



Ufuk University

Graduate School of Social Sciences

Department of English Language Teaching

**VOCABULARY TESTING TECHNIQUES AND IRAQI EFL
LEARNERS' VIEWS ON THEM**

Master's Thesis

Raad Thaidan

Ankara, 2015

**VOCABULARY TESTING TECHNIQUES AND IRAQI EFL
LEARNERS' VIEWS ON THEM**

Raad Thaidan

Ufuk University

Graduate School of Social Sciences

Department of English Language Teaching

Master's Thesis

Ankara, 2015

KABUL VE ONAY

2

Rasul Thaidan tarafından hazırlanan "Kelime Test ,Teknikleri ve Iraklı İngilizce Öğrenen Öğrencilerin Bu Tekniklere dair Görüşleri" başlıklı bu çalışma 07 .01. 2015 tarihinde yapılan savunma sınavı sonucunda başarılı bulunarak jürimiz tarafından Yüksek Lisans Tezi olarak Kabul edilmiştir.

Prof. Dr. Gülsev Pakkan (Başkan)

Yrd. Doç. Dr. Gülşen Demir (Danışman)

Yrd. Doç. Dr. Neslihan Özkan (Üye)

Yukarıdaki imzaların adı geçen öğretim üyelerine ait olduğunu onaylıyorum.



Prof. Dr. Mehmet Tomanbay

Enstitü Müdürü

BİLDİRİM

Hazırladığım tezin tamamen kendi çalışmam olduğunu ve her tezimin kağıtve elektronik kopyalarının Ufuk Üniversitesi Sosyal Bilimler Enstitüsü arşivlerinde aşağıda belirttiğim koşullarda saklanmasına izin verdiğimi onaylarım:

- Tezimin tamamı her yerden erişime açılabilir.
- Tezim sadece Ufuk Üniversitesi yerleşkelerinden erişime açılabilir.
- Tezimin..... yıl süreyle erişime açılmasını istemiyorum. Bu sürenin sonundağım takdirde, tezimin/ raporumun tamamı her yerden erişime açılabilir.

Raad Thaidan

DEDICATION

To Allah, The Almighty for His help and endless support in enlightening my path of knowledge, i dedicate this humble work.

To my parents who were very enthusiastic, proud and supporting through my studying.

To my faithful wife who supported and encouraged me forwardly to achieve this goal away from home.

To my patient and wonderful kids, Mohammad and Noorelhudah who were tolerant during my study abroad.

Finally, to my true friends, i respectfully and affectionately dedicate this work.

Raad Thaidan

ACKNOWLEDGEMENTS

I would like to take this opportunity to show my sincere gratitude to all of people who helped me and gave me the possibility along the way to achieve this dissertation.

Here, I would like to express my deep appreciation and gratitude to my thesis supervisor, Dr. Gülşen Demir whose assistance , stimulating, and feedback helped me in all the time of research for and accomplishing this study.

In addition, I would like to extend my unceasing thanks and warm appreciation to the Head of Department in the Faculty of Education at Ufuk University Prof. Dr. Gülsev Pakkan who supported, encouraged, and paved my way through the different stages of study.

Furthermore, I would like to express my special thanks to Dr. Neslihan Özkan who helped and gave me the possibility to complete this study as well as my special thanks to my excellent lecturer in the Faculty of Education, Prof. Dr. Ahmet Kocaman whose lectures are still dug deeply in my mind.

Finally, I have furthermore to thank my faithful friends who supported and motivated me to continue this research and never give up.

To them all I am very grateful.

Raad Thaidan

ABSTRACT

Raad, Thaidan. Vocabulary Testing Techniques and Iraqi EFL Learners' Views on Them, Master's Thesis, Ankara, 2015.

The main topic in this study is testing vocabulary. The purpose is to find out how vocabulary may be assessed. The research is categorized into two major poles; the theoretical and practical parts.

Throughout the theoretical part in chapter two, the researcher shed lights on the core of testing itself in language teaching and explaining the basic divisions of tests in terms of standardization related issues. Then, this study moves to deal with reasons of testing as well as to investigate inner mostly concerning the criteria of a good test like reliability and validity concerns. Types of tests are also handled. Following a general outlook on testing, testing vocabulary is taken into focus. After that, various techniques of testing vocabulary would be presented with their examples.

The practical part of the study consists of three sections. So, to see the effectiveness of various vocabulary testing techniques with the Iraqi EFL learners, the researcher tried to address some of variable well-known techniques with four groups of students in Republic of Iraqi public schools. These groups were divided as two female groups from secondary school, and coded as A and B, whereas the other two groups were males from another secondary school and were labeled as C and D. After the data collection, the result and discussion of these results are in the center. Finally, chapter five ends with the conclusion and suggestions in accordance with the study aims.

Key words: English Language Teaching (ELT), Testing Vocabulary, Vocabulary Tests, Test-taker, Test-setter.

ÖZET

Raad, Tlıaidan. Kelime Test Teknikleri ve Iraklı İngilizce Öğrenen Öğrencilerin Bu Tekniklere dair Görüşleri, Yüksek Lisans Tezi , Ankara, 2015.

Bu çalışmanın ana konusu 'Kelime Ölçmedir'. Buradaki amaç kelime bilgisinin nasıl ölçüleceği konusudur. Araştırmanın çerçevesi teorik ve uygulamalı kısımlar olmak üzere iki kutba ayrılmıştır.

İkinci ve üçüncü bölümlerdeki teorik kısım boyunca bir dilde ölçme konusuna ışık tutuldu ve ilgili hususların standartlaştırılması açısından testin temel bileşenlerini açıklandı. Daha sonra test hazırlamanın sebeplerine ve çokça bilinen güvenilirlik ve geçerlilik konularına değinildi. Ayrıca örneklerle birlikte test tipleri ve çok yönlü test tekniklerini gösterildi.

Üçüncü bölümün üçüncü kısmında, kelime bilgisi konusu yer aldı ve genel olarak dilin doğal bir parçası olarak önemine değinildi. Bunun yanısıra kelime bilgisinin temel boyutlarına odaklanıldı ve neden kelime bilgisinin ölçülmesi gerektiği sorusuyla ilgilenildi. Bundan sonra değişik kelime bilgisi ölçme tekniklerini örneklerle belirtildi.

Diğer taraftan dördüncü bölümün uygulama kısmında derslerde kelime bilgisi gelişimini ölçmek için en uygun tekniklere ve aynı zamanda test edilen grupların eğitim ve test edilme süreçleri boyunca ket vurma etkisine değinildi. Böylelikle, bu çalışma ile daha önce teorik kısımlarda belirtilen tekniklere vurgu yapılmıştır ve Irak'ta bulunan devlet okullarındaki dört densesel öğrenci grubuyla çalışılmıştır. Ortaokul düzeyindeki bu gruplar bayan öğrencilerden oluşan iki gruba ayrılmıştır ve bunlar A ve B grupları olarak kodlanmıştır. Ayrıca benzer şekilde erkek öğrencilerden oluşan gruplarda C ve D olarak isimlendirilmiştir.

Daha sonra dördüncü bölümde nitel veriler toplanmış ve kelime bilgisini ölçmede en uygun ve etkili teknikleri bulmak için analiz edilmiştir. Böylelikle katılımcıların başarıları ve dil öğrenme ortamında bu metodlarla etkileşimleri stratejik öngörü oluşturmuştur. Son olarak, beşinci bölümde sonuç ve öneriler sunulmuştur.

Anahtar kelimeler: İngiliz Dili Eğitimi (İDE), Kelime bilgisini ölçme, sınava girenler, test edenler, değerlendiriciler/notlayıcılar

TABLE OF CONTENTS

ACCEPTANCE AND APPROVAL OF THE JURY.....	I
BILDIRİM.....	II
DEDICATION.....	III
ACKNOWLEDGEMENT.....	IV
ABSTRACT.....	V
ÖZET.....	VI
LIST OF ABBREVIATIONS.....	XIII
LIST OF TABLES.....	XIV
LIST OF FIGURES.....	XV
CHAPTER 1.....	1
INTRODUCTION	
1.0 PRESENTATION.....	1
1.1 BACKGROUND OF THE STUDY.....	1
1.2 PROBLEM OF THE STUDY	2
1.3 AIMS OF THE STUDY.....	3
1.4 IMPORTANCE OF THE STUDY.....	4
1.5 RESEARCH QUESTIONS.....	5
1.6 LIMITATIONS OF THE STUDY.....	6

1.7 DEFINITIONS OF SOME TERMS.....	6
CHAPTER 2.....	9
REVIEW OF LITERATURE	
2.0 PRESENTATION.....	9
2.1 TESTING IN LANGUAGE TEACHING.....	9
2.1.1 LANGUAGE TESTS.....	10
2.1.2 REASONS FOR TESTING.....	13
2.1.3 PRINCIPLES OF TESTING.....	15
2.1.3.1 Practicality.....	15
2.1.3.2 Reliability.....	16
2.1.3.3 Validity.....	18
2.1.3.4 Authenticity.....	21
2.1.3.5 Washback.....	21
2.1.4 PREPARING A LANGUAGE TEST.....	23
2.2 TESTING VOCABULARY.....	24
2.2.2 THE DICHOTOMIES OF TESTING VOCABULARY.....	27
2.2.2.1 Discrete vs. Embedded.....	27
2.2.2.2 Selective vs. Comprehensive.....	28
2.2.2.3 Context-independent vs. Context-dependent.....	29
2.2.2.4 Receptive vs. Productive.....	29
2.2.2.5 Size vs. Depth.....	30

2.2.1 BASICS OF TESTING VOCABULAR.....	31
2.2.3 VOCABULARY TESTING TECHNIQUES.....	32
2.2.3.1 Multiple-choice.....	32
2.2.3.2 Word Formation.....	34
2.2.3.3 Writing Sentences.....	35
2.2.3.4 Cloze-Test.....	35
2.2.3.5 Matching Item.....	37
2.2.3.6 Sentences Completion.....	38
2.2.3.7 Definitions.....	38
2.2.3.8 Dictation.....	39
2.2.3.9 Odd Out.....	40
2.2.3.10 Translation.....	41
2.2.3.11 Associations.....	42
2.2.3.12 Synonyms and Antonyms.....	42
2.2.3.13 Transformation.....	42
2.2.3.14 Substitution.....	43
2.3 SUMMARY OF THE CHAPTER.....	43
CHAPTER 3.....	44
METHODOLOGY	
3.0 PRESENTATION.....	44

3.1 RESEARCH DESIGN	44
3.2 PARTICIPANTS	44
3.3 DATA COLLECTION PROCEDURE	47
3.4 DATA COLLECTION INSTRUMENTS.....	47
3.5 DATA ANALYSIS	48
CHAPTER 4.....	49
RESULTS AND DISCUSSION	
4.0 PRESENTATION	49
4.1 RESULTS.....	49
4.1.1 Multiple-Choice	49
4.1.2 Odd Out	51
4.1.3 Matching Items	53
4.1.4 Sentence Completion	55
4.1.5 Cloze Test	58
4.1.6 Dictation	60
4.2 DISCUSSION OF THE RESULTS	63
CHAPTER 5.	64
CONCLUSION AND SUGGESTION	
5.0 PRESENTATION.....	64
5.1 CONCLUSION	64

5.2 SUGGESTIONS	65
5.2.1 For Learners	66
5.2.2 For Teachers	66
5.2.3 For Material and Curriculum Designer	67
5.2.4 Implications for Further Research.....	68
5.3 SUMMARY OF THE CHAPTER	68
REFERENCES.....	69
LIST OF APPENDICES.....	73

LIST OF ABBREVIATIONS

ELT	English Language Teaching
TOEFL	Test of English as a Foreign Language
EFL	English as a Foreign Language
L1	First Language
L2	Second Language
NAEP	National Assessment of Educational Program
NAGB	National Assessment Governing Board
ITP	Institute Test Paper
ETS	Educational Testing Services
IELTS	International English Language Testing System
GPA	Grade Point Average
Ls	Learners

LIST OF TABLES

Table 1: The Difference between Informal and Formal test.....	12
Table 2: The Types of Reliability and Ways of Enhancing It.....	17
Table 3: Demographics of group (A).....	45
Table 4: Demographics of group (B).....	45
Table 5: Demographics of group (C).....	46
Table 6: Demographics of group (D).....	46

LIST OF FIGURES

Figure 1: Practicality.....15

Figure 2: Reliability.....17

Figure 3: Authenticity.....21

CHAPTER 1

INTRODUCTION

1.0 PRESENTATION

This chapter lays out the groundwork of this study such as the aims of the study and its importance. Then, the hypothesis and some related definitions will be presented at the final part of this chapter.

1.1 BACKGROUND OF THE STUDY

In recent years, “testing and assessment issues have been begun to figure with increasing prominence on the agenda of educational authorities around the world. In the wake of greater demands for public accountability, an increasing concern with standards and the imposition of more stringent reporting requirements, educational institutions have introduced a variety of testing and assessment procedures in order to make decisions on selection, certification and achievement” (Elder, Brown, Grove, Hill, Iwashita, Lumley, McNamara and O’Loughlin, 1998, p.126).

With its role as a vital distinctive component of successful communication concerns in speaking, writing, listening, and reading comprehension, vocabulary is at the center of these testing procedures. It is clear that less attention is paid to teaching or assessing vocabulary as compared to grammatical structures. Language teachers are ill-equipped in terms of effective strategies of modern teaching as well as education skills for conducting tasks in the classroom. These humdrum or derailing strategies are adopted by school teachers due to the previously rooted impacts of Audio-lingual Method as a dominant method for many years along with lack of effective training. Hence, Ministry of Education in Iraq started to update the methods of teaching and curriculum affairs after 2003 in order to meet the latest pedagogical needs of English language teaching (ELT) and to recalibrate the educational system. But many teachers

are still resisting to switch their old styles in implementing the curriculum. Although “the task for the ethical language tester is to look into the future, to picture the effect of the test is intended to have, and to structure the test development to achieve that effect. This is what we refer to as effect-driven testing” (Fulcher and Davidson, 2007, p.144). In order to achieve that, as Kachru (2005) puts, it “goes a step further and emphasizes that teachers need both awareness and training to be equipped to make their learners cognizant of the rich variation that exists in English internally” (as cited in Celce-Murcia, Brinton and Snow, 2014, p.67).

Despite all the innovations in the field, presenting vocabulary in an instructional lesson in Iraqi secondary schools is usually done by following the traditional mode like listing vocabularies with translation from the second language (L2) to the first language (L1) in order to be tested in the same manner in the exams. However, “Translation has a necessary and useful role in L2 learning, but it can hinder learners’ progress if it is used to the exclusion of L2-based techniques” (Richards and Renandya, 2002, p.260), which is usually what is done. So, the lack of multifaceted techniques based on the contextualization of vocabulary while teaching it is frequently overlooked in Iraqi English classes. As a result, English language learners in secondary schools are not exposed to those educational methods in terms of testing vocabulary which in turn results in negative instructional washback outcomes.

1.2 PROBLEM OF THE STUDY

Testing vocabulary has not received a great deal of attention in Iraqi secondary schools. Thus, students are already exposed to overdoses of syntax, which inhibits learner’s cognitive abilities to master communicative or writing skills for example.

It is clear that for decades, school teachers have considered vocabulary as a means rather than a goal. As a result, English language learners have found themselves learning vocabulary to practice structural patterns. This has led to find imbalance between vocabulary knowledge and structure despite the interest of vocabulary which has been attracted since 1980s.

Vocabulary plays a crucial role in learning a language. It is also a well-known fact that testing vocabulary is as important as teaching it. However, no genuine learning occurs because some teachers lack a pre-service and in-service knowledge of training in testing or teaching vocabulary. Such instructors are accustomed to transmitting the knowledge they received during their earlier education.

Consequently, teaching and testing vocabulary are the inseparable parts of a whole and they are necessary to help learners to succeed with the language. The techniques to enable the learners (Ls) to use the vocabulary in L2 should be picked carefully both in teaching and testing.

1.3 AIMS OF THE STUDY

Since the Iraqi old curriculum versions of English language have been changed into a new version, which adopts the Communicative Approach, the method of teaching has to be changed accordingly. In the same vein, this study seeks to investigate the techniques of testing vocabulary and to shed light on the advantages and disadvantages of these techniques with a modern approach in mind.

Some other perspectives regarding teaching and testing vocabulary are worth checking out at this point:

- Exposing Ls to various strategies in testing vocabulary in order to ensure that they are trained with enough activities and sufficiently modeled materials. That is, expanding learners' vocabulary knowledge, then enabling them to use or to develop these instructional strategies inside and outside the classroom. Bachman and Palmer (2012) state "in situations where we want to assess student's achievement or mastery of material covered in a course of information, if students do not have equal opportunity to learn this material, then the assessment may be biased" (p.207).
- Encouraging English language teachers to notice the importance of finding extra classroom activities and presenting them to second language learners in order to

widen their vocabulary. As Bachman and Palmer (2012) put it, “different types of assessment tasks may be appropriate for different areas of language ability for a given assessment purposes or group of test takers, and these tasks may vary in their efficiency in terms of adequately measuring the specific area in a given amount of time”(p.391). Variety of materials given to the students helps them in the aforementioned different areas of language ability.

- Helping students to develop their fluency and accuracy through well-designed and modeled activities which include already encountered words.
- Encouraging instructors of foreign languages to take more responsibility in teaching and in assessing vocabulary in terms of latest trends.
- Having a good scope of testing theory as well as utilizing the time accurately in conducting vocabulary tasks through modeling the techniques. This gives students more opportunities, motivating them to succeed. As Martin and Sugarman (1993) note, “many difficulties in classroom management can be prevented by effective teaching” (as cited in Woolfolk, Winne and Perry, 2003, p.423).

This study aims to demonstrate how beneficial the aforementioned elements are in the teaching and testing of vocabulary in Iraqi English language classes. Additionally, in the greater scheme of things, it is aimed to point out the uselessness of ineffective or outdated classroom applications and replace them.

1.4 IMPORTANCE OF THE STUDY

Vocabulary has an indispensable role in language research and practice. In other words, vocabulary in foreign language learning can be considered as the heart of the four skills of language writing, namely reading, listening, and speaking.

It is also impossible to deny the significance of testing vocabulary, which brings some issues forward. For instance, stakeholders are required to get a sufficient repertoire of information about how to maneuver in testing vocabulary. This study can function as

an- attempt to remedy this situation. Moreover, this research can work as feedback on the multifaceted tactics of tests and techniques for ELT professionals.

In the last two decades, the investigation of vocabulary in the field of ELT has gained its salience over developing the standardization of tests such as scaling vocabulary size, knowledge, and growth of productive or receptive words. Therefore, tests are accepted as major components of instructional programs across the world. Moreover, learners who are stuffed with more vocabulary items will be more likely to exchange communication concerns.

As a result, those students will achieve better in four skills of a language. As Read (2000) too suggests, “vocabulary can be seen as a priority area in language area of language teaching, requiring tests to monitor the learner's progress in vocabulary learning and to assess how adequate their vocabulary knowledge is to meet their communication needs” (pp. 1-2).

Henceforth, English language teachers can make use of this study in teaching and testing vocabulary by adopting or adapting their scholastic instructional strategies as Brown and Hudson (2003) point out, “adjustment will then either have to make how the material is taught or how it is tested. Although time consuming and occasionally frustrating, this process strengthens the program of instruction by affecting not only what the teachers teach, but also how they teach” (p.270).

1.5 RESEARCH QUESTIONS

This study aims to answer the following questions:

1. How do Iraqi learners of English score in vocabulary exercises with various vocabulary testing techniques when they are used in classroom practices?
2. What do these learners think of these techniques used?
3. Is there a difference between female students' motivation and willingness regarding these techniques and male students'?

1.6 LIMITATIONS OF THE STUDY

The current study can be said to present several limitations.

First of all, the research is limited to testing vocabulary of English language to Iraqi students as English as a foreign language (EFL) in secondary public schools. Secondly, the study is concerned with the Ministry of Education in Iraq, Directorate of Education in Dhi-Qar province with the teacher and learners in the aforementioned district, and the new coursebook, Iraq Opportunities 8 of secondary schools. Finally, the findings of the study are limited to the academic year of 2013-2014.

1.7 DEFINITIONS OF SOME TERMS

It is important to shed lights of certain terms that are presented in the research study to clarify some concepts.

1.7.1 Tests

Tests can be defined as “devices to reinforce learning and to motivate the student or primarily as a means of assessing the student's performance in the language” (Heaten, 1990, p.5).

1.7.2 Stakeholder

In language testing, Mousavi (2009) explains that stakeholders “include language tests (and the larger group of applied linguists of whom they form part), teachers, parents, administrators, teacher educators, sponsors and funding bodies, government and official bodies, the public, various national and international examination authorities, members of working parties and curriculum committees, test takers, (and the larger group of learners of whom they form part)” (p.646).

1.7.3 Objective Testing

Objective testing consists of tests “in which the learning material is divided into small units, each of which can be assessed by means of a test item with a single correct answer that can be specified in advance” (Read, 2000, p.75).

1.7.4 Informal Assessment

Harris and McCann (1994) explain informal assessment as “a way of collecting information about student's performance in normal classroom conditions. This is done without establishing test conditions such as in the case of formal assessment. Informal assessment is sometimes referred to as continuous as it is over a period of time like a term of an academic year” (p.5).

1.7.5 Standardized Language Tests

These types of tests “are used in some educational settings for both first and second language assessment. For example, bilingual education programs in the United States routinely require that limited-English-proficient students take standardized tests in English in order to determine their participation in bilingual versus mainstream English classes. Most foreign students applying to universities in the United States and Canada are required to take the TOEFL in order to qualify for admission” (Genesee and Upshur, 1996, p.233).

1.7.6 Washback Effect

Shohany (1992) offers a fairly complete scope of this term as “the utilization of external language tests to affected and drive foreign language learning....this phenomenon is the result of the strong authority of external testing and the major impact it has on the lives of test takers” (as cited in Brown, 2005, p.242).

1.7.7 Task

Breen (1987) suggests that “task is therefore assumed to refer to a range of work-plans which have the overall purpose of facilitating language learning from the brief and simple exercise type to more complex and lengthy activities such as group problem-solving or simulations and decision-making” (as cited in Willis and Willis, 2007, p.12).

1.7.8 Subjective Test

Subjective tests are the ones which “require scoring by opinionated judgment, mostly based on insight and expertise, on the part of the rater. An example might be the scoring of free, written composition” (Birjandi, Bagheridious and Mossalanejad, 2004, p.20).

1.7.9 Achievement Tests

Achievement tests are similar to progress tests in that “they determine what a student has learned with regard to stated course outcomes. They are usually administered at mid-and end-point of the semester or academic year. Achievement tests are often cumulative, covering material drawn from entire course or semester” (Coombe, Folse and Hubley, 2010, p.xvii).

CHAPTER 2

REVIEW OF LITERATURE

1.0 PRESENTATION

This part provides a look into the theoretical framework of the study expanding from testing in language teaching to testing vocabulary in ELT.

2.1 TESTING IN LANGUAGE TEACHING

In recent years, language testing has been digging deeply in teaching genuine communication. Over the last decades, the mode of testing was based on subjective or intuitive norms. Then, a distinctive revolutionary movement targeted the humdrum of traditional methods of language teaching and testing, where the classical testing strategies were piled on personal impressive scales. It was followed by a scientific stage which sounded thunderously towards testing domain. As Madsen (1983) puts it: “During the scientific era, many changes occurred. Testing specialists with linguistic training entered the scene. Careful linguistic descriptions suggested that language mastery could be evaluated 'scientifically' bit by bit” (p.6).

Twentieth century was also marked as the year when ELT researchers started to engage in vocabulary assessment because vocabulary has started to play a vital role in language users' lives. Hence, “the measurement of vocabulary knowledge in second language learners is of interest not only to language teachers, who are often required to make assessment of development of their learners' language proficiency, but also to researchers and test-developers who seek to develop valid and reliable measure of second language knowledge and use” (Daller, Milton and Daller, 2007, p.xii).

Thornbury (2004) suggests that “there has been a revival of interest in vocabulary teaching in recent years. This is partly due to the recent availability of computerized databases of word or corpora, and partly due to the development of new approaches to language teaching which are more 'word-centered', such as Lexical Approach” (p.vi). Additionally, to support the previous scope, “formal translation

generative linguistics, which previously took syntax as the primary focus, now gives more central to the lexicon and how the lexicon is formatted, coded and organized. Chomsky the father of contemporary studies in syntax has recently adopted ‘lexicon-is-prime’ professions in his Minimalist linguistic theory” (Richard and Rodgers, 2001, p.132).

It has been understood that vocabulary can be seen as “a priority area in language teaching, requiring tests to monitor the learners' progress in vocabulary learning and to assess how adequate their vocabulary knowledge is to meet their communication needs” (Read, 2000, pp. 1-2). Elder et al. (1998) clarify the importance of vocabulary in learning a language stating “The multifaceted nature of word knowledge means that deepening the knowledge of individual words can be viewed as one aspect of progress in vocabulary learning in general, and in second language learning in particular” (p.241). For instance, “the importance of vocabulary knowledge has been recognized in the development of reading skills. As early as 1924, researchers noted that, the growth in reading power relies on continuous growth in word knowledge” (Pikulski and Templeton, 2004, p.1). McCarthy (2005) also adds that vocabulary “will still be the largest single element in tackling a new language for the learner and it would be irresponsible to suggest that it will take care of itself in some ideal world where language teaching and learning discourse-driven” (p.64). Despite all the evidence, unjustified priority has been given to grammaticality when compared to other linguistic items, which is frequently the case for testing vocabulary, too.

2.1.1 LANGUAGE TESTS

Language tests represent a vital backbone in the educational field. In this sense, Davies (1990) simulates language testing with error analysis: “In recent years it has found itself taken up as a methodology for probing and investigation of language ability (and therefore of language itself) much as error analysis was taken up, first by contrastive studies, and later by second-language acquisition research” (p.9).

Shin and Crandall (2014) explain a language test as in the following: “a test can measure specific knowledge (for example, accuracy with specific vocabulary words, grammatical structures, or writing mechanics) or overall proficiency in using the language. It can be developed by a teacher or be a high-stakes, standardized test given by an educational institution or ministry. It can involve choosing among multiple-choice items, filling in blanks, completing sentences, writing texts, or giving short oral answers and yielding a score (for example 90 percent)” (p.247).

In short, in language assessment, testing can be defined as “any procedure used to measure a factor or assess some ability. Included in this general and encompassing sense of the terms are a wide variety of language tests as well as statistical tests which determine the significance of experimental results” (Mousavi, 2009, p.711).

Bachman (1995) also points out that “to elicit a certain behavior of test-takers performance, a test measures the participants’ knowledge or their specific characteristics in order to take decisions. The test is designed to elicit a particular behavior which distinguished a test as a device of judgment to other kinds of measurement” (p.20). It is obvious that a test is an effective tool to measure the characteristics of individuals’ knowledge in specific competencies or general abilities. Similarly, Galton states that, “the use of tests is like drilling a hole into the test taker to discover what is inside” (as cited in Fulcher, 2012, p.33). Therefore, through language testing we can thoroughly imbibe usable information for teaching and learning requirements to enhance the educational goal oriented programs. Basically, “a test is a formal, systematic (usually paper-and-pencil) procedure to gather information about student's behaviors” (Coombe et al, 2010, p.xv).

Language tests can be divided into two basic categories as standardized tests and non-standardized tests.

Harris (1969) defines standardized tests as “instruments which are prepared by professional testing services to assist institutions in the selection, placement, and evaluation of students” (p.1). Birjandi and Mosallanejad (2010) report that, “these tests are administered by highly authentic experts and skilled staff of professionals which cover a highly commercial standardized norms (codes of practice) like scoring or

reporting affairs, etc.” (p.47).Therefore, such type of tests is highly reliable in its administration and scoring.

The most well-known international examples of this type of test is the Test of English as a Foreign Language (TOEFL), which is administered by the Educational Testing Services (ETS). Another influential standardized language test all around the world is International English Language Testing System (IELTS), which is administered by British Council.

The core quality of standardized tests is that “they specify a set of competencies (or standards) for a given domain, and through a process of construct validation they program a set of tasks that have been designed to measure those competencies” (Brown, 2004, p.67).

Non-standardized tests can be tagged as ‘teacher-made tests’ and they are “small scale and are generally prepared, administered, and scored by one teacher” (Birjandi and Mosallanejad, 2010, pp. 47-48). In this category of tests, the material being tested is taken from the content of a specific course or the objectives of a test can be based on the course objectives.

Mousavi (2009) summarizes the differences between the informal 'non-standardized' teacher-made tests and formal 'standardized' tests in Table 1.

FEATURE	TEACHER-MADE TEST	STANDARDIZED TEST
Sampling of content	Both content and sampling are determined by classroom teacher	Content determined by curriculum and subject matter experts; involves extensive investigations of existing syllabi, textbooks and programs, sampling of content done systematically
Administering and scoring	Usually no uniform directions specified	Specific instructions, standardized administration
Interpreting scores	Score comparisons and interpretations	Scores can be compared to norm groups. Test manual and other guides aid
Reliability	Reliability usually unknown; can be high if carefully instructed	Reliability high; commonly between 0.80 and 0.96, frequent is above 0.90
Quality of test items	Quality of test items is unknown unless test item file is used. Quality typically lower than standardized due to limited time and skill of teacher	General quality of item high. Written by specialists, pretested and selected on the basis of effectiveness
Construction	May be hurried and haphazard;	Uses meticulous construction procedures

	often no test blue-prints, item tryouts, item analysis or revision	including test blueprint, item tryouts, item analysis, and item revisions
Norms	Only local classroom norms are used	In addition to local norms, these tests typically make available national school district and school building norms
Purpose and use	Best suited for measuring particular objectives set by teacher and for intra-class comparison	Best suited for measuring broad curriculum objectives and national comparison for inter-class, school, etc.

Table1: Differences between standardized and non-standardized tests (p.708)

2.1.2 REASONS FOR TESTING

English language teachers and educational authority across the learning process need to evaluate and gather information of students and teaching ongoing process between periods of time as criteria for judgment or achievement. Many logical reasons of testing are of a great importance to instructors, students, and administrators because those who are engaged in this activity need information, motivation, and reassurance periodically.

Heaten (1990) also contributes stating “a good classroom test will also help to locate the precise areas of difficulty encountered by the class or by the individual students” (p.6), which means a test can guide both the teachers and the learners in comprehending or evaluating the ongoing learning and teaching process. Moreover, “an understanding of language testing is relevant both for those actually involved in creating language tests and also more general for those involved in using tests or the information they provide, in practical and research context” (McNamara, 2000, p.5).

Mousavi (2009) lists the motives for testing below:

- find out about a candidate's suitability to follow a course of study although this is not the case in state education;
- Find out how a student is progressing during a course of study and possibly identify problem areas before a course ends;
- Compare a student's performance with that of other students;

- Find out how much a student has learned during the course of academic year, i.e., compare what students can do at the end of the course compared with the beginning of the course;
- Give the students a sense of achievement;
- Foster learning through diagnosing the problematic areas;
- Enhance learning by making students aware of the course objectives;
- Adjust the students' personal goals;
- Give promotion to students;
- Gauge the ability of teachers;
- Predict the possible success or failure in language acquisition of certain individuals;
- Show and maintain the effectiveness and efficiency of instruction;
- See if the materials for instruction were at the right level of difficulty;
- Get realistic STANDARDS for groups or individuals by comparing our results with those in similar classes learning under similar conditions;
- Assess—for specific purposes—the level of mastery (proficiency) of those who have not had formal language training program;
- See if all the language skills are being equally emphasized;
- Find out whether any point needs reviewing;
- See if the same material should be used the next year or they need some modifications;
- Offer remedial work to those students who have fallen behind, either because they have moved, been ill, or need a longer period of time in which to acquire a specific body of knowledge;
- Provide for ability or interest placement (or grouping) so that students can work with others who are on similar points of the continuum of one of the integrated communication skills, preparing freer dialogues or play-lets, or engaging in other productive instructional activities;
- Improve our understanding of how, where, and when bilinguals use certain aspects of the language they have learned; and
- Learn more about language acquisition process at various age levels in order to recommend either an earlier or later starting point, or more effective instructional measures for a target group. (pp. 725-726).

In accordance with this sense, all testers, teachers, and researchers aim to create infallible classroom testing because testing needs to have an effective backwash on learning and teaching. Obviously, teachers should try their best to find out or create a fruitful environment when setting classroom tests to find a positive washback effect or to evoke students' potential abilities to interact with the course material. Additionally, Smith (1994) suggests that “the goal of current measurement-driven reforms in

assessment is to build better tests that will drive schools toward more ambitious goals and reform them toward a curriculum and pedagogy general more toward thinking and away from rote memory and isolated skills” (as cited in Cheng, Yoshinari and Curtis, 2004, p.15).

With the types of language tests and reasons for using them in mind, the characteristics of testing should be covered for a clearer picture of the field regarding testing.

2.1.3 PRINCIPLES OF TESTING

Brown (2004) identifies five cardinal criteria for ‘testing a test’, namely practicality, reliability, validity, authenticity and washback.

2.1.3.1 Practicality

Practicality can be defined as “the relationship between the resources that will be *required* in the design, development, and use of the test and the resources that will be *available* for these activities” (Bachman and Palmer, 2012, p.262). This relationship can be represented as in the following figure:

$$\text{Practicality} = \text{Available Resources} - \text{Required Resources}$$

- If Available resources \geq Required resources, then Practicality is *positive*, and the assessment is practical.
- If Available resources $<$ Required resources, then Practicality is *negative*, and the assessment is not practical.

Figure 1: Practicality (Bachman and Palmer, 2012, p.262).

What Bachman and Palmer (2012) try to show in this figure is that “the test will be impractical if the existent resources cannot cover the required affairs to implement the test, which means without finding sufficient or extra proper resources, this will obstacle to administer the tests” (p.263).

Brown (2004) mentions that if a test requires expensive equipment to be implemented like reprographic and paper facilities or unique environmental problems in terms of administration should be expected. In such circumstances, Weir (2005) emphasizes the fact that, many countries over the world have such obstacles like the shortage of human resources, the appropriate building, the equipment like smart boards or tape recording devices, loud speakers system, and even power supplements; therefore, administering tests under these conditions makes the tests impractical. In order to make a test more practical, a pilot administration as Heaten (1990) suggests can be done: “in the case of all large-scale tests, the time to be allowed should be decided on a result of a pilot administration of the test (i.e. a tryout of the test to a small but, representative group of testees” (p.167).

In short, there are four criteria affecting the practicality of a test; a test

- “is not excessively expensive,
- Stays within appropriate time constraints,
- Is relatively easy to administer, and
- Has a scoring/evaluation procedure that is specific and time-efficient”
(Brown, 2004, p.19).

As long as one sticks to these criteria, it is possible to make a test practical in every sense.

2.1.3.2 Reliability

Crocker and Algina (1986) define the term ‘reliability’ as in the following: “whenever a test is administered, the test user would like some assurance that the results could be replaced if the same individuals were tested again under similar circumstances” (as cited in Fulcher and Davidson, 2007, p.104). In other words, if the same test is administered to the same group of candidates in two different settings and occasions (with no overloaded practices or extra sub-tests), whether on Wednesday morning or Sunday afternoon, it should not yield differences in terms of the test rank grades or level of marks. In accordance with this, Bachman and Palmer (2000) visualize reliability as a consistency of measurement. (p.19) adding that “a reliable test score will be consistent across different characteristics of testing situations. Thus, reliability can be considered to be a function of the consistency of scores from one set of tests and test

tasks to another” (Bachman and Palmer, 2000, pp. 19-20). They sum up their views on reliability in Figure 2 below:

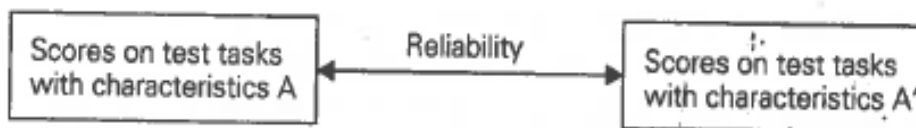


Figure 2: Reliability (Bachman and Palmer, 2000, p.20).

Brown (2004) mentions two features to meet reliability, namely consistency (reproducibility) and dependency.

Moreover, Genesee and Upshur (1996) come up with a table categorizing the types of reliability and ways of enhancing it:

<i>Type of reliability</i>	<i>Ways of enhancing reliability.</i>
<i>-Rater reliability</i>	<ul style="list-style-type: none"> <i>-Use experienced trained raters.</i> <i>-Use more than one rater.</i> <i>-Raters should carry out their assessments independently.</i>
<i>-Person-related reliability.</i>	<ul style="list-style-type: none"> <i>-Assess on several occasions.</i> <i>-Assess when person is prepared and best able to perform well.</i> <i>-Ensure that person understands what is expected (that is, those instructions are clear).</i>
<i>-Instrument-related reliability.</i>	<ul style="list-style-type: none"> <i>-Use different methods of assessment.</i> <i>-Use optimal assessment conditions, free from extraneous distractions.</i> <i>-Keep assessment conditions constant.</i>

Table 2: Types of reliability (p.60)

In line with the suggestions made by Genesee and Upshur (1996), students need to be provided with rubrics with descriptors concerning what they should know to obtain grades. These sufficient instructions will also make them familiar enough in order to meet all the standardized aspects of testing before engaging in the process.

2.1.3.3 Validity

As a significant characteristic of testing, validity is “the extent to which it measures what it is supposed to measure and nothing else” (Heaten, 1990, p.159).

Ensuring validity in a test is a challenging task since the end product should be the actual information which reflects the required characteristics that we want to know about. As Shin and Crandall (2014) put it: “Assessments have social and educational consequences; they must be fair, and what they are used for must match their intended purpose to be valid assessments” (p.252). Therefore, the test must aim to “provide a true measure of the particular skill which it is intended to measure: to the extent that it measures external knowledge and other skills at the same time, it will not be a valid test” (Heaten, 1990, p.159). In a nutshell, validity can be said to equal to “the meaningfulness of the scores, which defines a broad scope of concerns” (Luoma, 2008, p.184).

Brown (2004) lists five aspects of validity complementing each other, i.e. content validity, construct validity, criterion-related validity and face validity.

Brown (2004) explains what content validity stands for as in the following: “If a test actually samples the subject matter about which conclusions are to be drawn, and if it requires the test-taker to perform the behavior that is being measured, it can claim content-related evidence of validity, often popularly referred to as content validity” (p.22). Weir (2005) also contributes stating that context validity is “concerned with the extent to which the choice of tasks in a test is representative of the larger universe of tasks which the test is assumed to a sample. This coverage relates to linguistic and interlocutor demands made by the task (s) as well as the conditions under which the task is performed arising from both the task itself and its administrative setting” (p.19). In order to ensure that a test has content validity, one could follow Heaten’s (1990) suggestion: “When embarking on the construction of a test, the writer should first draw

up a table of test specification, describing in very clear and precise terms the particular language skills and areas to be included in the test. If the test or sub-test being constructed is a test of grammar, each of the grammatical areas should then be given a percentage weighting (e.g. the future simple tense 10 percent, uncountable nouns 15 percent, relative pronouns 10 percent, etc.” (pp. 160-161).

Criterion-related validity can be identified through determining how the raw scores are highly correlated to some external well-respected measure (criterion) of the same norms. In other words, “validity evidence is the strength of the predictive relationship between the test score and that performance on the criteria” (Fulcher and Davidson, 2007, p.5).

In this sense, Bachman (1995) states that “in collecting or gathering the statistical information obtained from the ongoing process of validation, there is an important link (coefficient of correlation) which indicates holding a relation among the scores of the test being validated with some other highly well-respected criteria as an indicator of the ability tested” (p.243). The author adds that, “the criteria in question may be level of ability as defined by group membership, individual's performance on another test of the ability in question, or their relative success in performing some task that involves this ability” (p.248). Criterion-related validity slides down into two sub-categories; *predictive validity* and *concurrent (status) validity*. The former reincarnates into the prediction of future performance.

According to Davies (1990), “this mode of validation correlates with the anticipation of the trial test judgment's results in terms of their future outcomes. For instance, we might plan to estimate a scholastic aptitude exam grades in order to expect candidates outcomes in English course or those students who are studying abroad foreign languages in the United Kingdom or in the States” (p.23).

Concurrent validity, on the other hand, is “based on a measure that is already at hand, usually another test, and in its most pure form concurrent validity can be established only when the test under scrutiny represents either parallel version of the criteria test or a simplified version of it” (Davies, 1990, p.24). As for concurrent (or

status) validity, it is possible to say that it is highly correlated with the already present performance of some other measure.

Messick (1989) defines *construct validity* as “an integration of any evidence that bears on the interpretation or meaning of the test scores” (as cited in Khodadady, 1999, p.138). Bachman and Palmer (2000) put it as “the extent to which the graders to interpret a specific test grade as indicators of the capacities or the characteristics of language ability knowledge that we aim to measure. (p.21). So, to grasp the meaning of this type of validity, one needs to identify ‘a construct’, which denotes as “an ability that provides the basis for a given test or test task and for interpreting scores derived from this” (Bachman and Palmer, 2000, p.21).

This suggests these abilities are already existent features which are reflected in test performance. There are certain components in language tests. Brown (2004) exemplifies one taking an oral interview into account as the performance to be evaluated. The grading criteria are fluency, grammatical accuracy, pronunciation, vocabulary knowledge, and sociolinguistic appropriateness. (p. 25). To sum up, construct validity suffers if the tester applies two of these components and drops the others.

It is suggested that “a consequential validity encompasses all the consequences of a test, including such considerations as its accuracy in measuring intended criteria, its impact on the preparation of test-takers, its effect on the learner, and the (intended and unintended) social consequences of a test's interpretation and use” (Brown, 2004, p.26).

As one of the most important factors in validity, face validity is based on subjective criteria. It is not empirically tested by an observer, but it is correlated with the physical vision of the supervisor. Face validity of a test is questioning “if it looks as if it measures what is supposed to measure” (Hughes, 2003, p.33). If a test components clear to other graders, instructors, and test-setters, its face validity can be set. Face validity in a test will likely be high if students encounter:

- A well-constructed, expected format with familiar tasks,
- A test that is clearly doable within the allotted time limit,

- Items that are clear and uncomplicated,
- Directions that are crystal clear,
- Tasks that relate to their course work (content validity), and
- Difficulty level that presents a reasonable challenge. (Brown, 2004, p.27).

2.1.3.4 Authenticity

The test-takers will perform actively when they face test tasks which include the norms of real-life situations implemented in the target language classroom activities. Teachers should depend on using authentic tasks to gauge learner's language skills. Thus, authenticity can be defined as “degree of correspondence between tests, tasks, and activity of target language use. Authenticity is the extent to which the tasks required on a given test are similar to normal real-life language use. Therefore, the higher the correspondence, the more authentic the test. Authenticity is, thus of interaction between the test-taker and the test task” (Mousavi, 2009, p.43). It relates the test task to the domain of generalization to which we want our score interpretations to generalize. Authenticity thus provides a means for investigating on the test to language use in the target language use domain, or to other similar non test language use domain (Bachman and Palmer, 2000, pp. 23-24).



Figure 3: (Authenticity. Mousavi, 2009, p.44)

As a result, the test-setters should adopt using authentic materials in testing and try their best in testing curriculum items in the context as it is in real life.

2.1.3.5 Washback

Washback refers to the impact which results from ongoing testing process on individuals, teacher, policy makers, classroom, and educational strategic system in

general terms. Alderson and Wall (1993) explain that, washback or backwash “is a term now commonly used in applied linguistics, refers to the influence of testing on teaching and learning, and has become an increasingly prevalent and prominent phenomenon in education”(as cited in Cheng et al, 2004, p.3).

Tests are a double-edged sword; they can have both harmful and beneficial effects on the stakeholders in the system. As Hughes (2003) illustrates, “If a test is regarded as important, if the stakes are high, preparation for it can come to dominate all teaching and learning activities. And if the test content and testing techniques are at variance with the objectives of the course, there is likely to be harmful backwash” (p.1). Person (1988) also points out “public examinations influence the attitudes, behaviors, and motivation of teachers, learners, and parents, and, because examinations often come at the end of a course, this influence is seen working in a backward direction-hence the term washback” (as cited in Cheng et al, 2004, p.7).

In order to remedy such effects, Lynch and Davidson (1994) describe an approach to criterion-referenced testing which involves “practicing teachers in the translation of curricular goals into test specification. They claim that this approach can provide a link between the curriculum, teacher experience, and test and can therefore; presumably, improve the impact of tests on teaching” (Alderson and Banerjee, 2001, p.214).

Similarly, Shohamy (1997) stresses that “the true power of tests is that of offering pedagogical benefits (i.e., promoting beneficial washback). This is exemplified in involving teachers in the test development process, and improving teaching through testing by considering concepts coming from innovation theory (Wall, 1996)” (as cited in Elder et al., 1998, p.147).

Otherwise, many of English language teachers ignore some classroom activities or simply bore learners with meaningless activities so that they can ‘catch up with’ the exam content, which results in negative teaching and learning experiences for both parties.

2.1.4 PREPARING A LANGUAGE TEST

Over the last two decades, a highly significant change has been seen in the language teaching practice area. The attention was focused on the role of a teacher in the classroom in terms of how to guide learners to maximize their educational outcomes. As a school teacher, it is not easy to construct a test individually because this creative process entails highly qualified experts as well as amount of time. A good extrapolated test with low-stakes or high ones requires careful planning because this critical item of testing carries an effective washback on the learning process.

Brown (2004) raises some questions to ponder before preparing a language test:

1-What are the targets of the test? Are we planning to measure learner's proficiency or for the placement purposes? Then we have to restrict the objective scope behind testing,

2-What is the objective of the test? What are we planning to find out? Holding proper objectives which include numerical issues, stepping from simple ones (forms and functions) of a syllabus to complex ones,
3-How will the test limitations reflect both the purposes and the objectives? We should be sure that the objectives are incorporated into a structure which weighs the variable competence being measured,

4-How will the test tasks be elected and the isolated ones arranged? The task activities should be practical as well as they ought to achieve context validity by presenting tasks which reflect those available in the coursebook being assessed. Moreover, reliability must be able to be evaluated by the testees.

5-What type of grading, scoring, and feedback? The reported results are varied in terms of the form or function of feedback so, grading the tests is connected with the washback effectively.

Ur (1996) also offers some guidelines to prepare a test:

“**Validity**, check that your items really do test what they are meant to!

Clarity, make sure the instructions for each item are clear. They should usually include a sample item and solution.

Do-ability, the test should be quite do-able: not too difficult, with no trick questions. Ask a colleague to read throughout it and answer the questions before finalizing.

Marking, decide exactly how you will assess each section of the test, and how much weighting (percentage of the total grade) you will give it. Make the marking system as simple as you can, and inform the testees what it is: write in the number of points allotted after the instructions for each question.

Interest, try to go for interesting content and tasks, in order to make the test more motivating for the learners.

Heterogeneity, the test should be such the lower-level students can feel that they are able to do a substantial part of the test, while the higher-level ones have a chance to show what they know. So include both easy and difficult items, and make one or more of the difficulty ones optional” (p.42).

2.2 TESTING VOCABULARY

Language is composed of various components, including vocabulary, grammar, and pronunciation. These bits are mixed in skills like listening or writing. So, it is possible to assess the mastery of each of these components individually as sub-skills. These cannot tell much about students’ mastery of language, but it helps teachers to diagnose the strong and the weak points in oral or written communication of their students.

Vygotsky defines a word as “a microcosm of human consciousness” (as cited in Thornbury, 2004, p.1). Thus, vocabulary can be considered as a skeleton of human language, which means without vocabulary communication will suffer. Thornbury (2004) also opines that vocabulary plays a tremendous role in learning and teaching languages: “All languages have words. Language emerges first as words, both historically, and in terms of the way each of us learned our first and any subsequent languages” (p.1).

Since vocabulary is so important, it is equally important to teach it with all its dimensions. In this sense, Ur (1996) identifies five aspects: form of a word, grammar, collocation, meaning and register.

As for the form of a word, “the learner has to know what a word sounds like (its pronunciation) and what it looks like (its spelling)...In teaching, we need to make sure that both these aspects are accurately presented and learned” (Ur, 1996, p.60). Therefore, focusing on a form will contribute to the student's learning. Long (1991) also suggests that “focus on form...overtly draws student's attention to linguistic elements as they arise incidentally in lessons whose overriding focus is on measuring or communication” (as cited in Cook, 2008, p.42).

It is important to pay plenty of attention to consolidating student's knowledge of grammatical structures and while teaching grammar. For instance, when an instructor teaches a new verb, “we might give also its past form, if this irregular (think, thought), and we might note if it is transitive or intransitive” (Ur, 1996, p.60). Introducing irregular plural forms of nouns while teaching vocabulary could be another example. As Francis (1995) explains the intermixed nature between lexical and grammatical description, “Particular syntactic structures tend to co-occur in a particular range of structures. In short, syntax and lexis are co-selected, and we cannot look at either of them in isolation” (as cited in Hunston and Francis, 2000, p.30)

Collocations refer to “the regular occurrence of words together in a language”(Richards and Rogers, 2001, p.133). Collocation words not used ALONE so, it is more efficient to learn the WHOLE and break it into parts than vice versa to improve fluency and native likeness. Learning how to collocate words and in what context, will help students to comprehend the mechanism of collocation. Firth (1957) argues that “you shall know a word by the company it keeps” (as cited in Hunston and Francis, 2000, p.230). Collocation, then is best seen as “part of a continuum of strength of association: a continuum that moves from compound words (*second-hand, record player*), including idioms (*out of the blue*) and phrasal verbs (*do up*), to collocations of more or less fixedness (*set the record straight, set a new world record*)” (Thornbury, 2004, p.7).

Ur (1996) categorizes meaning in general into the denotation which gives the literal meaning of a word in a dictionary verbatim such as, a dog which denotes to a kind of animal. The second one is connotation which reflects the metaphorical meaning of a word like, the dog which connotes to a friendship feature. Ur (1996) also opines that “a more subtle aspect of meaning that often needs to be taught is whether a particular item is the appropriate one to use in a certain context or not. Thus, it is useful for a learner to know that a certain word is very common, or relatively rare, or ‘taboo’ in polite conversation, or tends to be used in writing but not in speech, and in general much less common” (p.61).

In terms of relationships between words, there are some aspects of meaning which divide into several categories. The most familiar ones are (synonyms, antonyms, and hyponyms). Synonyms are words that have the similar meaning like (old- aged- ancient). Antonyms carry the opposite feature of meaning such as (old- new). Hyponyms are seen “when the meaning of one form is included in the meaning of another, the relationship is described as hyponym. Examples are the pairs: *animal / dog, dog / poodle, vegetable / carrot, flower / rose, tree / banyan*” (Yule, 2010, p.118).

In addition to all other aspects, register is another significant one. It is “a configuration of meaning” (Mahboob and Knight, 2010, p.95). According to Lee (2001), a register is “defined by lexico-grammatical and discoursal-semantic patterns associated with situations i.e. linguistic patterns” (as cited in Meyer, 2009, p.81). Bachman and Palmer (2000) distinguish between formal and informal register. McCarthy (1991) also states that register “is closely tied to choice of vocabulary words so that lexical choices are often significantly affected by the situation and the context of the conversation” (as cited in Wintergert and McVeigh, 2011, p.50).

Consequently, vocabulary “can be seen as a priority area in language teaching, requiring tests to monitor the learners' progress in vocabulary learning and to assess how adequate their vocabulary knowledge is to meet their communication needs” (Read, 2000, pp. 1-2). Similarly, Singleton (2000) also highlights the significance of vocabulary stating “interestingly, a child acquiring language appears to develop an awareness of words earlier than an awareness of how sentences are formed” (p.2).

In the same vein, Read (2000) emphasizes that, “the importance of testing vocabulary in use in order to gain a balanced picture of a learner's vocabulary” and adds that “Measures of vocabulary in use have the characteristics of having the assessment of vocabulary as part of a larger construct such as the ability to read informative texts, taking account of the total vocabulary context of the language use material, and involving the user in having to take account of average of contextual information” (as cited in Daller et al, 2007, p.41).

Assessing vocabulary is important both for achievement and diagnosis purposes. For instance, test-setters can collect infinite information about student's educational level. Results of tests can function as feedback given to learners.

It is also known that without testing there is no reliable means of knowing how effective teaching sequences are. As Thornbury (2004) puts it: “Testing provides a form feedback, both for learners and teachers. Moreover, testing has a useful backwash effect: if learners know they are going to be tested on their vocabulary learning, they may take vocabulary learning more seriously. Testing motivates learners to review in preparation for a test” (p.129). Hughes (2003) supports the very same idea: “Systematic teaching of vocabulary is desirable; vocabulary achievement tests are appreciated for their backwash effect” (p.179). Testing is also a way to revitalize learner’s vocabulary knowledge.

As a result of testing vocabulary, we can test learners’ proficiency, encourage their learning processes, and give feedback to students.

2.2.2 THE DICHOTOMIES OF TESTING VOCABULARY

Assessing vocabulary has infinite facets to be presented in testing process. The important issue for the test-setters before administering or processing the test is to decide on which aspects of vocabulary they plan to test and how to do it. In this sense, there are some dilemmas for test setters as Read (2000) lists such as discrete vs. embedded, selective vs. comprehensive, context-dependent vs. context-independent, receptive vs. productive and size vs. depth.

2.2.2.1 Discrete vs. Embedded.

In simple words, discrete test is assessing words as an independent constructs in order to gauge learner's knowledge characteristics at the word level. Therefore, “a discrete test takes vocabulary knowledge as a distinct construct, separated from other components of language competence” (Read, 2000, p.8). In addition, Mousavi (2009) defines a discrete test as “a test which measures knowledge of individual language item at a time. Therefore, each test item is aimed at one and one element of a particular

component of the language, such as phonology, syntax, or vocabulary. Such tests are designed to measure knowledge of performance in very restricted areas of the target language” (p.208).

In contrast, an embedded vocabulary test aims to assess vocabulary as a part of a large construct where students have to use it in an integrated format such as in a written essay (composition) or screen interview. So, such tests are usually used in international tests like TOEFL.

In accordance with this, students have to use various skills like listening or reading comprehension to answer sets of questions: “It is common practice to include in such tests a number of items assessing the learner's understanding of particular words or phrases in the text” (Read, 2000, pp. 9-10).

2.2.2.2 Selective vs. Comprehensive.

Another dichotomy that finds its place in testing vocabulary is about a test's being selective or comprehensive. This mode of assessment deals with the specification of vocabulary which included in the test. Selective vocabulary test is a kind of a prognostic tool, where specific words are selected to be a basal unit test.

Read (2000) argues that in a selective test, “a set of target words selected by the test-writers, and the test-takers are assessed according to how well they demonstrate their knowledge the meaning or use of those words. The target words either be selected as individual words and the incorporated into separate test items, or alternatively the test-writer first chooses a suitable text and then uses certain words from it as the basis for the vocabulary assessment” (p.10). The most familiar examples of vocabulary being selected are found in TOEFL reading comprehension part. Such tests are considered as competition tests.

On the other hand, a comprehensive vocabulary test measures the overall quality of words, as in spoken or written texts. Therefore, comprehensive mode does not look at specific words. As Read (2000) illustrates: “let us take a speaking test in which the learners are rated on various criteria, including their range of expression. In this case, the raters are not listening for a particular words or expressions but in principle are

forming a judgment of the quality of the test-takers overall vocabulary use” (p.11). As a result, this type of test is usually employed for placement and achievement purposes.

2.2.2.3 Context-independent vs. Context-dependent.

If words are not presented in a text to the testees, it means that these isolated or segregated words are categorized as context-independent. In this sense, “vocabulary is much more useful tested in context since it is the context that gives specific meaning and relevance to a word, thus creating a situation which is as linguistically valid as possible in the circumstances” (Heaten, 1990, p.53).

Therefore, when a student is engaged in a context to get the right answer as it happens in the reading passages of TOEFL.

The following example explains to which extent the testee must understand the meaning of the word *consume* according to the context:

“Humans have an innate ability to recognize the taste of salt because it provides us with sodium, an element which is essential to life. Although too much salt in our diet maybe unhealthy, we must consume a certain amount of it to maintain our wellbeing.

What is the meaning of *consume* in this text?

- a. Use up completely
- b. Eat or drink
- c. Spent wastefully
- d. Destroy” (Read, 2000, p.12).

2.2.2.4 Receptive vs. Productive.

Another issue in testing vocabulary, is to decide which words to include, namely receptive vocabulary (passive) or productive vocabulary (active). Productive vocabulary is the set of words that an individual can use in the writing or speaking process. They are words that are well-known, familiar, and used frequently.

Hiebert and Kamil (2005) state that, “receptive or recognition vocabulary is that set of words for which an individual can assign meanings when listening or reading. These are words that are often less well known to students and less frequent in use. Individuals may be able assign some sort of meaning to them, even though they may not

know the full subtleties of the distinction” (p.3). Receptive vocabulary, on the other hand, consists of words that the language user does not utilize actively.

It is clear that, students recognize passive vocabularies more than active ones, especially in reading passages or listening comprehension. It means that learners have more receptive vocabulary as compared to productive vocabulary, which is employed in writing and speaking activities.

2.2.2.5 Size vs. Depth.

Alderson and Freebody (1981) explain the distinction between size and depth as in the following: “The first may be called ‘breadth’ of knowledge, by which we mean the number of words for which the person knows at least some of the significant aspects of meaning....{There} is a second dimension of vocabulary knowledge, namely the quality or ‘depth’ of understanding so, we shall assume that, for most purposes, a person has a sufficiently deep understanding of a word if it conveys to him or her all of the distinctions that would be understood by an ordinary adult under normal circumstances” (as cited in Bogaards and Laufer, 2004, p.210).

Thus, the size of human beings’ vocabularies continue to expand on a daily basis and they are affected by the quality and type of language to which they are exposed as Albrechtsen, Haastrup and Henriksen (2008) put it: “Unsurprisingly, high frequent items are often acquired first due to the frequent and varied exposure; we get to these items in all types of language input. Low frequent items are often acquired and establish in our mental lexicon through reading input, and the acquisition of these words is often related to our educational experience in a more formal school setting and our widening literacy training and mastery” (p.58).

Vocabulary size has always received more attention because “despite the fact that the tests may seem superficial, they can give a more representative of the overall state of the learners' vocabulary than an in-depth probe of a limited number of words” (Read, 2000, p.115).

2.2.1 BASICS OF TESTING VOCABULARY

O'Dell (1997) “gives vocabulary a high profile in the syllabus and the classroom so that students can see its importance and understand that learning a language is not just about learning grammar” (as cited in McCarteen, 2007, p.20).

It is important that before engaging students to examinations, teachers need to comprehend the usefulness of assessment to students. It is also essential that test-setters and teachers be aware of the material being taught and included in the test. Richards and Renandya (2002) point out certain elements to teach in terms of vocabulary stating “language focused instruction thus includes focusing on pronunciation and spelling of words; deliberately learning the meanings of a word; memorizing collocations, phrases and sentences containing a word; and being corrected for incorrect use of a word” (p.270). These features are also the ones to be tested. While testing those aspects in addition to meaning, collocations, parts of speech, word's derivations, relative frequency, and specific register styles, how to do it should be another question to answer since contextualization or decontextualization plays vital roles in the process.

As Thornbury (2004) suggests: “if the purpose of the test is to predict the learner's reading ability for example, then a receptive test will be sufficient. A de-contextualized word test might be a valid test of reading ability. Moreover, it has been argued that de-contextualized tests encourage learners simply to learn long lists of words” (p.131). De-contextualized test items are easy to construct and score. For instance, in a dictation test, a teacher can assess spelling or mechanics of writing, without a context. This type of test is very practical to administer.

Whether a vocabulary test will stimulate the passive (receptive) vocabulary or the active (productive) one is another question. The former does not entail the invention of any words by test takers, an example of which can be seen below (Thornbury, 2004, p.131).

Choose the best word to complete each sentence:

1. The flight attendant asked the passengers to _____ attention to the safety demonstration.
a. Give. b. Devote. c. Pay. d. Lend.
2. A severe hurricane in the South Pacific has _____ many lives.
a. Claimed. b. Taken. c. Killed. d. Destroy.
3. The delegates blamed each other when the Peace talk broke _____.
a. Off. b.Up. c. On. d. Down.

The latter, on the other hand, requires the test-takers to invent the right word as an answer for each item (Thornbury, 2004, p.131).

Choose the best word to complete each sentence:

1. The flight attendant asked the passengers to _____ attention to the safety demonstration.
2. A severe hurricane in the South Pacific has _____ many lives.

2.2.3 VOCABULARY TESTING TECHNIQUES

There are multifaceted types of techniques in testing vocabulary. In this chapter the researcher is endeavoring to explain the most used ones with ample examples.

2.2.3.1 Multiple-Choice Items

This technique has traditionally a valuable structured nature. This format is useful due to the nature of reliability in terms of scoring objectively. These items are prepared with predetermined answers in mind, but they are difficult to design.

There are some guidelines to follow while writing multiple choice test items as Birjandi et al. (2000) lists:

- The stem should not start with a blank (meaningful learning),
- The stem should serve its function,
- The stem should carry as much of the information as possible,
- The item should not provide any clues for the candidates (intelligent guesses),
- The stem should not be so long and complicated,
- The alternatives should be of the same length and level of difficulty,
- The alternatives should be plausible,
- The alternatives should all be grammatical (error fossilization),
- The response should not be in the form of "all" of the above or "none of the above",
- The response should be only one, and
- The response should be distributed randomly. (p.54).

A single word's definition can be asked as in the following:

“A *knight* is
a- world champion.
b- chess piece.
c- playing card” (Barnes, Hines and Weldon, 1996, p.71).

Vocabulary can also be tested in sentences.

“July 4 is when Americans **observe** Independence Day.
a- Celebrate.
b- Recognize.
c- Demonstrate” (Burgmeier, 1991, p.57)

A multiple choice test item should be prepared carefully so that there will not be multiple correct answers like the one Madsen (1983) provides below:

(Poor) “She wrote _____ yesterday.
a- Letter
b- Gift
c- Friend
d- Book” (p.20).

Thus, while writing the distractors, it should be kept in mind that all of them are at the same level of difficulty. Harris (1969) provides a similar example stating “the


keyed response of the following vocabulary item, i.e., *nefarious*, is more difficult than its distractors, i.e., *delicious*, *prosperous* and, *courageous*”:

- Very evil
a- Nefarious*
b- Delicious
c- Prosperous” (as cited in Khodadady, 1999, p.50).

Therefore, we have to avoid unintentional clues which can help students to guess the right answer without having the specific skill or knowledge being assessed.

Mascull (2002) presents a perfect way in terms of contextualizing multiple-choice items, where the test-taker has to get the right option while engaging in a text such as:

Carla used to work for an Italian magazine publishing company. She talks about how she lost her job. Choose the correct form of the words in brackets to complete the text.



Edizione Fenice is a big magazine publishing company, and a very nice company to work for. I was director of a magazine called *Casa e Giardino*.

Then, Fenice was bought by an international publishing group. We had to have regular performance (1) (review/reviews/reviewer) with one of the new managers. After a few months they started laying staff (2) (off/on/out). Our own journalists were put on temporary (3) (contracts/contractual/contracting) or replaced by (4) (freelancer/freelancers/freelanced).

Then they started (5) (laid/lying/laying) off more senior people like me. The new owners said they wanted to make the company (6) (flat/flatter/flatten) and (7) (lean/leant/leaner). So I was made (8) (redundant/redundancies/redundancy). They offered to help me to find another job with (9) (outplacement/outplaced/outplacing) advice, but I refused.

(As cited in Thornbury, 2004, p.23).

2.2.3.2 Word formation

Many words can be modeled in various forms to become different parts of speech. The noun ‘power’, for instance, can be transformed into a verb (empower) or into an adjective (powerful). In this sense, students are asked to convert the form of a word to be best fitted in a specific sentence. This technique requires students’ comprehension of the text along with their knowledge of words. The test-takers are required to use their knowledge of derivation to convert the word to fit into a sentence:

“Change the word on the left into a suitable form to fill the gap:
 1- *Compose* On one occasion the opera was conducted by the

 2- *Place* _____ Have you seen my keys? I seem to have
 _____ them” (Thornbury, 2004, p.134).

2.2.3.3 Writing sentences

Here, Ur (1996) explains that in this technique, the test-setter offers specific words and then asks students to create sentences with the words given. This test encourages learners to use their knowledge (grammatical or semantic) in order to construct particular words in a context. (p.72).

Write complete sentences with these words.

archaeologists spicy Ancient Romans
 explorers restaurant factory

1 _____
 2 _____
 3 _____
 4 _____
 5 _____
 6 _____

(Penn, 2010, p.9).

2.2.3.4 Cloze-test

It is argued that such brand of tests are considered as a test of productive skills because it entails candidates to integrate grammatical, contextual, lexical, and pragmatic knowledge so as to be able to supply the missing words. The purpose of cloze-test is to trigger students' active vocabulary because the options are not provided in the gapped text.

The cloze tests function better if the assigned formats do not deal with topics extraneous to test takers' experience which preventing them from gaining an overall comprehension of the content. Cloze tests, however, are usually criticized for being a subjective technique and for testing reading instead of vocabulary.

One of the solutions offered is to give the initial letter of the word that the student is expected to fill the blank with so that the number of possible answers can be limited to some extent. Thornbury (2004) supports the use of cloze tests in testing vocabulary stating "it has been argued that C-tests are valid tests of overall vocabulary knowledge, and thus can usefully serve as placement tests" (p.133).

An example of a cloze-test is provided below:

(Bribery)

Instructions. Fill each space in the text below with a suitable word. (Note that only one word per space is allowed and that it must make grammatical sense.)

Bribery seems to be a common problem these days. Perhaps part of the problem is that it can be easily c _____ f _____ d (1) with the offer of rewards or other inducements. Rewards are undoubtedly positive – particularly if you are the one receiving them! Inducements too can be positive – depending on what they are intended to i _____ e (2)? For example, a _____ n _____ u _____ (3) scheme at work – or some other such inducement – to work hard, is perfectly acceptable and almost par for the _____ o _____ m (4) these days, or certainly very _____ c _____ m _____ n _____ l _____ c _____ (5). However, the offer of inducements by _____ c _____ n _____ t _____ c _____ i _____ (6) companies to those who make the decisions when _____ d _____ d _____ n _____ (7) for a contract to build something is perhaps not so acceptable. Additionally, the offer of an inducement to a policeman to perhaps look the _____ h _____ r (8) way after committing a minor _____ i _____ a _____ _____ c (9) offence should not be considered positively either – even though it might help _____ s _____ p _____ p _____ m _____ t (10) his rather _____ m _____ _____ g _____ (11) wages! The word inducement generally has positive connotations, but can be easily confused with the word bribe. Then again, even bribes can be looked on _____ f _____ v _____ _____ r _____ b _____ (12). Every parent in Poland, for instance, it seems, will face the _____ i _____ i _____ i _____ m _____ (13) at some point in their life of whether to offer their child some kind of bribe in order to encourage them to get better results at school! This may well have the _____ d _____ _____ i _____ _____ d (14) effect in the _____ h _____ _____ t (15) term – but is it really a good thing? Perhaps this is something that requires some deep thought..!

(Phillips, 2002, p.13).

2.2.3.5 Matching

In this technique, the students have to match the word to the correct option. Through matching, the goal is to measure learner's capacity in comprehending and recognizing synonyms or antonyms. Mousavi (2009) adds it is also possible to test “names of objects and names of groups or classes of objects, names of objects and occupations, names of objects and their pictures” (p.400).

This type of questions does not ask test-takers to invent or produce words, but only to match the given words. It is easy to construct those items as compared to multiple-choice items. Here, Brown (2005) argues that the options are supposed to be more than the words in order to prevent students from narrowing down the choices too quickly. Thus, the author suggests that, to fix this issue, it is possible to add more options than premises. Moreover, the author added to make the options shorter than premises and both of them should be related to one central theme. (p.50).

Example 1: Read the sentences. Then match the word with its definition below:

- 1-Red shoes are a new fashion **trend**. Many people wear them now.
- 2-Five years ago, no one knew who he was. Now he's **celebrated**. His name is in the newspaper every day.
- 3-That's the most **ridiculous** idea I've ever heard. It will never work.
- 4-Bill told a **joke** at dinner last night. We all laughed very hard.
- 5-Driving a car without wearing a seatbelt is **risky**. You can get hurt in an accident.
- 6-The person who wrote the article was very **critical** of the movie. He didn't like it at all.
- 7-He's a very experienced **journalist**. He works for a very important newspaper company.
- 8-I **admire** Elaine very much. She's a great artist and a very nice person.

- | | |
|----|--|
| a. | To respect someone very much |
| b. | A famous person |
| c. | Saying what is wrong with something |
| d. | A popular idea or change |
| e. | A thing that someone says that is funny, not serious |
| f. | A person who collects, writes, and publishes news |
| g. | Very silly or unreasonable |
| h. | Dangerous" |

(Brooks, 2011, p.5).

Example 2:

For Set One, match each word with its synonym. For Set Two, match each word with its antonym.

Synonyms**Set One**

- | | |
|---------------------|------------------|
| _____ 1. blurb | a. record |
| _____ 2. retrospect | b. advertisement |
| _____ 3. networking | c. identify |
| _____ 4. empathize | d. remember |
| _____ 5. chronicle | e. sharing |

Antonyms**Set Two**

- | | |
|-----------------------|---------------|
| _____ 6. malicious | f. open |
| _____ 7. confidential | g. ugly |
| _____ 8. elaborate | h. suspicious |
| _____ 9. exquisite | i. kind |
| _____ 10. gullible | j. simple |

(Olsen, 2010, p.76).

2.2.3.6 Sentence Completion

In this technique, the test-setter provides students with incomplete sentences. The students have to complete the sentences taking the prompts given into consideration.

Complete the answers.

How do I get to the bank?

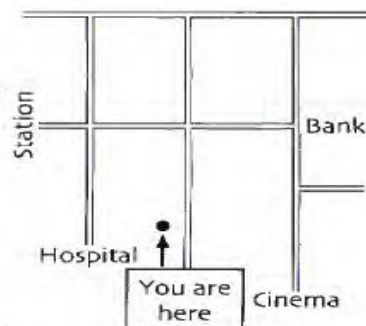
Go (a) *along* this road and turn (b) _____.

How do I get to the station?

Go (c) _____ this road. Then turn
(d) _____ and go (e) _____ on.

How do I get to the cinema?

(f) _____ this road and (g) _____.
Then (h) _____ again and (i) _____.



(Doff and Jones, 2007, p.59).

2.2.3.7 Definitions

When a teacher uses this technique, s/he asks students to provide definitions for the words given or the other way round. Mousavi (2009) states that the technique has “the advantage of being highly economical in terms both of the number of items that can

be included on a printed page and of the number of problems that can be answered in a short period of testing time” (p.181).

Example 1:

“Nap:

a- A brief sleep

b- A happy song

c- A sharp rock

d- A short meeting

A brief, light sleep:

a- Nap **c-** Stroll

b- Yawn **d-** Hug” (Mousavi, 2009, p.181).

Example 2:

1- Define a Jack. (Machine – lift and lower heavy things).

2- Define a pen. (Tool – write).

2.2.3.8 Dictation

Dictation is a technique in which a dictator dictates a text of a connected discourse to students with certain pauses. Baker (1989) explains that “dictation tests can give information about a wider range of features” (p.73), one of them being vocabulary.

Mousavi (2009) explains in this technique a dictated material of connected discourse is read, loudly to the examinees. This directed passage is about (50-100) words to be read to students with three broken modes of time. The first mode is at a normal speed so as to enable testees to decode the details (listen for gist). The next reading is done by breaking the dictated text into bursts or units, to store the details. The third reading is with normal speed and without breaking where the test-taker have to re-encode the text.

Example 1:

First reading at natural speed with no pauses:

Students go to university to get an education and find a good job. A CV or curriculum vitae is an essential job search tool. A CV is a document that lists your contact details, education, work history, skills, and achievement. A well-written CV is necessary to promote yourself to future employers. According to career advisors, students need to start working on their CVs as soon as they begin college. They recommend that CVs should be free of spelling and grammar mistakes. Things like this can make a bad impression on future employers.

Second reading with breaks at // and slower speed:

Students go to university // to get an education // and find a good job. // A CV or curriculum vitae // is an essential job search tool. // A CV is a document // that lists your contact details, // education, work history, // skills, and achievement. // A well-written CV // is necessary to promote yourself // to future employers. // According to career advisors, // students need to start working // on their CVs as soon as they begin college. // They recommend that // CVs should be free of spelling // and grammar mistakes. // Things like this can make a bad impression // on future employers.

Third reading at normal speed with no pauses:

Students go to university to get an education and find a good job. A CV or curriculum vitae is an essential job search tool. A CV is a document that lists your contact details, education, work history, skills, and achievement. A well-written CV is necessary to promote yourself to future employers. According to career advisors, students need to start working on their CVs as soon as they begin college. They recommend that CVs should be free of spelling and grammar mistakes. Things like this can make a bad impression on future employers.

(Coombe et al, 2010, pp. 100-101).

2.2.3.9 Odd One Out

Here, students are asked to pick the word standing out among a group of words. Ur (1996) states that “this technique is easy to score, and only meaning is being tested in this mode. Therefore, the teachers have to know that their students are already well familiar with the material being assessed” (p.72).

Example 1:

Which is the odd one out in these words and expressions relating to music and why?

Example: guitar, piano, violin, harp *The piano is the only one that does not have strings that you can see.*

- 1 contemporary music, electronic music, seventies music, 20th century music
- 2 soothing, relaxing, discordant, tuneful
- 3 orchestral music, chamber music, big band music, rock music
- 4 background music, soundtrack, contemporary music, dance music
- 5 jazz, blues, muzak, heavy metal

(McCarthy and O'Dell, 2002, p.40).

Example 2:

Spot the "odd word out." [The word that doesn't belong in each list.]

Example: radio, computer, video, television

Discuss the following words. Put a circle around the odd word out and say why it is the odd word.

1. tourist, visitor, traveler, student
2. investigate, determine, explore, inquire
3. elderly, intelligent, stupidly, talkative
4. utilize, uncover, reveal, disclose

(Nunan, 2003, p.16).

2.2.3.10 Translation

This technique is classically used to test how well the learners are able to decode speech or writing from native language into a target language, or vice versa. Thus, translation "is a type of measurement of cross cultural communication including the rendition of meaning from one language to another" (Mousavi, 2009, p.784).

It is suggested that there are two advantages of this brand of technique. "First, having a speaker talk freely about his home country produced a reasonably spontaneous text. Second, the requirement to translate ensures that test-takers have to process the semantic content of the text, which clearly meets Anderson's (1972) criteria of a test of comprehension" (Buck, 2007, p.82). But, this technique does not reflect wide areas of teaching situations, because over reliance on it may bear negative access to develop an independent L2 lexicon. Moreover, scoring of this technique may yield a problematic washback. Another drawback might be "finding exact equivalent across languages and it may be tricky to mark" (Ur, 1996, p.73).

Example:

Write In your language

It won't rain on Friday.

I'll see you tonight.

Will you be at the meeting tomorrow?

(Doff and Jones, 2007, p.157).

2.2.3.11 Associations

This technique, calls students ability to sort out or to recognize the words which are belong to a certain group given in a task.

Example: Underline the words that belong to *Dairy Products*:

(Butter – chess – yoghurt – jam – cheese – diaries – cream – coconut).

2.2.3.12 Synonyms and Antonyms

Here, students have to write down the similar or opposite meaning of given items either by matching two lists or by creating their own.

Example: “Look in the dictionary to find one more synonym for each of these words. Write a sentence with each new synonym.

1- Word: huge Synonym: _____

Sentence: _____

2- Word: laugh Synonym: _____

Sentence: _____

3- Word: funny Synonym: _____ ”

Sentence: _____

(Scanlon, 2011, p.112).

Example 2:

For each of the words find (a) a near synonym and (b) an antonym from the words in the box.

Example: innovative: synonym = creative; antonym = unoriginal

background	classical	creative	discordant	loud	modern
relaxing	rousing	serious	soft	tuneful	unoriginal

1 deafening 2 soothing 3 light 4 tuneless 5 contemporary

(McCarthy and O'Dell, 2002, p.46).

2.2.3.13 Transformation

In this technique, test-takers have to rewrite the given sentence keeping the same meaning.

Example:

Martin was jailed three weeks for his misdemeanor.

He _____.

Sarah as a cook has a brilliant recipes of foods.

She cooks _____.

2.2.3.14 Substitution

In substitution, test-takers are asked to rewrite the original sentence in the form which is indicated by the given words.

Example:

He writes a lampoon against his teacher.

You _____.

The temperature increased up in summer.

It _____.

2.3 SUMMARY OF THE CHAPTER

This chapter has provided the theoretical framework, which includes a look into testing in general terms and into testing vocabulary specifically. The theoretical principles that are dealt with here will serve as guidelines in the following chapters.

CHAPTER 3

METHODOLOGY

3.0 PRESENTATION

This chapter includes the details regarding the research aspect such as the design, participants, data collection instruments, procedures and so on.

3.1 RESEARCH DESIGN

The study was more of a qualitative nature than of quantitative. The setting for it was two secondary schools in Iraq.

The tested groups were assigned to have ongoing instructional and practices of vocabulary techniques in order to explore the effectiveness of these different methods via teachers procedural instructions.

The study aimed to find out;

1. How do Iraqi learners of English score in vocabulary exercises with various vocabulary testing techniques when they are used in classroom practices?
2. What do these learners think of these techniques used?
3. Is there a difference between female students' motivation and willingness regarding these techniques and male students'?

3.2 PARTICIPANTS

The participants in the study were students who were learning English as a foreign language in Iraqi public schools. The tested groups were chosen randomly from AL-Khawarizmy and AL-Nesour secondary schools with cooperation of the General Directorate of Education in Dhi-Qar province. The study was conducted in the second semester of scholastic year 2013/2014.

These selected students were divided randomly into four groups labeled accordingly as A, B, C, and D. Further details regarding the participants can be found below:

Name of the school	AL-Nesour
Number of the participants	15 students
Educational grade level	4 th grade/Preparatory
Gender	Female
Age	16-18 years old
Group description	This group of female students was highly motivated to learn English. They expressed their eagerness and interest to take part in the study. There were no obstacles or problems with the discipline. They were very positive throughout the process.

Table 3: Demographics of Group A.

Name of the school	AL-Nesour
Number of the participants	15 students
Educational grade level	4 th grade/Preparatory
Gender	Female
Age	16-18 years old
Group description	This group was studying English language for seven years in public schools. They enjoy learning English as a foreign language. They were also enthusiastic about participating in the study. They were similar to Group A on many accounts.

Table 4: Demographics of Group B.

Name of the school	AL-Khawarizmy
Number of the participants	15 students
Educational grade level	4 th Grade/Preparatory
Gender	Male
Age	16-19 years old
Group description	This group had problems with their attention span. They were also used to the humdrum of classical methods of teaching English language, which resulted in a negative classroom atmosphere.

Table 5: Demographics of Group C.

Name of the school	AL-Khawarizmy
Number of the participants	15 students
Educational grade level	4 th grade/Preparatory
Gender	Male
Age	16-19 years old
Group description	This group was eager to learn and it consisted of students who were moved from intermediate 3 rd . grade after passing the Public Bachelor exam. But this group was still suffering from the traditional styles of teaching English. Therefore, they needed a lot of instructional effort.

Table 6: Demographics of Group D.

All participants had Iraq Opportunities 8 as their coursebook in English classes. The set of books the participants used consisted of the coursebook itself, the workbook and the teacher's book.

3.3 DATA COLLECTON PROCEDURE

The first step in the data collection procedure was to outline it with the teachers that were expected to assist throughout the study. Several points such as the classroom proceedings, scoring issues were clarified in advance.

After it was ensured that all the participants were at the same point in their English classes, i.e. 'Unit 4-Cyberspace' (see Appendix 1), data collection started.

During the eight weeks the study was conducted, the participants were provided with vocabulary exercises including techniques such as multiple-choice, odd one out, matching, sentence completion, cloze test and dictation. Afterwards, they were given vocabulary tests with these techniques. Following the administration of each test, the participants were asked of their ideas on the exercises and tests that they were given.

3.4 DATA COLLECTION INSTRUMENTS

There were two main data collection instruments used in the study. First one was the tests given to the participants (see Appendices 3-9) and the second one was the written questionnaires and oral interviews conducted with the participants about their thoughts on the techniques used in the classroom. The questions asked were:

- Are you happy or bored with this experience?
- How do you find the instructional conditions; were they good or bad? Why?
- Do you think the instructors' feedback, clues, and other educational styles can help you to learn more?
- How long did you study vocabulary at home?
- Did you follow your instructor's advices? If not why?
- What do you think about the way that we did the activity?

The interviews were conducted in Arabic.

3.5 DATA ANALYSIS

The data gathered as result of the administration of the vocabulary tests and the interviews was analyzed separately. First, the tests were scored and checked if there was a difference between female and male participants in terms of achievement. Then, the interviews were analyzed to see what the participants thought of the vocabulary testing techniques used in the classroom.

CHAPTER 4

RESULTS AND DISCUSSION

4.0 PRESENTATION

This chapter focuses on the results of the study as well as the discussion in terms of the research questions asked in the very beginning.

4.1 RESULTS

The tests applied and the interviews conducted with the participants yielded important results. Each and every technique used is handled separately. Moreover, the results of the tests and the interviews are also dealt with separately to provide a more detailed insight into the data.

4.1.1 Multiple-choice

After being given multiple choice exercises for some time, participants were also given a test with the same type of questions. Special attention was paid by the instructors to make sure that the questions were free from extraneous clues (see Appendix 2).

Group A

Being familiar with the technique, this group reacted positively as it was expected due to their eagerness to learn English as well as the strength of their knowledge.

The achievements of this group over the three multiple-choice tasks reflected a highly scale of dealing with the tests. The grade point average was 4.5, as 97 % of success. This group mirrored a kind of familiarity with interacting to classroom doing tasks. Therefore, during test taker's involvement in the tests, the instructor observed the positive consequences of the whole process. This meant that, this group was promising and had the challenging spirit.

Group B

In this group, the procedural testing process was the same as compared with other groups. The test takers' attitudes toward learning English and taking tests were studious. A good classroom environment and preparing well to take tests reflected positive outcomes. The grade point average was 4.5, where they achieved around 91 % of success.

Group A and B have recycled the contents of tasks towards mastering and gearing into more promising materials. Therefore, these groups were able to enhance their learning vocabulary.

Group C

The achievement results of this group were expected especially with this undernourished group. Group C lacked motivation and background knowledge. The instructor of this group tried to give feedback to these participants in addition to providing them with a written feedback report in order to motivate and prepare them to increase their participation. The grade point average was 1.00, and their percentage achievements of success reached 53 %.

However, it was observed that, this group needs much effort to raise their achievement. Oral, written report feedback, using audio or visual aids, and breaking down the tasks into bits etc. in a continuous way would turn positively on student's attitudes and pave the way for furthering their vocabulary.

Group D

This group was also crippling just like the previous one. Their achievements were not satisfactory as compared with the first two groups. It seemed that this group had not been exposed to good discipline.

The participants' results were similar to group C; they got 1.00 as an average scale of doing the exercises. 55% of students were successful; therefore, the strategy of the instructor was to keep giving ongoing oral, written, and other styles of feedback in order to fix student's educational process.

The three given tasks were based on selection strategy. It was preferred to start with easy techniques of receptive vocabulary so as to slide down to engaging students for more productive tasks.

Participants' Views on Multiple Choices:

Groups A and B

They explained that they were excited about taking part in this experience, which reflected the positive environment of their instructional classes. They were zealous groups as they spent over one hour assigned to study English at home. As mentioned earlier, these two groups of female participants were eager to explore and perceive knowledge. The affective filter of these groups was very low.

Groups C and D

The students of these groups explained that they were not bored because the instructional conditions were totally different as compared with previous classical styles. This reflected that their affective filters were already fluctuated affecting on their results. Most of them spent less than thirty minutes to study English materials.

They reacted positively as the instructors presented them the oral and written or other useful kind of feedback in order to increase their mastery of subjects' matters. Three students from group C hesitated to answer last questions, but they commented that teachers were used to loading students with the burden of homework.

These impressions reflected the lack of motivation and a shortage of positive instructional environment in the classroom. It is believed that the instructor's ongoing feedback and a good sufficient motivation would break student's stagnancy down.

4.1.2 Odd One Out

This technique required the participants only to identify the item that did not belong to the given material. Their task was to tick or circle the answer. This task would be supportive and a kind to motivating for groups, especially groups C and D to engage or prepare them for more productive activities (See Appendix 3).

Group A

This group was doing and progressing well. In this activity, the test takers proved that they had the ability to discriminate the word which does not go with the others. They did the task with full achievement. The average scale was 4.5; they passed the test successfully with 100 %, which reflected their mastery of how to learn and deal with vocabulary items in a good manner.

The female tested group explained that their task was good and they were confident to deal with the test items. It was an opportunity to learn more. Here, the instructor commented that she expected the same positive outcomes of the next tasks.

Group B

They were exposed to the same test and their achievement was really good. Their total average of achievements was 4.5, where 93% of students passed the test successfully. It was seen that groups A and B processed the testing items with no obstacles whether in participating actively in the testing or doing the homework activities.

This was not a final judgment until checking the outcomes in the next activities which required using productive vocabularies in different task areas or how to employ them in the context comprehensively. But in general perspective of views, the female groups as compared with male ones were more active and gave the instructors self-confident impressions.

Group C

According to the type of task, it was not a surprise with this group that they turned up with an achievement rate of 60%. The grade point average was 2.00. In my opinion, we postponed to engage participants with tasks that required using productive vocabulary knowledge to improve their level of mastery gradually.

Group D

This group fluctuated in its performance; ten of fifteen participants passed the test. Their mark scale average was 2.00, and the percentage level was 61%. However,

they reported that they had time to work and cooperate with each other to enhance and develop their vocabulary learning strategy on the succeeding days.

Participants' Views on Odd One Out:

In order to investigate the impact of the methods and their effective results on the tested groups, the assigned tasks were practical and economical to students in terms of scoring, administrating, and evaluating.

Moreover, the activities were free from difficulties. But, they were designed to evoke student's thinking skills. There were advantages that reincarnated into the positive results of groups A and B and a good start for the undernourished groups C and D. In general, students' attitudes toward the instructional environment were promising.

4.1.3 Matching

In this task, the matching items were presented to the test takers with two columns of information. The instructor asked the participants to study a specific reading passage as homework. Then, on the next day the four tested groups had a test. They were asked to match nine items from the first column with the ones in the second column to form compound nouns (see Appendix 4).

Group A

This active group was cooperated with teacher's instructions. The average scale in matching test items was 4.5, where fourteen of tested students had done well. It seemed that their instructor was fully satisfied with such ongoing excellent achievement. The female students were following and responding to instructional materials comprehensively.

Group B

This group was progressing well, and the participants had a good sense with engaging in classroom instruction. They kept their level of good results. The grade point average was 4.00; this reflected that 87% of students passed the test successfully. This group in general was stepping steadily in terms of the all targeted vocabularies being taught in the classroom.

Group C

It seemed that this group started to interact positively. It was decided to expose the groups C and D to various vocabulary materials such as textual or visual ones. These steps were to increase their educational level as well as to modify their behaviors.

Such procedures required certain amount of extra time to create a positive classroom environment and have the participants at that in positive attitudes toward learning. In matching items, their average scale was 2.50, where the percentage of success was 67%. In this term, we considered this result was satisfactory and good pointer as compared with previous bad ones.

Group D

The result of this group was expected in this type of test. The test takers' average points were similar to group C. Five of students of this group did not prepare well for the test. That was due to the continuous efforts of both parties (intensive instruction and good planning).

Participants' Views on Matching:

In dealing with these exercises, the materials being taught were clear for the participants. They received the instructional materials and linked them with the target vocabulary items. The test takers had prepared their homework (reading a passage) without knowing that the vocabulary of this passage would be employed as a matching test.

So, it can be said that they were progressing well, except for some problematic points of groups C and D. On the other hand, groups A and B results reflected the amount of time they spent to study the required materials. In general, the washback effect was clear. As a result of the lowering of the affective filter of the tested groups, we decided to replace the written questionnaires into oral ones as a kind of interview so as to listen and interact with students openly.

4.1.4 Sentence Completion

In achieving these task, test takers had to complete the sentences in accordance with the sense of given words or they had to create their own sentences (see Appendix 5).

The focus in this section was not on spelling or mechanics of the tested items, but other aspects of language like grammaticality and lexical meaning of vocabularies. The grading of this task was based on subjective criteria. This was a good opportunity to observe the participants' productive vocabulary range.

After the completion of the test activity, we planned to check out participants' attitudes toward the two given exercises. In the first task, the test setter offered the participants incomplete sentences including words that were to be tested whereas in the second part, the tested groups had to create their own sentences with given words.

Group A

Some participants of the group dealing with the first task reflected a kind of a rush in answering the items. For example, in the first sentence, they produced the following pattern; *the baby is hungry, so please feed her.*

In the second task the first item was done as a sample in order not to make the test takers focus on the target word as an important component in constructing the sentence. We checked out some good answers as *My parents always eat sea foods, I meet my friend Suha in the café.*

The average grade scale was 3.50, where 77% of success was seen on both tests. One of the participants created the following sentence, *I don't like eating out.* Generally, the shortage of vocabulary decreases the performance of this group, but this was a general phenomenon.

Group B

Regarding this group who was exposed to the same instructional materials it is possible to say that they made the same grammatical mistakes, but the focus was to

employ the right vocabulary items. One of student answers of item 4 in the first task was, *we all like free times*. This kind of creativity was appreciated.

The percentage of success of this group was around 70% to get 3.00 of grade point average. What was important here was to find practical benefits of instructing and modeling techniques to the tested groups.

Group C

With this group expectations were not very high, but it was considered that the average scale result was as a good pointer as compared to the types of activities which required them to comprehend and to create their own responses.

However, they enjoyed the new tasks and they had many regular mistakes due to being in a rush to answer or not having the ability to recognize. For example, one student confused between *hungry to angry*, stated that, *the baby is hungry, so please play with him*. The idea may be accepted.

This group got a 50% rate of achievement in the end; the average mark scale was 1.00. Those results were satisfactory because it seemed that there was improvement in the participants' performances. The promising point here was the slow progress in the productive tasks. The instructor of this group gave candies to them to reinforce their behavior.

Group D

The participants of this group stated that it was a good idea to create their responses. The exercises were a kind of open ended activity which helped them to write their responses. The tested groups were in need of being in a challenging environment. They made many mistakes just like their friends. For instance, one of the students created the sentence, *'I drink café'*.

However, after that, they had better achievement rates in the next techniques. The grade point average of group D was 1.00, where 57% of students passed the test successfully. The current pointers of groups C and D were satisfactory, regardless of

structural and spelling mistakes, but they understand the meaning of the tests components to base a good strategy later on.

Analysis of interview with tested groups

As it was mentioned previously, an oral interview was preferred with the participants. It was just like a mini workshop to exchange thoughts. Here, we put our considerations of some issues into action. Firstly, some students through answering the written questionnaire might answer randomly or carelessly.

Secondly, having a talk with students was aimed to create a more positive environment and to improve the classroom instructional conditions. The axis of the interview focused on the participants' attitudes of vocabulary instruction as well as their interpretations of the classroom discipline process.

The tested groups were enthusiastic to share their own ideas. They stated that the vocabulary instruction sessions were enjoyable because the instructors presented many vocabulary games like crosswords where the students had to find their names or certain words. The classroom conditions reflected friendliness where some students of groups C and D reported that they did not use time wisely; another one said that his inadequate attention to instructional affairs affected his educational and testing state.

Others stated that their lack of familiarity with studying second language vocabulary decreased their ability to handle the target language. But, groups C and D explained that the new instructional strategy and classroom atmosphere changed their attitudes and behaviors. In general, the participants' responses to the aforementioned axis indicated that the whole process was considered as an effective one by the majority of the participants.

As far to the last two exercises (incomplete or writing sentences), the first task was easier than the other. The students achieved the first one with different answers. Their scores were totally varied, but they were engaged in dealing with the aforementioned tasks. It was a satisfactory thing due to their ability to create their own responses which were imaginative and promising a lot.

The tested groups were able to achieve and meet the challenge to find suitable answers depending on their logical thinking and imagination. Some participants provided us with excellent answers while others tried to express themselves. It was important to realize that the students achieved the tasks comprehensively.

It can be inferred that these techniques were very effective and suitable. The best evidence reincarnated into creating positive backwash, administrating the tests practically, activating students' background knowledge of vocabulary, and evoking participant's mentality and imagination, and achievability of these tasks.

4.1.5 Cloze-test

In this method, the participants were exposed to a scientific instructional topic from their coursebook, Iraq Opportunities. The task was to prepare the text to take an exam later. In this passage "Telephone History", thirteen words were chosen and were erased to be replaced with correct vocabulary. The words were numbered randomly and it was aimed to see whether the participants would be able to recognize or recall the vocabulary items that they had already learned and use them in a logical manner in a context or not (see Appendix 6).

The passage included many important words. This test was considered as a challenge to the instructional materials of the instructors and to the target groups to prove their competence in dealing with the technique.

Groups A and B

These two groups expressed their enjoyment in dealing with this test. It was thought that this positive attitude was due to the friendly classroom environment as well as the systematic method of instruction. In addition, the time to take this type of test was clearly suitable.

However, both groups achieved the test successfully, where fourteen of students from group A passed the test. Seven of test takers got full marks. The average scale of this group was 4.5 as 93% of achievement. These outcomes reflected the stability in their progressive educational career.

At the same time group B proved that the test takers had a good ability to digest the vocabulary instructional materials. They achieved the task successfully, where thirteen of testees passed the test in good outcomes. The grade point average was 4.00. The percentage of success reached 87%.

Groups C and D

In these groups, a good participation came to the prominence. Their results for the first time increased as 73% of success to paralleling the other good groups. In group (C), eleven of testees passed the test. Their scores leveled between (7 to 11 marks).

This was considered as a positive pointer and satisfactory one. The grade point average of group C was 3.00. On the other hand, group D was doing well just like the others in order to step forward. Here, twelve of the test takers completed the exercises successfully. The average scale of this group was 4.00, as 80% of good achievement.

It seemed that holding a comparison among these techniques which required students to use receptive or productive skills was important. In the true/false items and cloze test, we were really satisfied with the participants' outcomes.

They engaged with the material given and employed their background knowledge in the context in a good manner. They used the required vocabularies with an effective strategy. Moreover, the preparation of these techniques and correction were fairly quick and economical.

Analysis of interview with tested groups

The tested groups got more interested in the learning process, as their vocabulary knowledge increased and adopted the strategy of modeling the vocabulary test tasks. These tasks were achievable. In addition, the instructors were encouraging the participants to modify their efforts by providing them continuous feedback. Therefore, It was concluded that our students could make use of such techniques practically. They really needed to be motivated or guided in order to retain or digest the vocabularies being taught. Moreover, the testees believed that the techniques could enhance and develop their vocabulary learning perspectives in the future.

The test takers after doing these tests expressed their approval and especially the friendly classroom environment. They accepted the instructors as friends to support them so as to improve their educational state, especially with groups C and D, whose attitudes changed positively and moved steadily day by day.

4.1.6 Dictation

This method was considered as an effective tool which can be employed to assess language skills. This useful method can gauge the students' weaknesses in phonological and lexical field.

It was decided to have two applications of two kinds of dictations; the first one was the standard mode and the second was called paused dictation. The assigned materials of the tests were taken from the participants' coursebook and the length was shortened to 100 words.

We selected the dictated texts materials which included the vocabularies that student's had already been exposed to or had done. Here, the testees would hear a stream of sounds where they had to decode them into a successive of words written on paper.

Standard dictation

In this type of dictation, we planned to test spelling and student's ability to write down a text. A text was dictated three times with different modes. First, the tester read the text in a normal speed in order to enable test takers to decode the details (listen for gist). The next reading was done by breaking the dictated text into bursts or units (sentence by sentence), to store the details (Appendix 7).

The third reading, the tester read the whole text once again with normal speed and without breaking where the students had to re-encode the text. Then, the testees had several minutes to check the dictation themselves. The level of difficulty in the text was average.

The evaluation of dictation was not an easy task as there were many types of spelling mistakes, and dropping out words or phrases. So, scoring of this brand bore a problematic environment. Here, Mousavi's (2009) scoring example was adopted.

Groups A and B

In all tested groups, there were many regular and expected mistakes, such as wrong forms (distortion), errors of deletions or insertions, and spelling ones. As a result, correcting the dictated text was done according to the criteria followed. But it required a great deal of time to mark because the answers sheets had to be checked twice each and exchanged by two instructors together.

The test takers' common mistakes varied from paper to paper; these mistakes reflected some problem with misspelling like confusing them with words that carry the same pronunciation as in *plain* to *plane* or *son* to *sun* etc. The average mark of group A was 4.00; it was a good result just like group B, whose grade point average was 3.50.

Groups C and D

These two groups made many mistakes as compared to previous groups. But, generally their achievement was considered as satisfactory in terms with their crippling start. The average point grade of group C was 2.50, ten of students passed the test successfully.

On the other hand group D achieved the same grade level which reflected the parallel background knowledge.

Paused dictation

In this activity the test takers were given a passage with some deletions (gapped text). The task was to fill the gaps with the suitable word. There were twelve gaps in the text (Appendix 8).

The same procedural action was followed to deal with this task. The erased words in the test were exposed to the participants over the instructional sessions. This can be considered as a good opportunity to them in order to try their vocabulary retention and ability to spell them correctly. If there would be some spelling mistakes, it was decided that those participants should be given half mark just to motivate them.

Most of the dictated words were nouns and verbs. In general, this task was easier and faster to correct, when compared to previous one.

Group A

The average mark of this group was 4.00 where the test takers of achievement reached 83%.

Group B

The percentage level of success in this group was satisfactory. It was 75% which reflected a kind of stability with their vocabulary knowledge. The grade point average was 3.50.

Group C

The grade point average was 3.00 where 70% of students did the test successfully. It was a supportive pointer with this group due to good positive outcomes and fruitful interaction with the whole process.

Group D

This group also did in a satisfactory way to get 65% of success. The grade point average was paralleled to group (C), where they achieved 2.50. We were satisfied with these two groups because if we compare their negative outcomes at the first start which required to make only a guess of the tasks given. Consequently, they produced and spelled vocabulary items in a good manner. Their strategic scope changed gradually.

Analysis of interview with tested groups

The achievability of these tasks reflected through the good satisfactory outcomes of tested groups. Moreover, this method evoked student's ability to provide us with details of dealing with spelling as well as the capacity to recall details. The testees decoded the tasks comprehensively.

The gapped dictation was more effective than standard one because students reacted, positively towards the gapped one. Moreover, they did not find any obstacles in dealing with it. Additionally, teachers did not require assigning much time to administer the task.

What was important for me that, we did not recognize any kind of poor backwash doing the activities. The candidates were progressing gradually and reacted positively. The test takers stated that, it was interesting to deal with the whole process.

4.2 Discussion of the Results

The techniques which were tried out were to investigate the strengths and weaknesses of these methods. In general perspectives and according to student's outcomes (score results and questionnaires), these methods were useful and effective for both the participants and the instructors.

The test takers were able to employ vocabulary items in various linguistic situations. During the process of holding the exercises and student's response to written or oral questionnaires, we could identify in which areas they had attitudinal problems of learning English and doing tests.

In terms of making use of these techniques, it was clear that multiple-choice, odd one out and matching would be placed into the category of less productive methods. The test takers had to select the answer from the options. They did not produce anything actively, but these methods can tell us to which extent students mastered the material being examined.

On the other side, the method of sentence completion was fully appreciated. Its use revealed a lot about the participants' strengths and weaknesses of their vocabulary background knowledge. Moreover, how they can employ these words in a context. These methods were categorized as productive ones which reflected the participants' knowledge characteristics.

In the standard and paused dictation test, students had to decode the vocal sounds into a written mode, they did many spelling mistakes, but they can identify their mistakes to deal with. This strategy was also adopted in dealing with cloze test, where the participants found some troubles with.

CHAPTER 5

CONCLUSION AND SUGGESTIONS

5.0 PRESENTATION

In this chapter, the major findings would be discussed of the research. Here, the scope of the discussion will cover two items, conclusion of the study as well as suggestions for; learners, teachers, materials, coursebook designers, and for further studies. These points would be presented individually to serve the main aim of the study.

5.1 CONCLUSION

Testing vocabulary has taken its important effective part in the scene of educational measurement. This field in the last two decades has taken its importance. Researchers and language-teaching specialists in the field of ELT, with a specific interest in testing or learning vocabularies have a progressive need for finding the effective assessments tools.

The current study thereby was designed to make a change in the humdrum of classical methods in testing vocabulary. For this reason, the perspective scope of the research was based on the theoretical part of testing which paved the way for investigating the multifaceted methods that deal with vocabulary.

However, many of testing vocabulary methods has been applied, where many suitable test tasks were designed in the formal testing activities. These methods reflected positive educational outcomes both for learners and teachers. In terms of this vision, “methodologists, thereby, are expected to modify their point of views, concurrent with the developments in the present scenario of ongoing globalization, and change the ways they construct and prescribe their classroom activities, techniques, and instructional innovations” (Hosseini, 2011, p.295).

In this sense, this referred to the effectiveness of adopting these methods and the validity of modeling them in a linguistic environment of a target vocabulary. These methods reflected on participants' interest in taking the tests.

Therefore, these methods yielded rich results. Here, the participants found themselves in challenging situations and they were motivated to take an active part in the different stages of the study. The best evidence of the success was reincarnated into the good results of the tested groups, especially with groups C and D, as well as the assistant teachers who were happy with the results the participants got. Ellis and Tod (2015) believe that “from a behavior for learning perspective, the importance consideration is the compatibility of any framework with the principle of at least protecting and when possible, enhancing the three relationships (with self, with others and with the curriculum) and fostering the development of positive learning behaviors. Any practice where a potential detrimental effect on these relationships for learning could reasonably be predicted should be avoided” (p.101).

Making use of these results is crucial because they reflect the same goals that direct the instructional activities, assessment; instructional planning of many cases and serves as a measure of teaching or testing effectiveness. In that way it would be possible to diagnose and build our strategic vision for treatment requirements in the field of testing vocabulary.

Consequently, it was found that good assessment was an integral part of good instruction. Both good instruction and testing need to reflect the goals of students' assessment to measure classroom objectives.

5.2 SUGGESTIONS

Students' vocabulary background knowledge plays a critical role in learners' academic success. Developing strategies to deal with vocabulary will be a guide in creating or adapting effective methods in language teaching and testing concerns. In the findings of the study, modeling of vocabulary testing methods and adopting to build a pedagogical strategy on a good theoretical and empirical investigation and to ensure the effectiveness of each method will support and guide teachers and learners to reach the coursebook objectives.

The following suggestions are provided to further and enhance the process of teaching English.

5.2.1 For Learners

Students are considered as the target groups of an educational process. They can be considered as a fertile source of language material input. Therefore, the learner should take her/his an active role in terms of participation in scholastic activities or interacting with their instructors and with classroom colleagues. In this sense, there are some important perspectives of students' roles that should be engaged with adopted:

- Students should take responsibility action in dealing with the whole instructional or testing affairs.
- They have to cooperate and interact with other students and they should also share each other's experiences, feelings, and knowledge.
- They have to assign a specific period of time to reading some extra books in accordance their educational levels.
- They need to build their own strategies in learning and keeping active vocabulary.
- They have to take an active part in classroom activities and testing tasks.
- They should make use of instructional materials inside and outside of school.
- They have to consider themselves as partners in the educational training program, but not as listeners or passive receivers.
- They should make use of language components and try to employ them in their English learning process.
- They have to appreciate the teachers' or classmates' feedback and make full use of it.

5.2.2 For Teachers

Teachers should play an effective role in the development of second language skills. They establish the instructional goals, develop curricula, conduct classes, and evaluate students' achievement. As an organizer, teacher should be responsible to help students to overcome their learning problems.

The new philosophy which had come into the prominence due to the revolutionary theories of ELT recognized her/his role as a friend to learners. In addition, teachers should have a good comprehension about the theory of testing and identify the

most effective methods to deal with vocabulary. Moreover, they should have sufficient knowledge of the basic principles of applied linguistics and be familiar with the updated affairs of ELT. Here, teachers can make use of accomplishing these strategic instructions in order to stimulate learners to become more successful in adopting more productive learning strategic styles.

Therefore, they have to develop learners' language skills through presenting real life and authentic materials, as well as providing the class with variable activities like formal and informal tasks, take active part in oral drills, and use vocabulary items in a context.

Teachers should design and manage the classroom activities using visual aids, gestures, relia, and authentic real objects. In a nutshell, teachers should be reflective because 'good teachers are always good testers'.

5.2.3 For Material and Curriculum Designer

Materials in the coursebook should be described and organized well enough in order to help teachers in making use of them with no confusions or obstacles. Many factors have to be engaged to the course book materials;

- 1- "Language is functional and must be contextualized,
- 2- Language development requires learner engagement in purposeful use of language,
- 3- The language use should be realistic and authentic,
- 4- Classroom materials will usefully seek to include an audiovisual component,
- 5- Learners need to develop the ability to deal with written as well as spoken genres,
- 6- Effective teaching materials foster learner autonomy,
- 7- Materials need to be flexible enough to allow for individual and contextual differences,
- 8- Learning needs to engage learners both affectively and cognitively" (Richards and Renandya, 2002, p.67).

These perspective, were supposed helping learners in order to develop their skills to use and master foreign language. So, it is clear that learners would cooperate actively when they involve in a meaningful and authentic purposeful activities.

Participation of learners in such activities is important. Therefore, applying the different methods of teaching, and the findings of the coursebook Iraq Opportunities, which was analyzed, would contribute in respect to the coursebook designers.

In accordance with the application of new curriculum of English language, teachers need to be familiarized with the methods for vocabulary learning strategies training. There is an urgent need to recognize the importance of training teachers.

5.2.4 Implications for Further Research

This research study was conducted to explore the most effective methods of testing vocabulary in English Language. So as to confirm the outcomes of the study and other previous ones, more research could be done. Further research in this area is required to investigate and penetrate deeply to explore or to create more effective methods to test vocabulary such as the role of modern technology in developing productive or receptive vocabulary skills, regarding the classroom-based and computer-assisted vocabulary instruction.

5.3 Summary of the Chapter

In this chapter, the conclusions of the research as well as the suggestions were presented. The conclusion part explained the importance of adopting the most effective methods to deal with testing vocabulary

The suggestions' part focused on the possibility to make use of this study to further teachers' experiences in dealing with testing vocabulary in Iraqi high public schools.

Finally, some suggestions were presented for further studies which might help a researcher in this field to take a clear vision of the findings of this study.

REFERENCES

- Albrechtsen, D., Haastrup K., & Henriksen, B. (2008). *Vocabulary and Writing in a First and Second Language* (1st.ed.). BRT: Alister Cumming.
- Alderson, J Charles & Banerjee Jayanti. (2001). *Language Testing and Assessment*. Cambridge: Cambridge University Press.
- Birjandi, P., Mosallanejad, P. (2010). *An Overview of Testing and Assessment*. Asfahan: Sepahan Press.
- Birjandi, P., Bagheridous, E., & Mossalanejad, P. (2004). *Language Testing* (2nd.ed.) Tehran: Shahid Mahdavi Publications.
- Brown, J. Dean. (2005). *Testing in Language Programs*. New York: McGraw-Hill Board.
- Brown, J. Dean & Hudson, T. (2003). *Criterion-referenced Language Testing*. Cambridge: Cambridge University Press.
- Brown, J. Dean. (1995). *The Element of Language Curriculum*. Boston: Heinle &Heinle press
- Brown, H. Douglas. (2004). *Language Assessment Principles and Classroom Practices*. Longman: Pearson Education Press.
- Burgmeier, A.(1991). *Lexis*. U.K.: Prentice-Hall, Inc.
- Bogaards, P& Laufer, B. (2004).*Vocabulary in a Second Language* (vol.10). Amesterdam: John Benjamin's publishing Company.
- Bachman Lyle F. & Palmer Adrian S.(2000). *Language Testing in Practice* (3rd.ed.). Oxford: Oxford University Press.
- Bachman Lyle F. & Palmer Adrian S.(2012). *Language Assessment in Practice* (1st.ed). Oxford: Oxford University Press.
- Bachman Lyle F. (1995). *Fundamental Consideration in Language Testing* (3rd.ed.). Oxford: Oxford University Press.
- Baker, D. (1989). *Language Testing* (1st.ed.).Cambridge: Cambridge University Press.

- Barnes, A., Hines, J. & Weldon, J.(1996) Have Fun with Vocabulary. U.K.: Penguin Books.
- Birjandi, P., Bagheridoust, E., & Mossalanejad, P. (2004). Language Testing (2nd.ed.) Tehran: Shahid Mahdavi Publications.
- Ellis Simon and Tod Janet. (2015). Promoting Behavior for Learning in the Classroom. NY: Rutledge (Taylor & Francis Group).
- Cheng, L. Yoshinori, W., & Andy, C. (2004). Washback in Language Testing. London: Lawrence Erlbaum Association Press.
- Celce-Murcia, M., Brinton Donna M., & Snow, M. Ann. (2014). Teaching English as a Second or Foreign Language (4th.ed.). Boston: Heinle Cengage Learning Press.
- Coombe, C., Folse, K., & Hubley, N. (2007). Assessing English Language Learners. U.S.A.: University of Michigan Press.
- Cook, V. (2008). Second Language Learning and Language Teaching (4th.ed.). UK: Hodder Education Press.
- Davies, A. (1990). Principles of Language Testing (1st.ed). U.S.A.: Blackwell Press.
- Daller, H., Milton J. & Daller J. Treffers. (2007). Modeling and Assessing Vocabulary Knowledge (1st.edition). Cambridge: Cambridge University Press.
- Doff, A. & Jones, C. (2007). Language Links. Cambridge: Cambridge University Press.
- Elder, C., Brown A., Grove N., Hill K., Iwashita N., Lumley T., McNamara T. & O'Loughlin K. (1998). Studies in Language Testing. BRT: Cambridge University Press
- Fulcher, G. (2012). Practical Language Testing. UK: Hodder Education Press.
- Fulcher, G. & Davidson, F. (2007). Language Testing and Assessment (1st.ed.). London: Routledge Press.
- Genesee F. & Upshur John A.(1996). Classroom-Based Evaluation in Second Language Education (3rd.ed). Cambridge: Cambridge University Press.
- Hosseini, S. Mohammad Hassan. (2011). Beyond the Present Methods and Approaches to ELT/ Education: The Crucial Need for a Radical Reform. Tehran: Jungle publication Press.
- Harris, M. & McCann, P. (1994). Assessment. Scotland: Heinemann Press (Oxford).

- Hiebert, Elfrieda, H. & Kamil, Michael, L. (2005). *Teaching and Learning Vocabulary*. London Lawrence Erlbaum Associates Press.
- Hughes, A. (2003). *Testing for Language Teachers* (2nd.ed.). UK: Cambridge University Press.
- Heaten, J. B. (1990). *Writing English Language Tests* (3rd.ed.). UK: Longman Press.
- Harris, M., Mower, D. & SikorzySka, A. (2010). *Iraq Opportunities*. London: York Press.
- Huston, S, & Francis, G.(2000). *Pattern Grammar*. Amsterdam: John Benjamin Publishing Press.
- Khodadady, E. (1999). *Multiple choice Items in Testing: Practice, and Theory* (1st.ed.). Iran: Rahman Publications.
- Luoma, S. (2008). *Assessing Speaking* (4th.ed.). UK. Cambridge University Press.
- Mahboob, A. & Knight, N, K. (2010). *Applicable Linguistics*. London: Continuum International Publishing Group.
- McCarthy, M. & O'Dell, F. (2002). *Test your English Vocabulary in Use* (3rd.ed.). UK: Cambridge University Press.
- Mousavi, S. Abbas. (2009). *An Encyclopedic Dictionary of Language Testing* (4th.ed.). Tehran: Rahman Press.
- McNamara, T. (2000). *Language Testing* (2nd.ed.). UK: Oxford University Press.
- McCarthy, M. (2005). *Discourse Analysis for Language Teachers* (16th.ed.). Cambridge: Cambridge University Press, 2005.
- Madsen, Harold, S. (1983). *Techniques in Testing* (1st.ed.). UK: Oxford University Press
- McCarteen, J. (2007). *Teaching Vocabulary* (1st.ed.). London: Cambridge University Press.
- Meyer, Charles, F. (2009). *Introducing English Linguistics* (1st.ed.). UK: Cambridge University Press.
- Nunan, D. (2003). *Practical English Language Teaching* (1st.ed.). New York: McGraw-Hill Press.

- Olsen Amy E. (2010). *Interactive Vocabulary (General Words) (4th.ed.)*. England: Pearson Education Inc.
- Phillips, C. (2002). *Increased your Word Power*. Poland: Altravox Press.
- Penn, J. (2010). *Food around the World (1st.ed.)*. Oxford: Oxford University Press.
- Pikulski John J. & Templeton, S. (2004). *Teaching and Developing Vocabulary: key to Long- Term reading success*. U.S.A.: Houghton Mifflin co.
- Richards, J. & Renandya, W. (2002). *Methodology in Language Teaching (1st.ed.)*. London: Cambridge University Press.
- Read, J. (2000). *Assessing Vocabulary (1st.ed.)*. Cambridge: Cambridge University press.
- Richards, C. J. and Rodgers, T. R. (2001). *Approaches and Methods in Language Teaching*. Cambridge : Cambridge University Press.
- Shin, J. Kang & Crandall JoAnn (Jodi). (2014). *Teaching Young Learners English*. Boston: Heinle Press.
- Singleton, D. (2000). *Language and the Lexicon*. Oxford: Oxford University Press.
- Scanlon, J. (2011). *Q: Skills for Success (Listening and Speaking 1)*. UK: Oxford University Press.
- Thornbury, S. (2004). *How to Teach Vocabulary (4th.ed.)*. BRT: Pearson Press.
- Ur, P. (1996). *A course in Language Teaching (1st.ed.)*.UK: Cambridge University Press.
- Weir, C. J. (2005). *Language Testing and Validation (1st.ed.)*. BRT: Palgrave Macmillan Press.
- Willis, D. & Willis, J. (2007). *Doing Task-based Teaching*. London: Oxford University Press.
- Woolfolk Anita E., Winne Philip H. & Perry Nancy E. (2003). *Educational Psychology (2nd.ed.)*. Canada: Pearson Education Press.
- Wintergerst Ann C. & McVeigh Joe. (2011). *Tips For Teaching Culture*. UK: Longman Press.
- Yule, G. (2010). *The Study of Language (4th.ed.)* UK: Cambridge University Press.

LIST OF APPENDICES

Appendix 1: Plan for Unit 4 Cyberspace in Iraq Opportunities English Coursebook 8.....	74
Appendix 2: Multiple-choice.....	75
Appendix 3: Odd one out.....	76
Appendix 4: Matching	77
Appendix 5: Sentence Completion.....	78
Appendix 6: Cloze-test.....	79
Appendix 7: Dictation (standard).....	80
Appendix 8: Dictation (paused).....	81
Appendix 9: Results of multiple-choice items of group (A).....	82
Appendix 10: Results of multiple-choice items of group (B).....	83
Appendix 11: Results of multiple-choice items of group (C).....	84
Appendix 12: Results of multiple-choice items of group (D).....	85
Appendix 13: Results of odd one out of groups (A, B, C, and D).....	86
Appendix 14: Results of matching items of groups (A, B, C, and D).....	87
Appendix 15: Results of sentence completion of groups (A, B, C, and D).....	88
Appendix 16: Results of cloze-tests of groups (A, B, C, and D).....	89
Appendix 17: Results of standard and paused dictation of groups (A, B, C, and D).....	90

Appendix 1

Plan for Unit 4 Cyberspace in Iraq Opportunities English Coursebook 8

4 CYBERSPACE		
Warm-up (p.49)	Vocabulary: science, science fiction, the environment	Listening: predictions
13 Tomorrow's World (pp.50-51)	Vocabulary: Internet words Grammar: Predictions: <i>will</i> and <i>going to</i>	Reading: magazine article
14 Websites (pp.52-53)	Function: telephoning/making suggestions Pronunciation: important words	Listening: radio programme; dialogue Listening Strategies: important words Speaking: going out
15 Virtual Reality (pp.54-55)	Grammar: First and Second Conditionals	Reading: websites; dialogues
16 Virtual Tourism (pp.56-57)	Vocabulary: collocation ("do" and "make") Linking: addition	Reading: tourist website Reading Strategies: matching topics/paragraphs Speaking: planning a weekend Comparing Cultures: New Zealand
Communication Workshop (pp.58-59)	Writing: an Internet page (planning/note-taking) Speaking: planning a holiday resort	Listening: a song
Review (p.60)	Grammar and vocabulary revision	Pronunciation: /r/

Literature Spot: Introduction (pp.61-62); Focus on Poetry (pp.63-69)

Science Topics (pp.70-79)

Writing Help (pp.80-81)

Grammar Summary (pp.82-84)

Mini-Dictionary (pp.85-94)

Appendix 2

Multiple-choice test

Student code # ()

Group (A)

Choose the correct answer from a, b, c or d.

1- When tourists visit a country, it's interesting to see the.....way of local national's life.

A. tradeshunal. B. traditionl. C. tradishonal. **(D) traditional.**

2- The.....of a country usually holds monthly meeting in the capital city.

A. goverment. **(B) government.** C. government. D. governmant.

3- Sydney has already beautiful.....

(A) harbor. B. habour. C. habor. D. harber.

4- Do you have in your country acity like Auckland ?

A. dinemic. B. dainamick. **(C) dynamic.** D. dinamytic.

5- Auckland has some of the best alluring..... in New Zealand.

(A) beaches. B. beches. C. beetches. D. beashes.

Answer :

1 D2 B3 A4 C5 A

Appendix 3

Odd One Out

Student code # ()

Group (b)

.....

Circle the "odd word or expressions out", which doesn't belong in each list).

- 1) ~~1) Hacker.~~ Cybercriminal. Cyberterrorist. Cyberia magazine.
- 2) Ancient civilization. Climatic change. Global warming. Earth's climate.
- 3) Canoe. ~~Textiles.~~ Yacht. Ship.
- 4) Wedding. Carnival. Rock climbing. Festival.
- 5) Bazar. Biscuits. Supermarket. Mall.
- 6) Before. Later. After that. Next

5/6 five

Appendix 4

Matching

Student code # ()

Group (B)

Match words in the boxes to make compound words

Example: solar power

- | | |
|-------------|----------------|
| 1-Solar | A- power. |
| 2-virtual | B- hotel. |
| 3-cyber | C- reality. |
| 4-space | D- mail. |
| 5- luxury | E- program. |
| 6- genetic | F- space. |
| 7- computer | G- fiction. |
| 8-web | H- shuttle. |
| 9-c- | I- engineering |
| 10-science | J- site. |

8
To eight.

Answer:

- | | | |
|----|---|---|
| 1 | A | ✓ |
| 2 | C | ✓ |
| 3 | F | ✓ |
| 4 | H | ✓ |
| 5 | B | ✓ |
| 6 | D | ✓ |
| 7 | E | ✓ |
| 8 | G | ✓ |
| 9 | I | ✓ |
| 10 | J | ✓ |

Appendix 5

Sentence Completion

Student code # (2)

Group (A)

.....
 Finish these sentences in any way you want to complete the sentence.

- 1- The baby is hungry, so please... feed her milk -
- 2- Noor and Layla are... studying Arabic
- 3- My parents always eat... in home
- 4- We all like... not travel in winter
- 5- Some people don't buy... fast food
- 6- Ahmed doesn't... sucess in the examination

Write sentences about what to do this weekend. Make use of the words given.

- 1- (film).....ex. I like to watch a comical film in a movie.
- 2- (sport)... sport is good for people
- 3- (café)... We drinking tea in cafe
- 4- (shopping)... I shopping in the market
- 5- (museum)... we visited the museum last month
- 6- (eating out)... my father don't like eating out

Appendix 6

Cloze-test

Student code # (15)

Group (C)

Fill each space in the text below with a suitable word. (Note that only one word per space is allowed).

- | | | |
|----------------|------------------|----------------|
| 1-Drums. | 7- age. | 13- telephone. |
| 2- inventor. | 8-message. | |
| 3- deaf. | 9-pioneer. | |
| 4- microphone. | 10-born. | |
| 5-communicate. | 11- internet. | |
| 6-optics. | 12- connections. | |

"Telephone History"

The history of the¹³.....begins at the start of human history. Man has always wanted to¹.....from a distance. People have used smoke signals, mirrots, jungle-~~carriers~~ pigeons and semaphore to get a².....from one point to another. But a phone was something new. In some way Francis Bacon predicted the telephone in 1627, however his book *New Utopia* only described a long speaking tube. Areal telephone could not be invented until the electrical².....began. The electrical principles needed to build a telephone were known in 1831, but it wasn't until 1876 that the idea became a reality. Before then, a telephone might have been impossible to even think of.

A³.....in the field of telecommunications, Alexander Graham Bell was¹⁰.....in 1847 in Edinburgh, Scotland. He moved to Ontario and then to the United States, setting in Boston, before beginning his career as an²..... Throughout his life, Bell was interested in the education of⁵.....people. This interest led him to invent the⁴.....and, in 1876, his "electrical speech machine". He set up the first telephone exchange in New Haven, Connecticut. By 1884, long distance¹².....were made between Boston, Massachusetts, and New York city.

Bell imagined great uses of his telephone, but would he ever have imagined telephone lines being used an amazing revolution in the telecommunication industry. Today, deaf people are able to use a special display telephone to communicate. Fiber⁶.....are improving the quality and speed of data transmission. Actually, your ability to access this information relies upon telecommunication technology. Bell's "electrical speech machine" made the⁸.....possible.

Appendix 7

Dictation (Standard)

A.

20/30 Twenty
N.B. clearly your handwriting is poor

Every evening I am sitting at the computer and I say what are we going to have for dinner tonight. My family did some of its shopping on the internet since the supermarket stated online. But my daughter don't like it. She prefer to go to the supermarket with me to buy chocolate while my son loves the way of shopping online. We click on plain biscuit when we want a little. But there are some problems with internet shopping. The supermarket are giving the customers a time when they will bring the internet shopping to the customer's home. The customers have to be at home at the time. All your customers want their internet shopping brought to their home at the weekend or in the early evening say the supermarket.

B.

Shopping on the internet 18/30 eighteen

Every evening I'm sitting down at the computer and I say what are we going to have for dinner tonight. My family did some of its shopping on the internet since the supermarket stated online. But my daughter don't like it. She prefer to go to the supermarket with me to buy chocolate while my son loves the way of shopping online. We click on plain biscuit when we want a little. But there are some problems with internet shopping. It cost 5 dollars every time your shopping was brought on the internet. The supermarket are giving the customers a time when they will bring the internet shopping to the customer's home. The customers have to be at home at the time. All your customers want their internet shopping brought to their home at the weekend or in the early evening say the supermarket.

Appendix 8

Dictation (Paused)

Student code # ()

Group (8)

"Charles Gray"

Charles was a college ~~Prof~~ ~~essor~~ with a huge six-bed room house and a ~~fortune~~ of \$2 million. Today he ~~lives~~ in a small caravan where there is only second-hand ~~fortune~~. There is a small ~~garden~~ outside with a few fruit ~~trees~~. Charles grows some ~~vegetables~~ and a few ~~flowers~~. He ~~gets~~ his clothes and a lot of other things from ~~garage~~ shops. Charles was tired of being a person who had everything, in a world where many ~~people~~ have nothing.

8
12 eight -

Appendix 9

- Multiple-choice Results. (Group A).

Students' Code #.	Drill (1). Total score (10 M.).	Drill (2). Total score (14 M.).	Drill (3). Total score (10 M.).
S1	8	10	9
S2	10	12	9
S3	6	10	9
S4	8	13	8
S5	8	11	9
S6	10	10	8
S7	6	8	9
S8	10	13	10
S9	8	11	9
S10	6	8	7
S11	8	9	10
S12	8	10	8
S13	4	7	7
S14	10	14	10
S15	6	9	8

Appendix 10

Students' exercises results.

- Multiple-choice. (Group B).

Students' Code #.	Drill (1). Total score (10 M.).	Drill (2). Total score (14 M.).	Drill (3). Total score (10 M.).
S1	8	13	8
S2	6	12	7
S3	4	9	9
S4	6	10	8
S5	6	6	4
S6	9	14	7
S7	6	10	6
S8	8	13	9
S9	8	8	8
S10	9	11	10
S11	8	9	9
S12	8	10	10
S13	4	9	7
S14	6	9	8
S15	6	10	9

Appendix 11

Students' exercises results.

- Multiple-choice. (Group C).

Students' Code #.	Drill (1). Total score (10 M.).	Drill (2). Total score (14 M.).	Drill (3). Total score (10 M.).
S1	4	6	8
S2	6	7	4
S3	8	9	9
S4	6	7	7
S5	4	5	4
S6	8	8	6
S7	8	6	4
S8	4	5	5
9	6	9	6
S10	10	7	4
S11	4	9	7
S12	8	12	6
S13	6	9	4
S14	6	6	7
S15	4	8	9

Appendix 12

Students' exercises results.

- Multiple-choice. (Group D).

Students' Code #.	Drill (1). Total score (10 M.).	Drill (2). Total score (14 M.).	Drill (3). Total score (10 M.).
S1	6	10	8
S2	4	5	4
S3	5	6	9
S4	4	6	4
S5	8	9	7
S6	4	8	4
S7	6	10	6
S8	4	6	4
S9	6	8	7
S10	4	5	4
S11	6	7	8
S12	4	6	6
S13	8	13	7
S14	8	8	7
S15	4	5	4

Appendix 13

Students' exercises results.

- Odd one out.

Group (A)	Total score 6M.	Group (B)	Total score 6M.	Group (C)	Total score 6M.	Group (D)	Total score 6M.
S1	4	S1	5	S1	2	S1	5
S2	6	S2	4	S2	5	S2	2
S3	5	S3	4	S3	2	S3	4
S4	4	S4	2	S4	4	S4	6
S5	5	S5	5	S5	4	S5	2
S6	6	S6	4	S6	2	S6	4
S7	6	S7	5	S7	6	S7	5
S8	5	S8	6	S8	5	S8	1
S9	4	S9	6	S9	2	S9	5
S10	3	S10	5	S10	4	S10	2
S11	5	S11	3	S11	2	S11	5
S12	4	S12	4	S12	4	S12	2
S13	5	S13	4	S13	2	S13	4
S14	5	S14	3	S14	5	S14	5
S15	6	S15	4	S15	4	S15	2

Appendix 14

Students' exercises results.

- Matching items.

Group (A)	Total score 10M. Drill # 1	Group (B)	Total score 10M. Drill # 1	Group (C)	Total score 10M. Drill # 1	Total score 5M. Drill # 2	Group (D)	Total score 10M. Drill # 1	Total score 5M. Drill #2
S1	8	S1	7	S1	4	4	S1	4	2
S2	10	S2	6	S2	7	5	S2	7	4
S3	7	S3	4	S3	4	4	S3	3	5
S4	6	S4	8	S4	6	5	S4	8	4
S5	7	S5	7	S5	4	2	S5	7	4
S6	6	S6	6	S6	4	5	S6	4	2
S7	5	S7	5	S7	8	4	S7	6	3
S8	8	S8	9	S8	7	4	S8	4	2
S9	9	S9	10	S9	4	3	S9	10	5
S10	10	S10	9	S10	9	4	S10	8	4
S11	10	S11	10	S11	3	5	S11	3	4
S12	6	S12	8	S12	4	4	S12	7	3
S13	4	S13	7	S13	10	4	S13	4	4
S14	7	S14	6	S14	3	2	S14	9	5
S15	8	S15	4	S15	8	5	S15	4	2

Appendix 15

Students' exercises results.

- Sentence completion

Group (A)	Total score 12M. Drill # 1	Total score 15M. Drill # 2	Group (B)	Total score 12M. Drill # 1	Total score 15M. Drill # 2	Group (C)	Total score 12M. Drill # 1	Total score 15M. Drill # 2	Group (D)	Total score 12M. Drill # 1	Total score 15M. Drill # 2
S1	10	11	S1	10	11	S1	5	7	S1	5	7
S2	9	10	S2	5	7	S2	9	10	S2	6	5
S3	5	6	S3	12	11	S3	4	7	S3	5	6
S4	8	10	S4	11	12	S4	8	11	S4	10	11
S5	9	9	S5	5	6	S5	4	6	S5	9	12
S6	10	12	S6	9	10	S6	6	9	S6	8	10
S7	12	14	S7	8	7	S7	5	6	S7	4	7
S8	11	15	S8	9	10	S8	8	10	S8	9	11
S9	9	10	S9	5	7	S9	5	7	S9	10	12
S10	5	7	S10	7	9	S10	10	9	S10	5	7
S11	10	11	S11	10	10	S11	4	6	S11	8	10
S12	9	9	S12	12	9	S12	11	8	S12	7	9
S13	7	6	S13	4	6	S13	9	7	S13	5	6
S14	5	8	S14	10	8	S14	10	8	S14	9	11
S15	8	7	S15	9	8	S15	5	7	S15	4	6

Appendix 16

Students' exercises results.

- Cloze-test.

Group (A)	Total score 13M.	Group (B)	Total score 13M.	Group (C)	Total score 13M.	Group (D)	Total score 13M.
S1	10	S1	9	S1	5	S1	9
S2	11	S2	6	S2	10	S2	7
S3	8	S3	10	S3	6	S3	8
S4	9	S4	11	S4	8	S4	6
S5	12	S5	8	S5	7	S5	7
S6	13	S6	7	S6	7	S6	9
S7	10	S7	9	S7	8	S7	10
S8	9	S8	12	S8	9	S8	9
S9	8	S9	13	S9	9	S9	5
S10	11	S10	11	S10	10	S10	9
S11	10	S11	8	S11	9	S11	8
S12	9	S12	7	S12	5	S12	7
S13	8	S13	6	S13	9	S13	7
S14	6	S14	10	S14	8	S14	7
S15	10	S15	9	S15	5	S15	5

Appendix 17

Students' exercises results.

- Standard dictation #1.
- Paused dictation #2.

Group (A)	Total score 30M. Drill # 1	Total score 12M. Drill # 2	Group (B)	Total score 30M. Drill # 1	Total score 12M. Drill # 2	Group (C)	Total score 30M. Drill # 1	Total score 12M. Drill # 2	Group (D)	Total score 30M. Drill # 1	Total score 12M. Drill # 2
S1	17	6	S1	15	5	S1	11	5	S1	14	5
S2	18	7	S2	16	6	S2	15	6	S2	15	6
S3	20	8	S3	13	6	S3	16	7	S3	15	6
S4	22	7	S4	15	7	S4	17	7	S4	16	5
S5	18	6	S5	17	8	S5	16	6	S5	13	4
S6	13	5	S6	19	7	S6	18	7	S6	10	7
S7	19	6	S7	17	6	S7	17	8	S7	16	8
S8	23	7	S8	18	6	S8	13	7	S8	17	7
S9	20	8	S9	16	7	S9	11	5	S9	15	5
S10	16	7	S10	20	6	S10	15	4	S10	11	7
S11	14	5	S11	17	8	S11	15	6	S11	18	6
S12	17	6	S12	13	4	S12	16	6	S12	17	6
S13	18	6	S13	18	9	S13	12	7	S13	16	7
S14	13	6	S14	14	5	S14	13	5	S14	16	8
S15	17	7	S15	13	5	S15	15	5	S15	13	5

ÖZGEÇMİŞ

Kişisel Bilgiler

Adı Soyadı : Raad Thaidan

Doğum Yeri ve Tarih : Iraq 1974

Eğitim Durumu

Lisans Öğrenimi : English Language Teaching Department- Dhi-Qar University

Yüksek Lisans Öğrenimi : English Language Teaching Department- Ufuk University

Bildiği Yabancı Diller : English

Bilimsel Faaliyetleri

İş Deneyimi

Stajlar

Projeler

Çalıştığı Kurumlar : English Language Institution

İletişim

E-Posta Adresi : Radhamod@ymail.com

Tarih : 07.01. 2015