants and asked them to write the Chinese equivalents of the 15 English words by presenting no keyword or mnemonic picture. Following the posttest, the instructor also requested all the participants of the keyword group to write down their perceptions toward vocabulary learning through the keyword method. As a consequence, it was revealed that students' pretest and posttest performances increased 8.6 words in the keyword method instruction whereas they increased 4.15 words in the traditional method instruction. This suggested that more beneficial effects were observed in the keyword treatment than the traditional treatment. The students of the keyword group turned out to gain and recall more target vocabulary items following the instructional treatment.

The purpose of the present study is to investigate the effect of the keyword method on ESP vocabulary learning and recall. Having worked in Civil Aviation Cabin Services Program of the private university approximately for three years where I conducted this quasi-experimental study, I noticed that the aviation students' level of English vocabulary knowledge is too inadequate to do their job as flight attendants. Considering that these students are supposed to have communicative competence in English language when they start work as flight attendants in an airline company, their lack of ESP vocabulary knowledge I observed in my classes led me to this study. The following research questions are addressed in this study:

1. Will there be a significant difference between the students who are instructed in the keyword method and those who are instructed through the traditional method in the retrieval of ESP vocabulary knowledge?

   1.1 Will there be a significant difference in the retrieval of ESP vocabulary knowledge at receptive and productive levels?

2. What are the EFL learners' perceptions toward learning ESP vocabulary through the keyword method?
3. Is there a difference between the experimental and control groups in the use of target vocabulary in oral communication?

Chapter 3

Methodology

3.1 Philosophical Paradigm: Quantitative and Qualitative Research

A paradigm can be defined as a fundamental set of beliefs guiding an action (Guba, 1990). The word paradigm also refers to a number of research methods, criteria and viewpoints that researchers adopt in the process of producing knowledge (Fossey, Harvey, McDermott, & Davidson, 2002). Philosophical paradigms consist of a point of view toward the nature of reality (ontology), how the researcher knows what he or she is knowledgeable about (epistemology), the roles that values in the research have (axiology), the research language (rhetoric) and the methods adopted in the process (methodology) (Creswell, 2003).

In this regard, quantitative and qualitative methods are built on a foundation of philosophical assumptions; and these assumptions may differ from one another by the methods used. To make it clear, quantitative research is originated from positivism in that it is designed by testing a theory or hypothesis, and with surveys, scales and questionnaires.

Qualitative research, on the other hand, is established on a base where researchers reflect their own worldviews, paradigms and sets of beliefs to the research (Creswell, 2007). There are a number of qualitative research designs. Among the most commonly used ones, interpretative studies, case studies, ethnography, narrative inquiry, historical studies, content analysis and phenomenological studies are major approaches (Ary et.al, 2010).

In consideration of the philosophical paradigms reviewed above, a quasi-experimental research design was adopted in the present study. In this approach, sampling of the participants is difficult to do randomly. This is because the existing participants of the study are assigned to experimental and control groups, which are two intact classes.

3.2 Research Design

In the present study, a quasi-experimental research design was adopted with the purpose of investigating the effect of the keyword method on vocabulary retention in short-term and long-term. The data was collected through the use of mixed method, namely, both quantitatively and qualitatively. On the basis of experimental research design, the participants of the experimental group received keyword method treatment whereas the participants of control group were taught by traditional method. To collect quantitative data, the students in

both groups were administered a vocabulary knowledge scale (VKS) developed by Paribakht and Wesche (1993) as pre-test, posttest and delayed-test. The students' vocabulary knowledge and retention was tested with the help of the quantitative data, which was obtained through pre-, post, and delayed tests given to the students in experimental and control group. Also, structured interviews were done with students in the experimental group in order to gather qualitative data, which was obtained to reach more reliable findings. The qualitative part of the research design here explored students' perceptions about the efficiency and criticism of the keyword method along with, in general, what opinion they are of about the keyword method.

3.3 Target Population and Participants

The present study was conducted with 32 students who have been studying in Civil Aviation Cabin Services Program of a private university. They range from 17 to 25 in age. As for gender, 21 of them are female and 11 of them are male learners. Their consent was asked for this experimental study. They are all second year students and already knowledgeable about how to use basic structures of English on the ground that they received English courses at elementary and pre-intermediate levels when they were studying in their first year. Since the current study is an experimental one, participating students were selected from two intact classes whose English levels were found same in the wake of a placement test. Each of these two classes consists of a total of 16 students participating in the present study. In other words, among 32 participants, 16 students are learners of daytime education while the rest 16 are registered to evening education.

In the Civil Aviation Cabin Services Program, English classes are divided into three groups for daytime education and two groups for evening education. The first groups are always the most proficient ones in terms of English knowledge and the present study was conducted with the most proficient group of both daytime and evening education. Being taught intermediate level English besides Aviation English, the second year learners have 70 points and above out of 100 in the exams held so far.

3.3.1 Setting. This experimental study was carried out in Civil Aviation Cabin Services Program at a private university in İstanbul, Turkey. The 2 year program educates flight attendants and serves within the scope of vocational school. To enroll in this program, students are requested to attend the university entrance exam. Approximately 100-120 students are yearly accepted to the program and allowed to do registration. Around 80-85 of

these students are enrolled in daytime education and the remaining 35-40 students, who receive an acceptance with a bit lower score from the university exam, have to study in evening education. The purpose of the program is to prepare students to become flight attendants and to ensure passengers' safety and security on domestic and international flights. The program also aims to help them reach B2 level English knowledge and actively use it in their professions. Civil Aviation Cabin Services students are also provided with a mock-up classroom to develop their practical skills in departmental courses. It is equipped with an aircraft model including economy and business class, window and aisle seats, lavatory, galley, overhead lockers, seatbelts and even tray-tables. The graduates of the program are awarded diploma with the intent of becoming flight attendants and get a job in airline companies.

The program is as follows: Two academic years are divided into four academic terms as fall and spring. Education is, in principle, given under the titles of *Cabin* and *English*. When the list of cabin courses namely departmental courses is made, in the first year/fall term, students have to take *Introduction to Civil Aviation, Organization and Responsibilities at Aviation Companies, Courtesy and Protocol Rules, and Aviation Terminology*. Additively, they have to select two compulsory courses – *Turkish language I and History of Turkish Republic I* – along with one elective course between two options – *Social Responsibility or Information and Communication Technology*. For the spring term of the first year, compulsory courses that students have to take are *Basic safety rules, Communication Skills and Basic Aeroplane Instruction* besides *Turkish language II and History of Turkish Republic II* courses. Finally, to be able to complete the freshman year fully, they have to select *English Speaking Skills* which is a restricted elective course in this program. When they become sophomore, they are responsible for five compulsory courses in the fall term, the third term of the overall 2 year-lasting program. These are called *Emergency Safety Rules, Elocution and Announcement, Medical Issues and First Aid, Communication and Passenger Affairs, and Basic Service Codes in Cabin.* Moreover, *Ticket and Passenger Services,* a departmental elective; and a second foreign language as an elective course, *Arabic I or Spanish I,* are other required courses of the third academic term. In the last term of the second year, students have to do course registration for *Personal Development, Crew Resource Management, Hazardous Material Instruction, Basic Airport and Flight Security, Interview Techniques* and *World Cultures*. All the courses listed except English are instructed in Turkish in the present program and aim at enabling students to gain theoretical knowledge and practical skills besides being able to integrate these two. One advantage of studying in this program for

existing students is that cabin classes are always divided into two groups because of excess classroom size. This division system is valid for each course in the program and is intended for making students' performance better and offering more particular education service to students.

Civil Aviation Cabin Services Program does not have an English preparatory school for the students within its own structure. All freshman students are asked to take 10 hour-Basic English course weekly, which is instructed in elementary for fall term and pre-intermediate level for spring term. The program students have to registrate in another English course taught weekly 2 hours only. The name of the course is *English Speaking Skills*. This departmental elective course appears in the course registration list on the system for the spring term of the freshman year only. It aims at enabling first grade students to gain speaking ability with the use of basic sentence structures and prepares them, in advance, for the next year's Aviation English course which requires being much more communicative. In the second year, students are asked to take 8 hour-Aviation English course per week. There are a total six English instructors who have currently been working in this program. All of them were appointed to Civil Aviation Cabin Services Program of the university after they completed at least 1 to 3 teaching years at English Preparatory School within the body of the same university. Since then, they all have been teaching Aviation English courses along with Basic English courses in the related program. Currently, the six instructors of English language have been working here voluntarily. Just before they started their academic teaching year in this program, these instructors received a 3 week-lasting training which was held by Education Centre of Turkish Airlines. The primary purpose of this training was to provide guidance to the instructors on to what extent English knowledge is expected from flight attendant applicants. They were also suggested the textbook that they would be using for Aviation English course. This ESP textbook, selected in cooperation with Education Centre of Turkish Airlines, is called English for Cabin Crew (Gerighty & David, 2011).

English classes in Civil Aviation Cabin Services Program are grouped in accordance with the results of a placement test administered to students at the very beginning of a term. Both first and second grade students in daytime education are grouped into three whereas the students in evening education are grouped into two. Thus, the program has two groups for first grade students and another two groups for second grade students. These groupings are sorted in the way that group 1 is the best and group 3 is the worst in English knowledge. For the continuity of education quality, all the groups are re-determined just before the beginning

of each academic term. This is because there might be an increase or decrease in students' English knowledge and language production performance in time.

Assessment of the program students in this study is done through the evaluation of five items. These are, in turn, a quiz (15%), a mid-term (25%), a speaking project (10%), teacher evaluation (10%) and a final exam (40%). First, a quiz covering 15% is given to students a couple of weeks before a mid-term exam. The quiz, averagely 2 pages, measures students' vocabulary knowledge and ability to use English, namely grammar knowledge. The reason why the quiz is previously applied is to present students the general format of mid-term exam following a couple of weeks. The only difference between quiz and mid-term exam is that there is no listening and reading skill to be measured in the quiz while they are both included in the mid-term. However, students often engage in listening and reading activities on each page of their textbooks within English classes. Apart from this, the quiz and mid-term exam share a lot of similarities regarding the content. For instance, they both include a vocabulary part which has matching words or phrases with their definitions, completing the missing letters of a word or phrase and filling in the blanks with the words or phrases in the box. Additionally, English grammar knowledge is measured by a number of questions such as circling the correct option among the given choices, making positive, negative and question forms by putting the given words in an order, and writing the verbs in brackets in a correct tense form. As in mentioned above, in the mid-term exam, reading and listening skills are particularly assessed. Question types for both can generally change from marking statements as true or false to choosing the correct option, and from filling in the blanks with one word or phrase from the text or listening to answering the questions fully in a written way. 3-4 weeks after the mid-term, students are asked to submit a speaking project which affects the overall grade 10%. They are supposed to record a 5-10 min-video with their classmates in groups of three or four. The target language students are requested to use here is already presented by English teachers when the project is first announced. This speaking project not only encourages the program students to be involved in a team work by helping them to learn from each other but also guides them to productively use the language they have been taught. The fourth assessment criterion is teacher evaluation which is of a proportion of 10%. This evaluation is calculated by the teachers who co-teach the same class. Each of the teachers individually gives a point to every student in the class and then the points given by each are totalized and averaged. The averaged point is the teacher evaluation of that student. In teacher evaluation, students' in-class performance, their attitudes and behaviors, and how actively

they take part in-class activities are considered. Also, it is essential to attend classes regularly so as to get a high teacher evaluation point. At the end of the term, students have to take a final exam with the highest proportion of 40% among all the criteria above. In final exam, they are asked all the same parts in terms of content as those of the mid-term exam, but this time the whole questions are multiple choices. Writing skill has the least necessity and importance to aviation students when the objectives of Civil Aviation Cabin Services Program and the reason behind its establishment are taken into consideration.

To clarify, the primary goal of aviation students is to be able to acquire English language which is common within an aircraft cabin and to produce it accurately and appropriately in communication both with foreign passengers and colleagues aboard. For this purpose, they are taught terminology and vocabulary knowledge used in aviation sector. They also learn vocabulary knowledge and phrases necessary during food and drink service, meeting passengers' needs and dealing with passengers' problems. Besides all these, they are expected to know how to make announcements in English and how to give instructions which are essential to foreign passengers during take-off and landing.

3.4 Procedures

In this section, I will provide you with detailed information about data collection procedures mentioning type of sampling, data collection instruments, data analysis, and reliability and validity of the study.

3.4.1. Data collection procedures. To begin with, the researcher got permission from the Ethics Board Committee, the Head of Vocational School and lastly the Head of Civil Aviation Cabin Services Program at the private university where the present study was conducted. The whole data was collected during October and the first half of November 2017; and analyzed in the middle of November.

According to their results in the placement test held at the beginning of the term, two intact classes (one daytime and one evening education) were assigned to the experimental group (the keyword group) and the control group. Both groups had the same instructor and the same teaching materials besides teaching hour and duration. A couple of weeks before the treatment, the placement test taken from the Teacher's Resource of New Success Intermediate Course Book (Pearson Publishing) had already been applied to identify the homogeneity of their English level. This test, administered by Civil Aviation Cabin Services Program of the

university where the present study was conducted, consisted of a hundred multiple-choice questions. It measured Aviation students' grammar and vocabulary knowledge. No listening, reading, writing or speaking skill was included in the test.

A few days before the treatment, all the participants were given a vocabulary knowledge scale (pre-test) including 17 target words/phrases so as to evaluate their vocabulary knowledge prior to the training. During the treatment, the experimental group was instructed through the keyword method while the control group was taught under traditional vocabulary learning method. The experiment of teaching the target ESP vocabulary items was completed in four days below, each of which was once a week.

On the first day of the experiment, the instructor delivered a speech on the importance of vocabulary knowledge in learning a foreign language. Then, she explained the role that the keyword method plays in memorizing vocabulary and illustrated the logic behind the keyword method. Just after, the teacher modeled the keyword method for her students: For one thing, she taught the students keywords for two sample words which she had previously selected from the aviation course book *English for Cabin Crew* (Gerighty, 2011). Secondly, she presented a mnemonic picture and established a keyword linkage for these words. A story for each word, which is preferably silly, was told by the teacher. The stories also included the keywords for the new words and the meanings of the words. Then, their Turkish meanings were declared by the teacher.

On the second experiment day, only 10 of 17 target words/phrases, each of which had been put on a separate slide of PowerPoint in advance, were taught to students. The teacher started the treatment by repeating the previous two samples in the process of introducing the keyword method. Then, she continued by presenting the first target word/phrase on the first slide. The target word was accompanied by a reminder keyword in students' native language (e.g. HOLD = LD sigara). At the top of the slide, there appeared a picture that can be of more than one image. A short story with 5-6 lines was demonstrated just below the picture. There was a relationship between the story scenario and the images. Lastly, an example sentence appeared at the bottom of that slide in order to show how to use the target word in a sentence. So far the teacher read each items on the slide aloud too while students were reading what is on board and looking at the visuals. All the stories and example sentences were made and written within aviation context. The teacher continued each target word/phrase on the slides in

the same way until she finished 10 target words. So, session 1 of the keyword experiment was completed.

On the third day of the experiment, the teacher made repetitions of the first session's 10 target words/phrases with the experimental group. All the vocabulary items from 1 to 10 were repeated in the way that they were first taught with visuals, silly stories and example sentences. On the fourth day, session 2 was completed by the teacher. In this session, the remaining 7 words/phrases were taught in the same procedure as followed in session 1. There were 2 hour-English classes. In the second hour the repetitions of the whole 17 target vocabulary taught so far were made while the experimental group learnt the rest 7 words/phrases in the first hour. To exemplify, the teacher re-read the stories belonging to each target vocabulary by opening every slide on PowerPoint one by one. She showed the pictures to the students again by connecting them to the keyword and the story.

The control group learnt the words and phrases through the traditional vocabulary learning. The students were provided with the word meanings by the presentation of definitions or synonyms. For instance, they were asked to match the words a-g with their definitions 1-7. Except for vocabulary learning method, everything else was applied to the control group exactly the same (treatment date, the target words, repetitions, pre-, post-, and delayed- tests, etc.)

A day after session 2 ended, the instructor immediately gave a vocabulary knowledge scale posttest to both the experimental and the control group in order to measure how much they were able to recall. The posttest was completely the same as the pre-test administered before the treatment. It consisted of the 17 English words and all the students were kindly asked to write down the Turkish equivalents for the English words selected from aviation context. In addition, they were requested to use them in a sentence if they could. After the posttest, the keyword group was interviewed and asked the following questions: 1) Do you think that the keyword method is effective to memorize vocabulary in aviation English? If yes, why?, 2) What would you criticize about the keyword method?, 3) What are your general opinions of the keyword method? Can you write your perceptions toward learning aviation vocabulary through the keyword method? 14 days later, the instructor administered a delayed test to the participants of both groups. Here, the researcher wanted to investigate how much vocabulary they could store in their long-term memory. The delayed-test was given in the same format again with 17 target words/phrases. However, there was a little change in the

order of the items. The instructor purposely changed the places of all the words so as to see if the students could still remember those even 14 days after the treatment.

Lastly, a week after the VKS delayed-test, picture description tasks were done with six aviation students from each group so as to measure the participants' vocabulary use in oral communication. They were provided with five pictures to describe in the way that they would use the taught vocabulary items from aviation context orally. The purpose was to have students produce more than a target word while describing one picture. Besides, the five pictures were chosen in a way that they would require the use of all the seventeen ESP vocabulary items.

3.4.1.1 Types of Sampling. In a quasi-experimental study, it is most of the time difficult to assign the participants to experimental and control group in a random way. For this purpose, nonrandom procedures were employed in order to select the participating students of this study. The researcher adopted nonprobability sampling in terms of acceptability in the present study. A total of 32 students who were selected from two intact English classes of Civil Aviation Cabin Services Program at a private university were involved in this quasi-experimental study. Currently, they have been studying English for specific purposes (Aviation English) at this private university in İstanbul, Turkey.

3.4.1.2 Data collection instruments. In an attempt to investigate the effectiveness of the keyword method on short-term and long-term recall in aviation English vocabulary learning and also to explore students' perceptions toward the keyword method, totally four data collection instruments were used. Data was collected quantitatively via vocabulary knowledge scale as pre-, post- and delayed-test and qualitatively via structured interviews and picture description task. Each instrument is described in detail below.

3.4.1.2.1 Vocabulary Knowledge Scale (pre-, post- and delayed). In the present study, vocabulary knowledge scale (VKS) (Paribakht & Wesche, 1993) was used as the instrument of the quantitative data three times in total and the differences in the pre-, post- and delayed-test scores belonging to the experimental and control groups were examined and analyzed. The pre-test and post-test were exactly the same and consisted of 17 target vocabulary items. They were prepared with random selection of the 17 vocabulary items which were supposed to be unfamiliar to the participants of both groups yet (see Appendix A). On this vocabulary knowledge scale there were numbers from 1 to 4 on the right side of each vocabulary item; and there was a *definition* and *sentence* part just below every

vocabulary item. Students were asked to mark only one number (1-4) depending on their vocabulary knowledge on the sheet. To elaborate, number 1 meant 'I've never seen this word/phrase before', number 2 meant 'I've seen this word/phrase before but I don't know what it means', number 3 referred to 'I know what this word/phrase means but I'm not sure how to use it', and lastly number 4 referred to 'I know what this word/phrase means and I can use it in a sentence'. Additionally, those students who marked number 3 had to write a definition to the related vocabulary item in Turkish or English. This was necessary to prove that they really knew what the word meant. Similarly, those who marked number 4 had to write a sentence in which they had to use that word along with its definition. So, they would evidence that they both knew the meaning of the word and had the capability to use it in a sentence. The delayed-test was prepared with the same vocabulary items and administered following the same format and procedure. The only difference of the delayed-test was that vocabulary items were given in a different / changed order in case students might have memorized word order in pre- and post-test.

Firstly, VKS was administered as a pre-test (see Appendix A) to find out whether students previously know the words on the scale or not; and to ensure that students are familiar with none of the 17 target vocabulary items. In the wake of the pre-test, the experimental group learnt these words/phrases through the keyword method. On the other hand, the control group did not receive any keyword instruction at all. They studied the words traditionally by definitions and synonyms. At the end of the 4-week period, both groups were applied the VKS as a post-test having the same words (see Appendix A). The post-test, which was linked to the short-term memory in the current study, was given to find out if there was any difference between the effect of the keyword method and the traditional vocabulary learning method. Finally, fourteen days after the post-test, the instructor gave the groups the VKS as a delayed-test this time (see Appendix B). The purpose of the delayed-test was to measure how many of the vocabulary items the group participants recalled in long term. In other words, we can say that the delayed-test was associated with the long-term memory. Thus, the researcher collected the quantitative data fully.

3.4.1.2.2 Structured interviews. In order to conduct the qualitative part of the present study, structures interviews were done with the treatment group. The treatment group participants, who consisted of a total of 16 students, were interviewed on the day when the post-test was given. They were expected to express their opinions about the utility of the keyword method for aviation English vocabulary learning. In these interviews, aviation

students were asked a few questions regarding the pros and cons and their perceptions of the keyword method (see Appendix C). Since they were not instructed under the keyword method, the control group was not involved in this qualitative part of the study. Namely the participants of the control group did not have any interviews with the researcher at all. Consequently, the researcher collected the qualitative data to conduct this quasi-experimental study.

3.4.1.2.3 Picture description tasks. In order to measure the participants' vocabulary use in oral communication, picture description tasks were done a week after the VKS as delayed-test was given. Six students from each group participated in these interviews. They were shown five pictures and asked to describe these pictures which would require them to use the taught vocabulary items from aviation context orally. It was aimed to have students produce more than a target word in the description of one picture. Also, the five pictures were selected in a way that they would call for the use of all the seventeen ESP vocabulary items.

3.4.2. Data analysis procedures. For the objectives of this study, so as to analyze the data of the first research question, the scores of vocabulary knowledge scales administered to both the experimental and control group as pre-, post and delayed test were analyzed quantitatively with the use of the Statistical Package of Social Sciences (SPSS) version 20.0. Through descriptive statistics, the mean scores obtained from vocabulary knowledge scales were compared between the two group participants. In addition, a Univariate ANOVA test was employed to make an analysis of group differences in the retrieval of ESP vocabulary items and to suggest the effectiveness of the keyword method on ESP vocabulary learning.

To analyze the differences between the two groups which were instructed through the keyword and the traditional methods, their vocabulary knowledge at receptive and productive levels were measured by the same data analysis procedures; namely, through SPSS and ANOVA test.

In order to answer the second research question regarding the experimental group students' perceptions of the utility and criticism of the keyword method in ESP vocabulary learning, the qualitative data gathered with structured interviews was analyzed. For this, content analysis was done. Firstly, the gathered data was transcribed by the researcher. Later, content analysis was done by the identification and categorization of qualitative data according to the theme and content. Lastly, quotations from the participants were used to

reveal positive or negative perceptions of the experimental group students on the keyword method in ESP vocabulary learning.

For the third research question, the data gathered through picture description task was analyzed through the VKS scoring formed by the researcher herself (see Appendix D). First, the recordings were transcribed and read meticulously by the two teachers, one of whom is native. Then, the teachers marked the students' sentences in accordance with whether they used the words or not and the appropriateness of the words in terms of grammatical accuracy and semantic. Finally, the percentages belonging to both groups were calculated. The vocabulary use of experimental and control groups in oral communication were compared in percentage.

3.4.3. Validity and reliability. Validity and reliability are of vital importance in both quantitative and qualitative research. Validity refers to how correct, meaningful, appropriate and useful the researcher's inferences are. Reliability is about consistency of scores. That is, it refers to how consistent scores or answers obtained from one administration of an instrument to another, and from a set of items to another item. When the case is the design or selection of the data collection tools which the researcher will use, validity and reliability are two important concepts to be considered (Fraenkel & Wallen, 2008). Considering validity and reliability, Guba (1981) highlights a number of criteria:

For one thing, internal validity (credibility) means ensuring that the researcher measures what she or he needs indeed. In the present study, internal validity is provided by means of triangulation of the data. Moreover, the researcher had a chance to collect the data from the participating groups in regular intervals just as she intended at the beginning owing to working in the department of that private university where the current study was conducted. Each individual in the treatment and control groups consented to participate in advance when the researcher asked for their consent.

Secondly, external validity (transferability), which refers to extending the study findings to other situations, was provided in this quasi-experimental study. Adequate information regarding the institution, setting, participants, etc. was presented in detail by the researcher.

Thirdly, referring to obtaining the same results every try within the same context by using the same methods and participants, dependability was provided by the presentation of the sufficient information on how the stages and process of the study developed. The

researcher illustrated data collection instruments, procedures and data analysis very clearly. Therefore, other researchers were given an opportunity to replicate this study.

The last criterion is confirmability which is known as reaching the study results without any prejudices and characteristics by the researcher. In this regard, the implementation of the quantitative and qualitative method together helped to decrease the researcher bias. The administration of both vocabulary knowledge scales and structured interviews made this study valid and reliable.

3.5 Limitations

In this study, there are some limitations to be considered. To begin with, the number of the students who participated in the study is limited to 16 for per class, a total of 32 students for both the experimental and control classes. Additionally, all the data was collected from the participants who have been studying in the sophomore year of Civil Aviation Cabin Services Program at the same university. In other words, conducting the study in only one institution is another limitation seen in the study. If this quasi-experimental study had been carried out in some other institutions too, it would be much more valid and reliable. Furthermore, the researcher could have had more participants' perspectives toward the efficiency or pros and cons of the keyword method.

This study investigates the effect of the keyword method on ESP vocabulary learning; therefore it can be said that the current study is focused on vocabulary competence of the learners. That means the researcher did not take the learners' competence in some other skills such as reading, listening and writing into account.

Chapter 4

Results

4.1 The Results of the Quantitative Data

The results of all the quantitative data obtained from Civil Aviation Cabin Services program students with B1 English level by means of a total of 96 vocabulary knowledge scales are presented in this section. The data here is regarding if there is a significant difference between the experimental and control groups in recalling vocabulary items.

To begin with, I would like to present some information about the participating population of the current study in the way that they are all divided into groups and times.

Table 1

The number of vocabulary knowledge scales administered to both groups

		Value Label	**N**
Group	1.00	experimental	48
	2.00	control	48
Time	1.00	pretest	32
	2.00	posttest	32
	3.00	delayed	32

As shown in Table 1, the total number of vocabulary knowledge scales given in this study was 96. A total of 48 of these scales were given to the experimental group while the rest 48 were administered to the control group. When we look at how many times VKS was given, we see that VKS was applied three times for both groups as pretest, posttest and delayed. In this situation, if we consider the same number of participants (32) for each of pretest, posttest and delayed, there appears 16 experimental group participants and 16 control group participants.

Table 2

Pairwise comparisons of the experimental and control groups

Time	(I) Group	(J) Group	Mean Difference (I-J)	Std. Error	**Sig.**[b]	95% Confidence Interval for Difference[b]	
						Lower Bound	Upper Bound
1.00 pretest	1.00 experimental	2.00 control	.688	1.538	**.656**	-2.369	3.744
	2.00 control	1.00 experimental	-.688	1.538	**.656**	-3.744	2.369
2.00 posttest	1.00 experimental	2.00 control	23.875*	1.538	**.000**	20.819	26.931
	2.00 control	1.00 experimental	-23.875*	1.538	**.000**	-26.931	-20.819
3.00 delayed	1.00 experimental	2.00 control	24.000*	1.538	**.000**	20.944	27.056
	2.00 control	1.00 experimental	-24.000*	1.538	**.000**	-27.056	-20.944

The results of VKS in table 2 show that there was no significant difference between vocabulary knowledge of the experimental and control groups before the keyword method treatment (.656). However, it is seen that after the treatment there was a significant difference between both groups (.000) when we consider their recall of the target vocabulary words/phrases in short-term memory (2.00 posttest). Furthermore, this significant difference (.000) between the two groups keeps its validity in long-term (3.00 delayed test).

4.1.1 Comparative Effectiveness Between Experimental and Control Groups Before and After Treatment. As for the first research question on whether there will be a significant difference between the students who are instructed in the keyword method and the traditional method in the retrieval of ESP vocabulary knowledge, the mean scores of pre-, post and delayed applications of vocabulary knowledge scales to the two groups (experimental vs. control) were analyzed in a comparative way. The following table presents the descriptive statistics:

Table 3

Descriptive statistics of VKS as pretest, posttest and delayed for the experimental and control groups

Group	Time	Mean	Std. Deviation	N
1.00 experimental	1.00 pretest	24.69	3.11	16
	2.00 posttest	53.75	4.02	16
	3.00 delayed	52.75	5.49	16
	Total	43.73	14.25	48
2.00 control	1.00 pretest	24	3.61	16
	2.00 posttest	29.87	5.11	16
	3.00 delayed	28.75	4.28	16
	Total	27.54	5	48
Total	1.00 pretest	24.34	3.34	32
	2.00 posttest	41.81	12.94	32
	3.00 delayed	40.75	13.12	32
	Total	35.63	13.38	96

As can be clearly seen in table 3 above, both experimental and control group performed better in the VKS posttest than they did in the VKS pretest. Also, it was revealed that the keyword group showed an increase of 29.06 ESP words and the control group had an increase of 5.87 ESP words when both groups' performances on the pretest and posttest were considered. Finally, the findings from pretest and posttest prove that the keyword treatment had much more beneficial effects than the traditional method.

Table 3 demonstrates that the students in the experimental group gained more knowledge of the target ESP vocabulary items after the two groups were instructed in the keyword treatment and the traditional treatment. The mean scores indicated that although the experimental (M=24.69) and control group (M=24) students were homogeneous in the VKS results of pretest, the students of the experimental group performed significant ESP vocabulary gains on the posttest by recalling much more target vocabulary items (M=53.75). The control group also gained ESP vocabulary knowledge on the posttest but in small measure (M=29.87).

Moreover, when we think that VKS given as delayed test is associated with long-term memory, it is obviously seen in table 3 that the experimental group students were able to recall more target vocabulary in long-term than the control group students could. There was little decrease in both groups' performance of recalling the target vocabulary on the delayed test. Nevertheless, the delayed test results reveal that the students who were instructed through the keyword method could retain considerably more lexical knowledge (M=52.75) than those who were instructed through the traditional method (M=28.75).

In order to support all the findings of the first research question above regarding if there is a significant difference between the students who are instructed in the keyword method and the traditional method in the retrieval of ESP vocabulary knowledge, the scores from the application times of VKS (pre-, post and delayed tests) were analyzed by the use of ANOVA. The following figure presents the results:

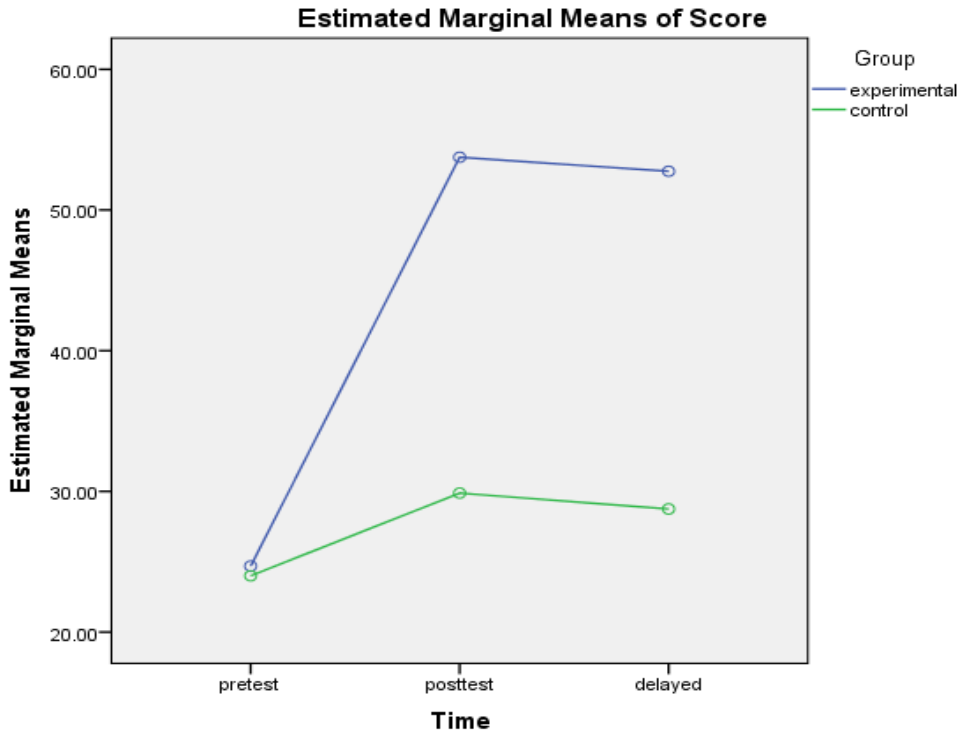

Figure 1 Estimated marginal means of score

In Figure 1 the researcher linked posttest with short-term memory while delayed test was associated with long-term memory. As can be seen in Figure 1, ESP vocabulary knowledge levels of the experimental and control group were almost the same before the

treatment (pretest). Both groups made improvement in the retrieval of the target ESP lexical items after the treatment (posttest). Also, both experimental and control group could recall a bit less number of ESP vocabulary items in long term (delayed). In other words, there was a slight decrease in their ESP vocabulary knowledge on the delayed-test when compared to their performance on the posttest. However, when both groups' means are taken into account, Figure 1 shows that the keyword group students recalled significantly more target vocabulary items than the traditional group both in short- and long-term memory. Therefore, the keyword method instruction proved its effectiveness and utility in ESP vocabulary learning.

4.1.1.1 The differences between the groups at receptive level. The differences between the experimental and control groups at receptive level are specifically mentioned in this section. The data regarding receptive vocabulary knowledge of both groups was collected by the scoring of number 3 for each ESP target word on the administered VKS, which is 'I know what this word/phrase means but I'm not sure how to use it'. Students' responses on VKS given as the post and delayed test were scored and summed up and how the experimental and control group differed from each other at receptive level was illustrated in the table below.

Table 4

Descriptive statistics of the differences between the experimental and control group at receptive level

Group	Time	Mean	Std. Deviation	N
1.00 experimental	1.00 posttest	30	13.01	16
	2.00 delayed	25.12	12.91	16
	Total	27.56	12.99	32
2.00 control	1.00 posttest	3	4.10	16
	2.00 delayed	3	3.29	16
	Total	3	3.65	32
Total	1.00 posttest	16.50	16.68	32
	2.00 delayed	14.06	14.57	32
	Total	15.28	15.58	64

Table 4 illustrates that the keyword method instruction had much better results in ESP vocabulary learning than the traditional vocabulary learning method at receptive level too.

The experimental group's mean score of receptive vocabulary knowledge (M=30) outnumbered the control group's (M=3) on the posttest. Similarly, the mean scores belonging to both groups on the delayed VKS demonstrated that the traditional method could not become as effective as the keyword instruction. When the retrieval of the target words/phrases in long-term memory was considered, the control group gained a mean score of 3 whereas the experimental group proved to be more efficient with a mean score of 25.12.

Furthermore, when we look at Table 5 below, we already see that there was a significant difference (.000) between the experimental and control group at receptive vocabulary knowledge level.

Table 5

Pairwise comparisons of the experimental and control group at receptive level

(I) Group	(J) Group	Mean Difference (I-J)	Std. Error	**Sig.**[b]	95% Confidence Interval for Difference[b]	
					Lower Bound	Upper Bound
1.00 experimental	2.00 control	24.563*	2.383	**.000**	19.796	29.329
2.00 control	1.00 experimental	-24.563*	2.383	**.000**	-29.329	-19.796

4.1.1.2 The differences between the groups at productive level. The comparison between the experimental and control groups at productive level is particularly made in this section. The data regarding productive vocabulary knowledge of both groups was obtained by the scoring of number 4 for each ESP target vocabulary item on the administered VKS, which is 'I know what this word/phrase means and I can use it in a sentence'. Students' responses on VKS given as the post and delayed test were scored and summed up and how the experimental and control group differed from one another at productive level was presented in the table below.

Table 6

Descriptive statistics of the differences between the experimental and control group at productive level

Group	Time	Mean	Std. Deviation	N
1.00 experimental	1.00 posttest	22.25	15.92	16
	2.00 delayed	20.50	17.65	16
	Total	21.37	16.56	32
2.00 control	1.00 posttest	2.50	5.03	16
	2.00 delayed	1.75	3.85	16
	Total	2.12	4.43	32
Total	1.00 posttest	11.50	14.78	32
	2.00 delayed	12	16.32	32
	Total	11.75	15.45	64

Table 6 suggests that the keyword group participants were much more productive in ESP vocabulary learning than the traditional group participants. The experimental group's mean score of productive vocabulary knowledge (M=22.25) outnumbered the control group's (M=2.50) on the posttest. Likewise, both groups' mean scores on the delayed VKS revealed that the experimental group students performed better than the control group students in the production of ESP vocabulary knowledge. They had a mean score of 20.50 by stating that they know what the target word/phrase means and they can use it in a sentence and also by writing a definition of the word along with a sample sentence while the students in the control group received a total of 1.75. To sum up, the keyword method proved to be more beneficial and successful at both receptive and productive level of ESP vocabulary learning not only in short-term but also in long-term memory.

4.2 The results of the qualitative data

This section presents the results of all the qualitative data collected from B1 level Civil Aviation cabin services program students via structured interviews. All the data is regarding aviation students' perceptions about effectiveness of the keyword method along with its

deficiencies. The qualitative data of this study was gathered through structured interviews which were done with the experimental group participants. Content analysis was done in order to group the obtained findings. Finally, by means of the content analysis of the qualitative data, students' perceptions in the treatment group of the efficiency and drawbacks of the keyword method in ESP vocabulary learning are reported. The following questions were asked to the treatment group participants in the structured interviews:

1. Did you find the keyword method effective in ESP vocabulary learning? If yes, why?

2. If you were asked to criticize the method, what criticism would you make?

3. What are your general opinions about the keyword method?

4.2.1 The perceptions of the students about the effectiveness of the keyword method. In this section, the perceptions of the students about the effectiveness of the keyword were included with their general opinions about the keyword method which is the answer to the third interview question. In other words, the structured interview questions 1 and 3 were analyzed together due to the content similarities of the students' answers.

To begin with, when the students were asked about their perceptions about the effectiveness of the keyword method, five of them explicitly stated that they found the method effective and catchy because they could remember the words more easily thanks to a keyword in their native language. Some excerpts are as follows:

> […] I found the method useful. I think it is effective more because vocabulary learning with the use of Turkish words makes it catchy. Thanks to associations, the target words can be stored in the memory for long-term (Student 3, Structured interview).

> […] It is a really good method in that it is catchy. Now I know it exactly as soon as I come across the word. If there was no keyword associated to the target word, it would be much harder to learn it (Student 1, Structured interview).

> […] The keyword method is really beneficial because it is always catchier to learn something by coding. In this method, a keyword from Turkish language

enables me to develop my vocabulary knowledge through coding (Student 7, Structured interview).

[…] Both written and visual reminders contribute to recalling … For instance, even when you do not remember what the related picture describe, there is another option to recall the target word. The key word … You may find the meaning of the target word based upon the key word (Student 9, Structured interview).

[…] In general I found the method effective. This is because we coded the target words by the use of key words. Furthermore, some key words that were formed were extremely similar to Turkish words, just as they are pronounced in Turkish language. Thus, it was simpler and catchier to memorize them (Student 13, Structured interview).

In addition, the treatment group students said that they found the keyword method efficient because it presents visuals too. Considering this point, five of the students stated:

[…] People should first know their own memory well, the way they learn best… Some of us have a better visual memory while some learn better via audial elements. I absolutely believe that I have got a strong visual memory (Student 5, Structured interview).

[…] Why I found the keyword method beneficial is that it consists of visual elements like images; and to me, I have really got a strong visual memory. When I study this way, I feel like improving my vocabulary knowledge better (Student 2, Structured interview).

[…] Yes the method is effective … In my opinion, besides acoustic associations, visualizing makes a good contribution to its effectiveness (Student 6, Structured interview).

[…] The method also provides brief stories that can remind learners of the word meaning. However, if your visual memory is strong enough, you can have a chance to remember the word again by means of images even if you may forget about what the story tells (Student 8, Structured interview).

> […] By means of this method, I can imagine the meanings of lots of newly-learnt words in my mind. This happens in my memory visually (Student 14, Structured interview).

Moreover, the students' responses showed that the students were positive about the method on the ground that it is effective and catchy owing to being silly and fun as illustrated in the following samples:

> […] The silliness of the visuals and the selected key words is interesting and thus successful. In general, people already tend to remember silly things more easily and find them intriguing (Student 10, Structured interview).

> […] I think the keyword technique makes vocabulary learning fun and easy. This is interestingly accomplished by silly stories (Student 8, Structured interview).

As a conclusion, Civil Aviation Cabin Services students found the keyword method effective and beneficial in ESP vocabulary learning and retention in that it provides them a keyword in their native language, meaningful visuals and silly and fun stories.

4.2.2 The perceptions of the students about the deficiencies of the keyword method. The students were also asked about their perceptions about the deficiencies of the keyword method with the second research question in the present study. Some students made constructive criticism on the formation of stories and pictures as follows:

> […] I think the pictures and photos could have been chosen more intriguing. In addition, the length of the stories must be taken into consideration. Sometimes they may be too long and boring. They may be formed not to bother learners. However, the presentation of all the stories within aviation context was a bonus in term of persistency (Student 15, Structured interview).

> […] Well, the sentences within the stories could have been more meaningful. Also, it would be more beneficial to make stories a bit shorter (Student 4, Structured interview).

> […] While making up the stories, more details can be considered in terms of content. Therefore, much more creative and efficient short stories could be created (Student 12, Structured interview).

Besides, ESP students made a few negative comments on the selection of the Turkish key words for each English target word although they usually agree that ESP vocabulary learning through the keyword method has more pros rather than cons.

> […] If I criticized this method, I would say that some of the key words selected from Turkish language are too difficult to keep in mind (Student 13, Structured interview).

> […] One thing I would criticize is that some key words do not remind me of original target words. I have difficulty in establishing between the two; but this happens only with some words (Student 8, Structured interview).

> […] The thing is that when I first look at the target word, the associated key word provided by the researcher may not come to my mind. Interestingly, I tend to associate it to a different key word which makes sense to me at the moment of practicing (Student 6, Structured interview).

Furthermore, two of the students stated that they were not happy with their productive performance even if they were able to recall the word meanings quickly with the help of key words, visuals and a set of funny stories.

> […] Many learners may not use these words in a sentence even though they have already got the meaning receptively. I think we need more practice in sentence formation by using the newly-learnt words (Student 9, Structured interview).

> […] For me, it is really easy to learn vocabulary items and recall their Turkish meanings later on. On the other hand, I may not know exactly where to put the word when I intend to make a sample sentence (Student 1, Structured interview).

Three aviation students expressed their dissatisfaction with the utility of the keyword method in ESP vocabulary learning roughly. They stated that stories are not sometimes creative enough; the method does not work well with each specific target word and that the associated key word may even be confusing rather than effective and mnemonic. As shown below:

> […] To be honest, although the keyword method seems effective to some extent, the content of stories does not remain in my mind. The stories must be improved and made more creative (Student 12, Structured interview).
>
> […] Though I can say that the method is useful broadly, I disagree that it works effectively in some words. Maybe this results from the differences between the researcher's perceptions and mine. Also, it may be in parallel with accumulation of knowledge (Student 15, Structured interview).
>
> […] I do not find this method effective. What comes to my mind later is not the meaning of the word but the key word only. I mean, I remember the key word as if it was the original word meaning. This situation may mislead me (Student 4, Structured interview).

It can be concluded that some of Civil Aviation Cabin Services students were not always happy with the use the keyword method in ESP vocabulary learning and retention since they feel some deficiencies regarding the method on the formation of stories and pictures and the selection of the Turkish keywords.

4.3 The Results of Vocabulary Use in Oral Communication

The results of the 12 Civil Aviation Cabin Services students' vocabulary use in oral communication are provided in this section. The findings obtained from both control and experimental group via picture description are presented descriptively and comparatively along with the calculated percentages.

To begin with how productive the six control group participants could be in the ESP vocabulary use in oral communication, it became evident that two of them did not generate any of the 17 target words or phrases at all. There was a tendency to describe the pictures which they saw through general English knowledge, not through Aviation English. They made lots of grammar mistakes even with the sentences formed in Basic English and including none of the target ESP words/phrases.

When it comes to those control group participants who used the expected target words and phrases while describing the pictures, they partially produced some of them. However, the problem with their sentences was that they were semantically right but grammatically wrong. To exemplify: *They are <u>brace position</u>* (Participant 9, semi-structured interview). In

this sentence, the participant 9 seemed to have chosen the appropriate word matching with the context in the picture. Nevertheless, it might cause misunderstanding in oral communication because it was used without the preposition 'in'. 'Brace position' here is normally supposed to refer to the passengers' position where they are at the moment. Since it was used grammatically wrong, it might be understood as if it referred to the subjects (passengers) themselves, which would be meaningless. In other words, wrong grammar use of the word directly and negatively affected the semantic meaning of it. Another example belongs to the participant 11: *The passengers doing brace position* (Participant 11, semi-structured interview). As seen in this sentence, the participant could select the suitable word and obviously express the existing situation in the related picture semantically correct. However, it is clearly seen that the sentence was formed structurally lacking and wrong.

Only two of the students in the control group were competent on the use of a few target words and phrases in oral communication. For instance: *There is a baby in a baby bassinet ... they're taking the baggage to the hold ...* (Participant 7, semi-structured interview). *The baby is in a baby bassinet ... The passengers are in brace position* (Participant 12, semi-structured interview). When we look at the sentences above, we clearly see that the participants 7 and 12 were able to recall some target ESP words and produce them in both semantically and grammatically correct way in oral communication.

On the other hand, the six experimental group participants of the semi-structured interviews turned out to be much more successful and productive than the control group participants in ESP vocabulary use in oral communication. In addition, the experimental group participants proved to recall much more target words and phrases in number. The target words/phrases produced by the participants are underlined in the quotes below.

> […] She is wearing life vest, pulling on the red cord ... maybe the passenger doesn't have exact change .. They are serving drinks and refreshments … He is lifting the buckle … He needs a bassinet for his son … They are in the brace position (Participant 1, Semi-structured interview).

> […] I see a flight attendant, she inflates vest … They are doing snack service and refreshment service … The pay is exact change … In picture A she is lifting the buckle … The baby is sleeping in bassinet … I see a flight attendant, she greets people … The people are in brace position. (Participant 5, Semi-structured interview).

In the examples by the participants 1 and 5 above, it is seen obviously that the target words and phrases were used with both semantic appropriateness and grammatical accuracy. However, four of the participants could not make grammatically full sentences in all cases even though they used a wide range of the target ESP words and phrases.

> […] I see a flight attendant wear life vest and pull on the red cord... flight attendants' refreshments service … maybe she needs exact change … passenger is lift the buckle … I see a baby in the bassinet … hold area … flight attendant hand out for the passengers … flight attendants greet passengers … the passengers brace position … the seats position upright position …. (Participant 2, Semi-structured interview).

As seen in the quote above by the participant 2, he or she produced a plenty of the target ESP words and phrases, some of which were formed with both semantic appropriateness and grammatical accuracy whereas some vocabulary items were only articulated within a suitable context but without grammatical accuracy.

Similar findings were obtained from the vocabulary use of the experimental group participants 3 and 4 in oral communication.

> […] She inflating the life vest with red cord … On the trolley there are drink, snack and refreshments… There is a passenger in the cabin, he is lift the buckle before flight … In Picture B there is a bassinet in the cabin … the passenger is going to the hold … she is handing out immigration forms to passengers … There are two flight attendants and they greet the passengers… The passengers are do brace position … All the seats are upright position… (Participant 3, Semi-structured interview).

> […] There is a woman, she inflating vest … there is no reclining seat in the cabin ... He refreshment service I think … A baby on the bassinet … there is a woman and she is lifting her buckle .. All baggage on the cargo hold … she is handing out some forms … they are greeting the passengers … these passengers are in brace position … (Participant 4, Semi-structured interview).

Although the participants 3 and 4 were able to produce a variety of the expected words and phrases such as inflate vest, lift the buckle, brace position, hand out, reclining seat,

upright, greet, refreshments and bassinet, we can observe that there partially appeared grammatical deficiencies in their sentences (e.g. He refreshment service I think).

Nevertheless, the experimental group participants ended up using much more in number among a total of 17 words and phrases than the control group participants. Furthermore, it was proved by the scoring of vocabulary use in oral communication that the experimental group performed much better when they are compared with the control group. The obtained percentages summarize briefly that the experimental group participants had vocabulary use in oral communication by 61% while the control group participants could use vocabulary items by 36% only. As a result, the keyword method proved to affect EFL learners' ESP vocabulary use in oral communication more positively than the traditional method did.

Chapter 5

Discussion and Conclusion

5.1 Discussion of Findings for Research Questions

The current research study attempted to investigate the impact of the keyword method on ESP vocabulary learning of EFL students both in short and long-term. It also aimed to examine the effect of the keyword method on students' receptive and productive vocabulary knowledge in ESP vocabulary learning and recall. In addition, the experimental group students' opinions about the keyword method to enhance their performance in ESP vocabulary learning were explored. Data was quantitatively collected by means of vocabulary knowledge scales administered at three different times as pre-, post- and delayed. Qualitatively, data was gathered from structured interviews done with the experimental group participants who received the target vocabulary items through the keyword method instruction. Also, a picture description task was administered to six students from both groups. In the following section, the findings with regard to every research question are discussed addressing the literature review of the present study. Also, pedagogical implications on a basis of the findings will be presented and suggestions for further research studies will be provided. .

5.1.1 Discussion of findings of RQ 1: Will there be a significant difference between the students who are instructed in the keyword method and those who are instructed through the traditional method in the retrieval of ESP vocabulary knowledge? The first research question attempted to investigate the effect of the keyword method on ESP vocabulary learning and recall in short and long-term memory both receptively and productively by comparing it with the traditional vocabulary learning method. The mean scores which were obtained from vocabulary knowledge scales administered to the experimental and control groups as the pre-, post and delayed test were compared. On the ground that Nation and Meara (2010) argue that English vocabulary consists of three main aspects related to form, meaning and use; on the VKS, the participants of this study were asked to write an example sentence by using the target word/phrase as well as a Turkish or English definition. Therefore, whether the participants really knew the form, meaning and use of the related word or phrase grammatically and semantically or not was measured in the current study. According to the findings, it was revealed that the keyword method had a considerable influence on the treatment group students' receptive and productive performance

in ESP vocabulary learning. The results of the ANOVA test demonstrated that there was a significant difference between vocabulary knowledge scales' scores of the students who received the keyword treatment while learning the target ESP lexical items and those of the students who were taught through the traditional vocabulary learning method via synonyms, definitions, so on. This significant difference also suggested that vocabulary learning and recall could be associated to the utility and efficiency of the keyword method. The application of the keyword method instruction was based upon associating the target vocabulary item with a reminder keyword from learners' first language. It additionally provided visual images accompanied by strange stories including the keyword and the meaning of the word/phrase for learners. In this study, the participating students of the keyword group were able to learn and remember the target ESP words and phrases later not only through the use of a keyword from L1 within a story which could be silly, strange or fun for the students but also through the use of visuals describing the context of stories.

The results of this quasi-experimental study are compatible with Schmitt's argument (2008) that a faster and better way of vocabulary retention and gaining productivity is created by intentional vocabulary learning. The results also supported that intentional vocabulary learning is more efficient than incidental learning in that it strengthens the process of learners' lexical development with the help of repetitions and memorization strategies, which is a situation learners can achieve by themselves in a short period of time (Elgort, 2011). Considering that the keyword method treatment in the present study took short time and that students can make up their stories on their own by linking the target word to a keyword in their native language when they wish, Elgort's opinion (2011) was proved right to a large extent.

Moreover, the findings also supported the results of the study conducted by Baleghizadeh and Ashoori (2010) who aimed to investigate the effect of the keyword and word list methods on immediate retention of English vocabulary in a classroom setting. The significant difference in the scores of the participants in the experimental and control group of this study confirmed that the keyword method application on teaching target vocabulary items to EFL learners is a more beneficial and effective way of enhancing learners' vocabulary knowledge and size when compared to the application of wordlist method. It was released that the experimental group which were taught through the keyword method instruction significantly outperformed the control group which was presented the target vocabulary words with the use of wordlist method.

Likewise, the findings of the current study are in parallel with the results of an experimental study carried out by Sagarra and Alba (2006). They attempted to investigate to what extent three vocabulary learning methods; rote memorization, semantic mapping and the keyword method are effective in learners of Spanish. The three group participants were all administered a total of six posttests having 8 items each; three for the immediate posttest and three for the 3-week delayed test. As a consequence of the mean scores obtained from these post and delayed tests, the keyword method had better long-term effects when it was compared with rote memorization and semantic mapping. This meant that deeper processing required by the keyword method via form and meaning associations produced the most retention in vocabulary learning. Besides, the impact of the keyword method instruction on students' vocabulary learning and recall is in line with another study conducted by Piribabadi and Rahmany (2014) who examined how effective the keyword method and word-list method instruction were on ESP vocabulary learning by making comparisons between the two methods. According to the results from a two-way ANOVA, engineering students in the keyword group gained highly better scores than those who were provided with the word-list method instruction. Consequently, the findings of the current study were in the same direction with the study by Piribabadi and Rahmany (2014) too.

It was argued by Atay and Ozbulgan (2007) that teaching and learning vocabulary explicitly can help students become more successful in the vocabulary learning process. The keyword method, an intentional vocabulary learning strategy, was proved right when the results of the current study showing the efficiency of the keyword method in gaining ESP lexical items were taken as a basis. The findings in this study are also in accordance with those of the study carried out by Chen and Hsiao (2010) who made an investigation so as to examine the efficiency of the keyword method training in ESP vocabulary instruction. They randomly assigned forty students at a technology university as the keyword and control groups. In the wake of the treatment, students' posttest scores revealed that there was a considerable increase of 8.6 words in the keyword group students' pretest and posttest performances while the students had an increase of 4.15 words in the traditional method instruction. This suggested that the keyword treatment yielded more beneficial effects than the traditional treatment. The keyword group ended up gaining and recalling more target words immediately after the instructional treatment.

5.1.1.1 Discussion of findings of RQ 1.1: Will there be a significant difference in the experimental and control group students' retrieval of ESP vocabulary knowledge at receptive and productive levels? Vocabulary was defined by Neuman and Dwyer (2009) in two categories as receptive and expressive, namely productive. They also underlined that effective communication calls for acquiring both. Alqahtani (2015) stated that vocabulary mastery is necessary so as to express ideas and understand what other people say. In an attempt to refer to this research question that aimed to investigate whether there was a significant difference between the experimental and control groups' performance in the retrieval of ESP vocabulary knowledge at receptive and productive levels or not, the data was gathered through the same VKS which was employed in data collection process of the first research question. However, among four numbers, each of which measured if the learner knew the target word or phrase, only responses given to numbers 3 and 4 on the VKS were examined and analyzed. For receptive performance, students' scores of number 3 for each target vocabulary item were summed up. On the other hand, the scores of number 4 were calculated again for each vocabulary item and summed up in order to measure students' productive vocabulary knowledge level. In other words, number 4 on the VKS students measured if students could grammatically and semantically use the target word/phrase in a sentence.

The findings indicated that the number of ESP words/phrases experimental group knew on the VKS (posttest) was much more than those of the control group. This means that the students in the experimental group outperformed the control group students in terms of receptive vocabulary knowledge. The findings also revealed that the keyword method became more effective at productive level of ESP vocabulary learning too. The experimental group's mean score of productive vocabulary knowledge (M=22.25) outnumbered the control group's (M=2.50) on the posttest. As for the delayed test, the results showed similarities in that the experimental group had a mean score of 20.50 while the mean score of the control group was 1.75. Shortly, the keyword method turned out to be more beneficial and effective at receptive and productive level of ESP vocabulary learning in both short-term and long term memory.

5.1.2 Discussion of findings of RQ 2: What are the EFL learners' perceptions toward learning ESP vocabulary through the keyword method? In order to address the second research question of the current study that investigated the perceptions of EFL learners about using the keyword method in learning ESP vocabulary items, data was qualitatively collected from structured interviews. These structured interviews were done only with the

participating students of the experimental group so as to get an idea of their positive and negative views about the effectiveness of the keyword method in ESP vocabulary learning. In the interviews, the experimental group students were asked a total of three questions regarding a) whether the keyword method is effective or not and if it is effective why, b) what they would criticize about the keyword method, c) what they think about the keyword method in general. The findings of the structured interviews revealed that the students who learnt the target ESP words and phrases through the key words as well as stories and visuals found this keyword method instruction effective and useful in order to study, practice and recall the target vocabulary items in short and long term.

The participants in the experimental group claimed that using the keyword method instruction made the vocabulary learning process more permanent particularly and they could store the lexical items in the memory for long-term because they associated the target words with a key word in their first language. Also, they asserted that the taught-words and phrases were much catchier for them thanks to the help of the Turkish words. Moreover, the treatment group students stated that recalling the target ESP vocabulary items got easier by means of both written and visual reminders. Some of them claimed they felt like improving their vocabulary knowledge better when their visual memories worked. In brief, using a Turkish key word associated to the target word and using images make a vast quantity of contributions to learning and recalling process of ESP vocabulary knowledge.

In addition, the findings are in line with the argument by Murcia (1997) that the keyword technique is an appropriate strategy among mnemonic devices in that the learners are reminded of the keyword from L1 as they see or hear the target word. As referred in the review of the literature, Atkinson and Raugh (1975) suggested that the keyword method develops not only immediate but also delayed retention of vocabulary in second language learning. In the present study, the results demonstrated that the experimental group students who received the target ESP vocabulary items through the keyword method instruction performed better in both immediate and delayed recall of the words and phrases when compared to the control group students who learnt the lexical items through the traditional method. Besides, the findings are also in concordance with an experimental study which was conducted by Baleghizadeh and Ashoori (2010) on the purpose of comparing the effect of the keyword and wordlist methods on immediate retention of English vocabulary knowledge. The experimental group was given vocabulary instruction with the use of the keyword method whereas the control group got vocabulary instruction through the wordlist method. In

accordance with the results of this study, the mean score belonging to the experimental group was higher than the mean score of the control group.

5.1.3 Discussion of findings of RQ 3: Is there a difference between the experimental and control groups in the use of target vocabulary in oral communication?

The third research question attempted to investigate how the keyword method affects the EFL learners' ESP vocabulary use in oral communication. In order to measure both groups' participants' vocabulary use in oral communication, picture description tasks were done a week after the VKS as delayed-test was given. Six students from each group participated in these tasks. They were demonstrated five pictures and asked to describe these pictures which would necessitate them to use the previously taught ESP vocabulary items orally. The aim of the researcher was to have students produce more than a target word in the description of one picture. Also, the five pictures were chosen in a way that they would call for the use of all the seventeen ESP target words and phrases.

Overall, the findings of the present quasi-experimental study are greatly in the direction of the literature review. Furthermore, they are supported by a number of various studies carried out to examine the effect of the keyword method on vocabulary learning and recall.

5.2 Pedagogical Implications of the Study

The present study has pedagogical implications in order to activate the use of the keyword method in ESP vocabulary learning and recall as an efficient and useful tool in ESP classroom settings. As mentioned before in the introduction chapter, I needed to conduct this study since I observed that Civil Aviation Cabin Services students' vocabulary knowledge and use was quite limited and most of the time insufficient. On the basis of all the obtained findings, it can be proposed that the keyword method and its properties such as meaningful visual images, strange stories presented within the ESP context and a key word from L1 to be acoustically associated with the target word or phrase in L2 should be used to teach lexical items in an ESP classroom by the teachers and instructors. Furthermore, ESP teachers should benefit from this method so as to make students have more practice on the target vocabulary by means of reminding them of the associated key word from the students' native language. It should be noted that this method was of a strong effect on vocabulary learning and retention in the aviation field of ESP. Therefore, it is of importance that curriculum developers and material designers take the functions and properties of the keyword method into account while

course materials are designed and developed. Thanks to the keyword method, learners have the potential to acquire a larger number of ESP words and phrases in a short period of time because the method supports intentional vocabulary learning with the aid of meaningful visuals, deliberate word links between L1 and L2 and stories related to the pictures.

The keyword method use in ESP vocabulary learning allows students to recall the vocabulary items which were previously taught to them in both short-term and long-term memory. Thus, EFL and ESP teachers can provide assistance for the effective strategy development of their students. It is necessary that the students be informed about the increasing effects of intentional vocabulary learning strategies. Also, the instructors should give them motivation to study and practice vocabulary items intentionally by themselves out of the class. Hence forth, the teachers and instructors need to explicitly teach strategies calling for using visual images, verbal or acoustic links and some other reminders efficiently and actively in their ESP classes.

Additionally, the findings indicate that applying the keyword method to the language classes in the field of ESP promotes autonomy in vocabulary learning and offers students the opportunity to work out and have practice on the target vocabulary independently. As a consequence, it is important that intentional vocabulary learning strategies should be incorporated in ESP classrooms so as to help learners in the immediate and delayed retrieval of the target words and phrases. Shortly, learning the target vocabulary knowledge intentionally by favor of the use of intentional vocabulary learning strategies like the keyword method must be taken as an advantage in an effort to motivate students to study and improve their vocabulary knowledge development not only in-class but also out of the language classes.

5.3. Conclusions

The purpose of this study was to investigate the impact of the keyword method, an intentional vocabulary learning strategy, on ESP vocabulary learning in both immediate and delayed retention. The study also aimed to find out Civil Aviation Cabin Services students' vocabulary knowledge performance at receptive and productive levels after the application of the keyword and traditional methods respectively to the experimental and control groups. Furthermore, it attempted to reveal the perceptions of the treatment group students toward the advantages and disadvantages of the use of the keyword method with the intent of ESP vocabulary learning. The experimental group students' outperformance rather than the control

group students demonstrates that learning vocabulary items intentionally with a strategy like the keyword method becomes more effective and permanent in second language vocabulary acquisition. Most of the students in the experimental group stated that they really enjoyed learning the target words and phrases in Aviation English through the visual images, verbal or acoustic links between first and second language and brief stories of the keyword method. In addition, they found the application of keyword method in ESP vocabulary learning really effective and useful particularly in storing and recalling a greater number of vocabulary items in memory.

Finally, the results of the present study confirmed that intentional vocabulary learning through the use of the keyword method motivated and enabled ESP learners to gain better learning and retention both receptively and productively in short and long-term.

5.4 Recommendations for Future Research

Even though the results of this quasi-experimental study revealed that the keyword method is an effective way to teach the target vocabulary items in the field ESP, the research study still has some weaknesses. To begin with, the limited number of students participating in the present study may not be adequate in order to come to valid and reliable conclusions about the efficiency of the keyword method for ESP vocabulary learning in EFL classroom settings. Apart from this, the study was carried out at an only one university which was a foundation university in İstanbul. It seems to be a barrier to generalize our overall findings obtained at the end of the study. For this reason, it is crucial to conduct further studies with participants enrolled in a variety of state and private universities.

In addition, similar studies can be carried out at different settings to investigate the impact of the keyword method on ESP vocabulary learning at receptive and productive levels. Besides, to draw more valid conclusions about the perceptions of the students who are studying in Civil Aviation Cabin Services Program of both state and private universities and who are instructed through keyword method toward this method has importance for future research.

In spite of the limitations above, the present study is still expected to make contributions to the increasing research studies in the future on the effectiveness of applying the keyword method instruction in ESP vocabulary learning.

REFERENCES

Akbari, Z. (2011). Vocabulary comprehension and learning in an ESP context: Strategy use and knowledge sources. *Asian ESP Journal, 7* (2), 5-27.

Alqahtani, M. (2015). The importance of vocabulary in language learning and how to be taught. *International Journal of Teaching and Education, 3* (3), 22.

Amiryousefi, M. & Ketabi S. (2011). Mnemonic Instruction: A way to boost vocabulary learning and recall. *Journal of Language Teaching and Research, 2* (1), 178-182.

Ashoori Tootkaboni, A. (2012). Recall of foreign-language vocabulary: Effects of keyword, context and wordlist instructional strategies on long-term vocabulary recall of EFL learners. *Journal of Theory and Practice in Education, 8* (1), 54-71.

Astika G. G. (1993). Analytical assessment of foreign students' writing. *RELC Journal*, 24, 61-72.

Atay, D. & Ozbulgan, C. (2007). Memory strategy instruction, contextual learning and ESP vocabulary recall. *Journal of English for Specific Purposes*, 26.

Atkinson, R. C. & Raugh, M. R. (1975). An application of the mnemonic keyword method to the acquisition of Russian vocabulary. *Journal of Experimental Psychology: Human Learning and Memory*, *104* (2), 126-133.

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

Barcroft, J. (2009). Strategies and performance in intentional L2 vocabulary learning. *Language awareness*, *18*(1), 74-89.

Campos, A., Rodríguez-Pinal, M. D., & Pérez-Fabello, M. J. (2014). Receptive and productive recall with the keyword mnemonics in bilingual students. *Current Psychology*, *33*(1), 64-72.

Cahyono, B.Y. & Widiati, U. (2008). The teaching of EFL vocabulary in the Indonesian context: The state of the art. *TEFLIN Journal. 19* (1).

Chen, I.J. & Hsiao, H.J.(2010). The effect of keyword method on ESP vocabulary learning.

Cobb T. (2007). Computing the vocabulary demands of L2 reading. *Language Learning & Technology*.

Creswell, J. W. (2003). Research Design: Qualitative, quantitative and mixed method approaches. (2nd ed.) Thousand Oaks, CA: Sage

Creswell, J. W. (2007). Designing and conducting mixed methods research. Thousand Oaks, CA: Sage

Daller, M.H. & Phelan, D. (2007). What is in a teacher's mınd? The relation between teacher ratings of EFL essays and different aspects of lexical richness. Testing and modeling lexical knowledge. Chapter:13. *Cambridge University Press*.

Davoudi, M., & Yousefi, D. (2016). The effect of keyword method on vocabulary retention of Senior High School EFL learners in Iran. *Journal of Education and Practice*, *7*(11), 106-113.

Demir, Y. (2013). The role of in-class vocabulary strategies in vocabulary retention of Turkish EFL learners. *Ilkogretim Online*, *12* (4)

De Ridder, I. (2002). Visible or invisible links: Does the highlighting of hyperlinks affect incidental vocabulary learning, text comprehension, and the reading process?

Diamond, L., & Gutlohn, L. (2006). Teaching vocabulary. *Retrieved from the reading rockets website: http://www. readingrockets. org/article/teaching-vocabulary*.

Elgort, I. (2011). Deliberate learning and vocabulary acquisition in a second language. *Language Learning*, *61*(2), 367-413.

Elgort, I., & Nation, P. (2010). Vocabulary learning in a second language: Familiar answers to new questions. *Conceptualising 'learning'in applied linguistics*, 89-104.

Fossey, E., Harvey, C., McDermott, F., & Davidson, L. (2002). Understanding and evaluating qualitative research. *Australian and New Zealand journal of psychiatry*, *36*(6), 717-732.

Fraenkel, Jack R., & Norman E. Wallen. (2008). How to design and evaluate research in education. 7th ed. New York: McGraw-Hill.

Guba, E. G. (1981). Criteria for assessing the trustworthiness of naturalistic inquiries. *Educational Technology Research and Development*, *29*(2), 75-91.

Guba, E. G. (1990). The alternative paradigm dialog. Newbury Park, CA: Sage

Hazenberg, S. & Hulstijn, J. H. (1996). Defining a minimal receptive second-language vocabulary for non-native university students: An empirical investigation. Applied Linguistics

Hornby, A.S., (1995). Advance learner's dictionary. New York: Oxford University Press.

Hu, M., & Nation, I. S. P. (2000). Unknown vocabulary density and reading comprehension. *Reading in a Foreign Language*, 13, 403−430

Hulstijn J.H. (1992) Retention of inferred and given word meanings: Experiments in incidental vocabulary learning. Vocabulary and applied linguistics. Palgrave Macmillan, London

Hung, H. T. (2015). Intentional vocabulary learning using digital flashcards. *English Language Teaching*, *8*(10).

Kelly, P. (1991). Lexical ignorance: The main obstacle to listening comprehension with advanced foreign language learners. *International Review of Applied Linguistics in Language Teaching*, *29* (2), 135-149.

Kilickaya, F., & Krajka, J. (2010). Comparative usefulness of online and traditional vocabulary learning. *TOJET: The Turkish online journal of educational technology*, *9* (2).

Koksal, O., & Çekiç, A. (2014). The effects of the mnemonic keyword method on 8th graders' L2 vocabulary learning. *Journal of International Scientific Publications*, *12*, 1030-1047.

Krashen, S.D. (1989). We acquire vocabulary and spelling by reading: Additional evidence for the input hypothesis. *The Modern Language Journal*.

Laufer, B. (1989). What percentage of text-lexis is essential for comprehension? In special language: from humans thinking to thinking machines, 316-323. Clevedon: Multilingual Matters.

Laufer, B. (2005). Focus on form in second language vocabulary learning. *Eurosla yearbook*, *5* (1), 223-250.

Laufer, B. & Nation, P. (1995). Vocabulary size and use: Lexical richness in L2 written production. Applied Linguistics, *16* (3). 307-322.

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

Loewen, S., Ellis, R. (2009). Incidental focus on form and second language learning. *Studies in Second Language Acquisition 27* (3), 361–86.

Mohebbi, H. (2013). Investigating vocabulary learning in second language classroom context: Recent findings, future outlook. *Advances in Asian Social Science (AASS). 4*(3). 882-886.

Morin, R. & Goebel, J. (2001). Basic vocabulary instruction teaching strategies or word? *Foreign Language Annals, 34* (1), -16.

Nation, I. S. P. (1990).Teaching and learning vocabulary. Boston, Mass.: Heinle&Heinle Publishers.

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

Nation I. S. P. (2001). How many high frequency words are there in English? Åbo Akademi University, Åbo: English Department Publications.

Nation , P . (2006). How large a vocabulary is needed for reading and listening? *Canadian Modern Language Review*, 63, 59 – 82.

Neuman, S. B., & Dwyer, J. (2009). Missing in action: Vocabulary instruction in pre-k. *The Reading Teacher, 62*(5).

Nunan, D. (1991). Language teaching methodology: A textbook for teachers. Sydney: Prentice Hall.

Paribakht, T. S. & Wesche, M. B. (1993). Reading comprehension and second language development in a comprehension-based ESL program. *TESL CANADA JOURNAL 11* (1).

Pellicer-Sánchez, A. (2016). Incidental L2 vocabulary acquisition from and while reading: An eye-tracking study. *Studies in Second Language Acquisition, 38* (1), 97-130.

Pellicer-Sánchez, A., & Schmitt, N. (2010). Incidental vocabulary acquisition from an authentic novel: do things fall apart? *Reading in a Foreign Language, 22* (1), 31.

Pigada, M. & Schmitt, N. (2006). Vocabulary acquisition from extensive reading: a case study. *Reading a foreign language*. 1-28.

Piribabadi, A. & Rahmany, R. (2014). The Effect of the keyword method and word-list method instruction on ESP vocabulary learning. *Journal of Language Teaching and Research*, Vol. *5* (5), 1110-1115 ACADEMY PUBLISHER.

Riahipour, P. & Saba, Z. (2012). ESP Vocabulary Instruction: Investigating the effect of using a game-oriented teaching method for learners of english for nursing. *Journal of Language Teaching and Research.* Academy Publisher. 3 (6), 1258-1266.

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

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

Schmitt, N. (2000). Vocabulary in language teaching. *Ernst Klett Sprachen*.

Schmitt, N. (2008). Instructed second language vocabulary learning. *Language teaching research, 12*(3), 329-363.

Schmitt, N. (2010). Researching vocabulary: A vocabulary research manual. *Springer*.

Schmitt, N. (2010). Vocabulary in language teaching. *Published by the press syndicate of the university of cambridge.*

Schmitt, N. & McCarthy, M. (1997). Vocabulary: Description, acquisition and pedagogy. *The University of Cambridge.*

Schmitt, N. & Schmitt, D. (2014). A reassessment of frequency and vocabulary size in L2 vocabulary teaching. *Language Teaching. 47* (4) 484-503.

Schonell, F. J., I. G. Meddleton & B. A. Shaw (1956). *A study of the oral vocabulary of* adults. *Brisbane: University of Queensland Press*.

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

Tabatabaei, O. & Hossainzadeh Hejazi, N. (2011). Gender differences in vocabulary instruction using keyword method (Linguistic Mnemonics). *Journal of Canadian Social Science, 7* (5), 198-204.

Tavakoli, M., & Gerami, E. (2012). The Effect of keyword and pictorial methods on EFL learners' vocabulary learning and retention. *Porta Linguarum.* ISSN: 1697-7467, 299-316.

Wang, Y. H. (2013). Incidental vocabulary learning through extensive reading: a case of lower-level EFL Taiwanese learners. *The Journal of Asia TEFL, 10*(3), 59-80.

Waring, R., & Takaki, M. (2003). At what rate do learners learn and retain new vocabulary from reading a graded reader? *Reading in a foreign language*, 15, 130-163.

Wilkins, D. A. (1972). Vocabulary. *Linguistics in language teaching*.

Yang, W. D. & Dai, W. P. (2012). Vocabulary memorizing strategies by Chinese university students. *Journal of International Education Studies*, *5* (1), 208-215.

Zeeland, H., & Schmitt, N. (2013). Incidental vocabulary acquisition through L2 listening: A dimensions approach. *System*, *41*(3), 609-624.

APPENDICES

A. VOCABULARY KNOWLEDGE SCALE (PRE- AND POST)

Appendix A: A Modified Version of the Vocabulary Knowledge Scale (Paribakht & Wesche, 1993)

1. I've never seen this word/phrase before.
2. I've seen this word/phrase before, but I don't know what it means.
3. I know what this word/phrase means, but I'm not sure how to use it (write definition in English or Turkish if you mark 3).
4. I know what this word/phrase means and I can use it in a sentence (use it a sentence if you mark 4).

1. Greet	**1**	**2**	**3 (write definition if you mark 3)**	**4 (write a sentence if you mark 4)**
Definition: Sentence:				
2. Hand out	**1**	**2**	**3 (write definition if you mark 3)**	**4 (write a sentence if you mark 4)**
Definition: Sentence:				
3. Head-rest	**1**	**2**	**3 (write definition if you mark 3)**	**4 (write a sentence if you mark 4)**
Definition: Sentence:				
4. Reclining seat	**1**	**2**	**3 (write definition if you mark 3)**	**4 (write a sentence if you mark 4)**
Definition: Sentence:				
5. Bassinet	**1**	**2**	**3 (write definition if you mark 3)**	**4 (write a sentence if you mark 4)**
Definition:				

Sentence:				
6. Inflate vest	**1**	**2**	**3 (write definition if you mark 3)**	**4 (write a sentence if you mark 4)**
Definition: Sentence:				
7. Upright	**1**	**2**	**3 (write definition if you mark 3)**	**4 (write a sentence if you mark 4)**
Definition: Sentence:				
8. Brace position	**1**	**2**	**3 (write definition if you mark 3)**	**4 (write a sentence if you mark 4)**
Definition: Sentence:				
9. Lift the buckle	**1**	**2**	**3 (write definition if you mark 3)**	**4 (write a sentence if you mark 4)**
Definition: Sentence:				
10. Assistance	**1**	**2**	**3 (write definition if you mark 3)**	**4 (write a sentence if you mark 4)**
Definition: Sentence:				
11. Pull on	**1**	**2**	**3 (write definition if you mark 3)**	**4 (write a sentence if you mark 4)**
Definition: Sentence:				

12. Exact change	**1**	**2**	**3 (write definition if you mark 3)**	**4 (write a sentence if you mark 4)**
Definition: Sentence:				
13. Off-loaded	**1**	**2**	**3 (write definition if you mark 3)**	**4 (write a sentence if you mark 4)**
Definition: Sentence:				
14. Stow	**1**	**2**	**3 (write definition if you mark 3)**	**4 (write a sentence if you mark 4)**
Definition: Sentence:				
15. Hold	**1**	**2**	**3 (write definition if you mark 3)**	**4 (write a sentence if you mark 4)**
Definition: Sentence:				
16. Bother	**1**	**2**	**3 (write definition if you mark 3)**	**4 (write a sentence if you mark 4)**
Definition: Sentence:				
17. Refreshments	**1**	**2**	**3 (write definition if you mark 3)**	**4 (write a sentence if you mark 4)**
Definition: Sentence:				

B. VOCABULARY KNOWLEDGE SCALE (DELAYED)

Appendix B: A Modified Version of the Vocabulary Knowledge Scale (Paribakht & Wesche, 1993)

1. I've never seen this word/phrase before.
2. I've seen this word/phrase before, but I don't know what it means.
3. I know what this word/phrase means, but I'm not sure how to use it (write definition in English or Turkish if you mark 3).
4. I know what this word/phrase means and I can use it in a sentence (use it a sentence if you mark 4).

1. Bassinet	**1**	**2**	**3 (write definition if you mark 3)**	**4 (write a sentence if you mark 4)**
Definition: Sentence:				
2. Pull on	**1**	**2**	**3 (write definition if you mark 3)**	**4 (write a sentence if you mark 4)**
Definition: Sentence:				
3. Refreshments	**1**	**2**	**3 (write definition if you mark 3)**	**4 (write a sentence if you mark 4)**
Definition: Sentence:				
4. Stow	**1**	**2**	**3 (write definition if you mark 3)**	**4 (write a sentence if you mark 4)**
Definition: Sentence:				
5. Greet	**1**	**2**	**3 (write definition if you mark 3)**	**4 (write a sentence if you mark 4)**
Definition:				

Sentence:				
6. Inflate vest	**1**	**2**	**3 (write definition if you mark 3)**	**4 (write a sentence if you mark 4)**
Definition: Sentence:				
7. Bother	**1**	**2**	**3 (write definition if you mark 3)**	**4 (write a sentence if you mark 4)**
Definition: Sentence:				
8. Brace position	**1**	**2**	**3 (write definition if you mark 3)**	**4 (write a sentence if you mark 4)**
Definition: Sentence:				
9. Lift the buckle	**1**	**2**	**3 (write definition if you mark 3)**	**4 (write a sentence if you mark 4)**
Definition: Sentence:				
10. Assistance	**1**	**2**	**3 (write definition if you mark 3)**	**4 (write a sentence if you mark 4)**
Definition: Sentence:				
11. Hand out	**1**	**2**	**3 (write definition if you mark 3)**	**4 (write a sentence if you mark 4)**
Definition: Sentence:				

12. Exact change	**1**	**2**	**3 (write definition if you mark 3)**	**4 (write a sentence if you mark 4)**
Definition: Sentence:				
13. Off-loaded	**1**	**2**	**3 (write definition if you mark 3)**	**4 (write a sentence if you mark 4)**
Definition: Sentence:				
14. Reclining seat	**1**	**2**	**3 (write definition if you mark 3)**	**4 (write a sentence if you mark 4)**
Definition: Sentence:				
15. Hold	**1**	**2**	**3 (write definition if you mark 3)**	**4 (write a sentence if you mark 4)**
Definition: Sentence:				
16. Upright	**1**	**2**	**3 (write definition if you mark 3)**	**4 (write a sentence if you mark 4)**
Definition: Sentence:				
17. Head-rest	**1**	**2**	**3 (write definition if you mark 3)**	**4 (write a sentence if you mark 4)**
Definition: Sentence:				

C. STRUCTURED INTERVIEW QUESTIONS

1. Did you find the keyword method effective in ESP vocabulary learning? If yes, why?

2. If you were asked to criticize the method, what criticism would you make?

3. What are your general opinions about the keyword method?

D. VKS SCORING IN ORAL COMMUNICATION

Possible scores	Meaning of scores
1	The word is not used at all.
2	The word is used with semantic appropriateness but without grammatical accuracy.
3	The word is used with both semantic appropriateness and grammatical accuracy.

E. SAMPLE PRESENTATIONS OF THE TARGET WORDS/PHRASES

My name is 'Lift'

<u>Lift the buckle</u>: Lift'e bakın

Lift isminde bir hostes, yolculara emniyet kemerinin nasıl bağlanıp nasıl çıkarıldığını göstermektedir. Tam o esnada bir yolcu lavaboya gitmek ister, ancak kemerini bir türlü çıkaramaz. Yanındaki yolcu yardımcı olmak ister ve şöyle der: "Lift'e bakın, kemerini nasıl çözüyor, **tokasını kaldırıyor.**"

e.g. To undo the seatbelt, please lift the buckle.

<u>Bother</u> : Bodur

1.40 boyunda bir kabin memurunun havayolu şirketine torpille girdiği düşünülmektedir. Diğer kabin memurları en az 1.60 ve üzeri boya sahip olduğu için bu kişiye Bodur lakabını verirler. Tüm kabin ekibi hakkıyla işe girerken, boy kriterine hiç uymayan birinin bodur haliyle aynı yerde çalışıyor olması ekibin **canını sıkmakta**, **rahatını bozmaktadır.**

e.g. I am sorry to bother you sir, but you must switch off your mobile phone.