

**GAZI UNIVERSITY
INSTITUTE OF EDUCATIONAL SCIENCES
DEPARTMENT OF ENGLISH LANGUAGE TEACHING**

145168

**VOCABULARY TESTING PROBLEMS
WITH SPECIFIC REFERENCE TO
MACEDONIAN STATE SCHOOLS**

145168

M.A. THESIS

By

Rufat OSMANI

Supervisor

Assist. Prof. Dr. Abdullah ERTAŞ

Ankara 2004

Eğitim Bilimleri Enstitüsü Müdürlüğü'ne

Rufat Osmani'ye ait **“Vocabulary Testing Problems with Specific Reference to Macedonian State Schools”** adlı çalışma jürimiz tarafından İngiliz Dili Eğitimi Anabilim Dalında YÜKSEK LİSANS TEZİ olarak kabul edilmiştir.

Başkan:

.....
Akademik Ünvanı, Adı Soyadı

Üye:

.....
Akademik Ünvanı, Adı Soyadı

Üye:

.....
Akademik Ünvanı, Adı Soyadı

ACKNOWLEDGEMENT

I would like to express my greatest thankfulness to Assist. Prof. Dr. Abdullah Ertaş, my thesis mentor for his great help and never ending patience during the preparation of my thesis. I feel providential to have worked with him.

I would also like to express my special thanks to the professors who taught me different subjects during MA studies at Gazi University: Prof. Dr. Aydan Ersöz, Prof. Dr. Abdülvahit Çakır, Assist. Prof. Dr. Paşa Tefik Cephe, Assist. Prof. Dr. Bena Gül Peker, Prof. Dr. İskender Sarıgöz.

I would like to thank the teachers of different high schools in Macedonia, participating in this study.

I wish to thank all my colleagues at South East European University for their great support and encouragement during my work preparation of this thesis.

Finally, my greatest appreciation to my family for their infinite endurance, encouragement and support and especially to my son Argjend for his little but meaningful motivation he gave me during this study.

ÖZET

‘‘ Makedonya’daki Yabancı Dilde Kelime Testlerin Hazırlamasında Karşılaşılan Problemler’’

Bu çalışmanın amacı, Makedonya’daki İngilizce öğretmenlerin Yabancı Dil sınavlarında kullandıkları kelime testlerini uygun bir şekilde hazırlayıp hazırlamadıklarını öğrenmektir. Devlet Orta Okullarında kelime testleri hazırlayan İngilizce Öğretmenlerin yeterlilik açısından ne derecede doğru olup olmadıkları tespit edilmeye çalışılacaktır.

Çalışma beş bölümden oluşur. Çalışmanın birinci bölümü, çalışmanın problemini, amacını, kapsamını, arka planını, ve yöntemini aktarır. İkinci bölümde, dil testlerin amaçlarını, öğretme metotları, kelime öğretme stratejileri ve kelime testlerin çeşitleri.

Üçüncü bölümde, öğretmenlerden toplanan veri analizleri sunulur. Dördüncü bölümde ise, önerilmiş bir kelime testi verilmektedir. Beşinci bölümde, sonuç ve öneriler sunulmaktadır.

ABSTRACT

“Vocabulary Testing Problems with Specific Reference To Macedonian State Schools”

The purpose of this study is to find out whether teachers of English in Macedonian State Secondary Schools are properly constructing vocabulary tests. Actually, to find out that the competency of English teachers at State Secondary Schools to construct vocabulary tests is not enough.

The study consists of five chapters. The first chapter of the study presents the problem of the study, the purpose, the background of the study and methodology of the study. In the second chapter will be presented, the aims of language tests in general, teaching methods and approaches, vocabulary teaching techniques, and vocabulary testing techniques. In the third chapter will be presented the analysis of the data collected from teachers. In the fourth chapter a suggested vocabulary test will be presented. Finally, in the fifth chapter conclusion and suggestions will be presented.

TABLE OF CONTENTS

	<u>Page</u>
ACKNOWLEDGEMENT	i
ÖZET	ii
ABSTRACT	iii
 CHAPTER I	
INTRODUCTION	1
1.1. Statement of the Problem	2
1.2. Aim of the Study	3
1.3. Scope of the Study	4
1.4. Background of the Study	4
 CHAPTER II	
REVIEW OF LITERATURE	5
2.0. Introduction	5
2.1. Aims of a Language Test	5
2.2. Language Teaching Method	8
2.2.1. The Grammar Translation Method	9
2.2.2. The Direct Method	10
2.2.3. The Audio-Lingual Method	12
2.2.4. Total Physical Response	13
2.2.5. The Silent Way	15
2.2.6. Communicative Language Teaching	16
2.3. Communicative Language Tests.....	17
2.4. Types of Vocabulary Knowledge	19
2.4.1. Active Vocabulary	20
2.4.2. Passive Vocabulary	21
2.5. Vocabulary Teaching Techniques	22
2.5.1. Using Games in Teaching Vocabulary	24

	<u>Page</u>
2.5.2. Using Songs in Teaching Vocabulary	28
2.5.3. Guessing the Meaning from the Context	31
2.5.4. Using Pictures	35
2.5.5. Guessing the Meaning by Form	37
2.5.6. Using Mime and Gestures	38
2.5.7. Using Realias	38
2.6. Vocabulary Testing Techniques	40
2.6.1. Subjective Language Testing	42
2.6.2. Objective Language Testing	44
2.6.2.1. Multiple Choise Items	45
2.6.2.1.1. Advantages of Multiple Choise Items	52
2.6.2.1.2. Disadvantages of Multiple Choise Items ...	54
2.6.2.2. Completion Items	56
2.6.2.2.1. Advantages of Completion Items	60
2.6.2.2.2. Disadvantages of Completion Items	61
2.6.2.3. Items Involving Synonyms	61
2.6.2.4. Matching Items	62
2.6.2.4.1. Advantages of Matching Items	65
2.6.2.4.2. Disadvantages of Matching Items	66
2.6.2.5. Definitions	67
2.6.2.6. Associated Words	69
2.7. Selecting Words to Test	70
2.8. Integrative Testing of Vocabulary with Other Language Skills	72
2.8.1. Testing Vocabulary with Reading	72
2.8.2 Testing Vocabulary with Writing	74
2.8.3 Testing Vocabulary with Speaking	75
2.8.4 Testing Vocabulary with Listening	76
2.9. Characteristics of a Good Language Test	78
2.9.1. Reliability	79

	<u>Page</u>
2.9.1.1. Factors Affecting Reliability Measures	80
2.9.1.2. Methods of Estimating Reliability	81
2.9.2. Validity	82
2.9.3. Practicality	85
2.10. Things to be done after Test Application: Item Analysis	86
2.10.1. Difficulty Level	87
2.10.2. Discrimination Level	88
 CHAPTER III	
DATA ANALYSIS AND INTERPRETATION OF THE RESULTS	91
3.0. Introduction	91
3.1. Interpretation of the Questionnaire Administered to the Teachers	103
 CHAPTER IV	
SUGGESTED VOCABULARY TESTS	110
4.0. Samples of Vocabulary Tests for the Fourth Year Secondary School Students	110
 CHAPTER V	
CONCLUSION AND SUGGESTIONS	121
5.1. Conclusion	121
5.2. Suggestions	124
 BIBLIOGRAPHY	 127
 APPENDICES	 133
APPENDIX A- Questionnaire For Secondary School Teachers	134
APPENDIX B- Answer Key For the Suggested Vocabulary Tests	139

CHAPTER I

INTRODUCTION

1.0 Introduction

Any doubts about the sense of testing grammar apply equally to the testing of vocabulary. A clear knowledge of vocabulary is essential to the development and demonstration of linguistic skills.

Read (2000:1) views vocabulary assessment as necessary in the sense that words are the basic building blocks of language, from which larger structures such as sentences, paragraphs and whole texts are formed.

From different points of view, vocabulary can be seen as a priority in 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.

Many reasons can be stated about why assessing vocabulary knowledge is regarded as a useful element in language teaching programs. Schmitt (2000:163) as the most important reason of testing vocabulary views the fact that "if vocabulary is considered to be a component of language knowledge, then it naturally needs to be assessed in some way". For some learners, mainly beginners, vocabulary may be about the only aspect of language they know well to be tested on.

Heaton (1990:79) states that for students not only the knowledge of the meaning of a word is essential but also an awareness of the other words with which it is generally used.

This study copes with a review of literature on vocabulary teaching and testing techniques that should be followed when constructing vocabulary tests.

1.1 Statement of the Problem

Vocabulary is important for a learner and for a language situation. Expressing something in a foreign language or communicating requires knowledge of vocabulary. Read (2000:39) classifies learning vocabulary into four processes by which learners add new words to their vocabulary. The first is the systematic procedure to memorize the form and meaning of words based on lists of foreign language together with their translation equivalents in the learners' native language. The second, is incidental learning of vocabulary where learners acquire knowledge of new words incidentally as they encounter them in context through their reading and listening activities. The third process focuses on the extent to which learners can suppose meaning of unknown words when they read or listen to them. The fourth process is learners' use of communication strategies to deal with situations where they lack the vocabulary to put across the meaning they want to express.

Schmitt and McCarthy (1997:3) agree with the first procedure stating that the learner's native language is one of the most important factors in learning foreign language vocabulary.

Richards (1976:81) supports the second procedure stating that "words do not exist in isolation. We can define their meanings through relationships with other words".

Testing is an important and indispensable subject in teaching and learning process. It has a wide scope including uses of testing; types of testing construction, administration, evaluation and interpretation of test scores and testing certain skills or elements.

However, these areas mentioned above do not mean anything without the requirements for the test. In short, requirements of tests constitute a vital aspect of testing. Without fulfilling the requirements of tests, a test is not a test and hence of no value. A good test possesses three qualities: reliability, validity and practicality. That is to say, "any test that we use must be dependable in the evidence it provides,

appropriate in terms of our objectives and applicable to our particular situation” as Harris (1969:13) states.

The above statements in a way summarize the language testing situation in Macedonian State Secondary Schools. The tests designed to measure the knowledge of vocabulary are not constructed to fulfill the requirements of vocabulary testing techniques.

1.2 Aim of the Study

The way vocabulary tests are constructed in Macedonian State Secondary Schools does not satisfy the needs and requirements of the learners, that is to say the competency of teachers to construct such tests is not well enough. Moreover, the language teaching approaches and different vocabulary teaching techniques will be mentioned and the main purpose of this study- the way how to construct valid vocabulary tests.

It should be emphasized also that the aim of this study is not to judge how teachers in Macedonia construct tests but in a way to find their needs in writing vocabulary tests and make them conscious of the importance of such tests.

Schmitt (2000:164) states that there are several purposes of testing vocabulary. The first one is to find out if students have learned the words they were taught during a course. The second purpose is to see the students' vocabulary gaps and in this way give more importance to that specific vocabulary. Another purpose of testing vocabulary is to measure how well words or word families are known.

1.3 Scope of the Study

This study will be carried out with a questionnaire administered to 30 teachers, teaching English as a Foreign Language at different Secondary Schools all around Macedonia.

The names of the schools that teachers were working are: Pance Poposki, Zlate Malakoski, Nikolla Shtejn, Gjoce Stojceski, Koco Racin, Zef Lush Marku, Kiril Pejcinovic, and Kliment Ohridski.

The purpose of the questionnaire was to see how teachers construct different language test types and the competency of the teachers in constructing vocabulary tests.

1.4 Background of the Study

All the teachers conducting this questionnaire and teaching in different schools in Macedonia use the Upper- Intermediate “New Headway” coursebook from Liz and John Soars. This book has been used from 1998 and is the most applicable book for Secondary schools in Macedonia. Regarding the issue of material evaluation and adaptation, all of the teachers agree on the appropriateness of this book and they view this book as in accordance with the learners needs. The book provides nearly 120 hours of language learning and it is covered in two semesters. The first semester teachers cover the first six units of the book and the proceeding six units are covered in the second semester.

Assessment during the year is done through achievement tests (approximately 2 achievement tests per semester) and according to the results gained from these tests, teachers discover the weaknesses of the students and in this way decide which parts need more practice and which parts are satisfactory.

CHAPTER II

REVIEW OF RELATED LITERATURE

2.0 Introduction

In the beginning of this chapter we will emphasize in general the purposes of language testing and its importance. Second, the focus will be on language teaching methods and the place of vocabulary in using these methods. Then, the types of vocabulary knowledge including active and passive vocabulary will be presented. Next, the techniques used in teaching vocabulary will be showed. In part five, the focus will be on the importance of vocabulary testing and testing techniques. In the sixth part, the way of selecting words to test. Then, the integrative testing of vocabulary with other language skills will be viewed. In part eight, we will focus on characteristics of a good vocabulary test such as reliability, validity and practicality. Finally, the focus will be upon things to be done after test application.

2.1 Aims of a Language Test

The most important consideration in the development and use of language tests is the purpose or purposes for which the particular test is intended. By far the most essential use of language tests is for purposes of evaluation in educational programs.

Evaluation has an indispensable part in language teaching process but it is not restricted to the context of education; it is a part of our everyday lives. However, in the educational context the results of a test may determine whether a learner moves into a higher level, class or repeats the course or year. Dickins and Germaine (1992:4) state that evaluation in an educational context should be systematic and

undertaken according to certain guiding principles using carefully defined criteria. No matter what the situation is, foreign language teachers are supposed to find out how well they have been teaching or in other words they have to find out the progress their students have made during a particular time. Gronlund (1976:6) defines evaluation as a systematic process of determining the amount to which instructional objective are obtained by the students. Educational evaluation has an important role in the education where the emphasis is on student learning progress.

Language tests also have a potentially important use in several areas of research. The information obtained from language tests can assist in the investigation of the very nature of language proficiency, in examining how individuals process language, in the study of language acquisition and slow destruction and in assessing the effects of different instructional methods and settings on language learning.

Linn and Gronlund (1995:115) state that the main purpose of testing and assessment is to obtain valid, reliable and useful information about the student achievement. It collects objective information that may be used in conjunction with subjective information to make better educational decisions.

Heaton (1975:1,2) emphasizes that the function of a test is to enhance motivation towards learning or to evaluate student achievement enabling the teacher to discover which parts of the language program have been found difficult by the class.

Kubiszyn and Borich (1993:14) also state that “tests represent an attempt to provide objective data that can be used with subjective impressions to make better, more defensible decisions”.

Heaton (1989:9) lists the reasons for language testing as:

- Finding about progress- to find out how well the students have mastered the language areas and skills that have been taught;

- Encouraging students- it can serve to show the students each set of goals which they have reached on their way to fluency and in this way increase their motivation;
- Finding about learning difficulties- diagnose specific weaknesses and problems the students may have encountered;
- Finding about achievement- in a way an achievement test is also like a progress test but it is usually designed to cover a longer period of learning;
- Placement test- it enables the teachers to sort students into groups according to their language ability.

On a parallel basis Tuckman (1975:43) lists the set of reasons of testing as:

1. To determine whether and to what degree each student is experiencing learning and development
2. To identify those students who are not learning and growing in particular area of their deficiency
3. To determine whether instructional inputs are in general, effective.

However, language testing does not include only the testing of language skills, it requires assessing of the words the students have learnt since the vocabulary knowledge is accepted as a main component of second/foreign language proficiency. Vocabulary is central to language and of critical importance to the typical language learner.

2.2 Language Teaching Methods

"The superior teacher has regularly gotten superior results regardless of the method."

(William E. Bull. *Spanish for Teachers: Applied Linguistics*, c.1965)

Human beings growing and living in given societies to varying degrees require new languages after they have learnt their first language. Stern (1983:21) defines language teaching as the activities, which are intended to bring about language learning. He puts under the concept of teaching the preparation of teaching materials, teaching grammars or dictionaries or the training of teachers as well as making the necessary administrative provision inside or outside the educational system. In other words language teaching is interpreted as to include all the activities intended to bring about language learning.

In his treatment of the historical developments in language pedagogy, Stern (1992:6) isolates three ways in which language pedagogy has aimed to renew and improve itself:

1. Innovation through change in teaching methods;
2. Innovation through language-related sciences and research;
3. Technological innovation.

Every teacher tries to develop a certain characteristic approach to instruction. McKean (1962:112) relates this to be a product of teacher's past experiences as a student in various schools, training program and testing as teachers work at teaching. The method of teaching is affected by the teachers' notion of the nature of the subject matter, the role and the aims of the school, and the relationship of teacher and students in the learning situation. Therefore, McKean (1962:113) leaves us to decide what method to use in our teaching experience stating that "teachers try hard to find an instructional approach which will be most beneficial for the students in their charge, which will be satisfying and challenging to themselves, and which will be judged by other professional workers, in the light of its results, as successful. And Celce-Murcia and McIntosh (1979:11) furthermore state that "in the teaching

situation, the achieved results are determined by the method used as the most important factor”.

Language educators have attempted to solve the problems of language teaching by focusing attention on teaching methods. In the next section we are going to discuss about the historical sequence of the language teaching methods and the way they have been used with reference to vocabulary.

2.2.1 Grammar-Translation or Traditional Method

This method was first used in teaching of classical languages, Latin and Greek. Here classes are taught in the students' mother tongue, with little active use of the target language. Vocabulary is taught in the form of isolated form list. Freeman (2000:18) states that the most emphasized areas of language are vocabulary and grammar. The students in order to show that they have understood the meaning and the use of a new vocabulary item; they make up sentences in which they use the new words. Grammar instruction provides the rules for putting words together; instruction often focuses on the form and inflection of words. Reading of difficult texts is begun early in the course of study. Little attention is paid to the content of texts, which are treated as exercises in grammatical analysis. Often the only drills are exercises in translating disconnected sentences from the target language into the mother tongue, and vice versa. Little or no attention is given to pronunciation.

Krashen (1982:127) summarizes this approach with following activities:

1. Explanation of a grammar rule, with example sentences
2. Vocabulary, presented in the form of a bilingual list.
3. A reading selection, which emphasizes the grammar rule and the presented vocabulary.
4. Exercises designed to provide practice on the grammar and vocabulary of the lesson, including translation in both directions, from native language to foreign language and from foreign language to native language.

Celce-Mursia and McIntosh (1979:3) characterize this method with following characteristics:

- Classes are taught in the native language, with little active use of the target language.
- Vocabulary is taught in the form of lists of isolated words.
- Grammar provides the rules for putting words together, and instruction often focuses on the form and variation of words.
- Reading of difficult classical texts is begun early.
- Texts are treated as exercises in grammatical analysis and little attention is paid to their content.

Krashen (1982:128) also claims that Grammar-Translation Method provides scraps of comprehensible input, where the model sentences are mostly understandable, but the focus is entirely on form and not meaning.

Language is tested through “written tests in which students are asked to translate from their native language to the target language or vice versa”. (Freeman, 2000:18)

According to Stern (1983:455) the major defect of this method is the overemphasis on the language as a mass of rules, which never let go the learner from the dominance of the first language.

2.2.2 The Direct Method

The interest in developing principles grew out of naturalistic principles of first-language acquisition and reflected the beginnings of the discipline of applied linguistics. The principles emphasized the importance of listening and speaking the foreign language. Meaningful contexts for learning, inductive teaching of grammar and avoidance of translation were some of the principles present. These principles

provided the foundation for the Direct Method. The new method, which emphasized communication in the target language, arose out of the need for more effective language learning.

Krashen (1982:135) characterizes that in this method the discussion and the classroom language is the target language, including the language of the exercises and teacher talk used for classroom management.

Celce-Murcia and McIntosh (1979:3) see this method as a reaction to the extension of the grammar- translation method and sum up it as follows:

1. Lessons begin with a brief anecdote or dialogue in the target language.
2. The material is first presented orally with actions or pictures.
3. The mother tongue is never used.
4. The favorite type of exercise is a series of questions in the target language based on the anecdote or dialogue and answered in the target language.
5. Grammar and culture associated with the target language are taught inductively.
6. Verbs are used first, and systematically conjugated much later.

According to Freeman (2000:23) this method appeared when the goal of instruction became learning how to use a foreign language to communicate through the use of demonstration and visual aids with no resource to the students' native language. In this method lessons begin with a dialogue using a modern conversational style in the target language. Material is first presented orally with actions or pictures. The mother tongue is never used. There is no translation. The preferred type of exercise is a series of questions in the target language based on the dialogue or an anecdotal narrative. Questions are also answered in the target language. Freeman (2000:29) also states that here vocabulary is emphasized more than the target language grammar. It is taught inductively and rules are generalized from the practice and experience with the target language.

The students in the Direct Method are evaluated by being asked to use the language; they might be interviewed orally or might be asked to write a paragraph about something they have studied.

Stern (1983:460) states that the break of this method was in that how to convey meaning of words without translating and how to protect against misunderstanding without reference to the native language.

2.2.3 The Audio-Lingual Method

Krashen (1982:132) as the main goal of this method states “having the student over-learn a variety of patterns to be used directly in performance”. If learned according to plan, students will finish up with an accumulation of sentences and patterns that will be of rare use in conversation, and also serve the conversational management goal, to some amount. The goal is also the memorization of the dialogue, not the comprehension of a message.

Freeman (2000:42) states that here language forms do not occur by themselves but most naturally within a context. Teachers should provide students with a good model. By listening to how it is supposed to sound, students should be able to mimic the model.

Celce-Murcia and McIntosh (1979:18) also define the techniques of audio-lingual method as:

- various kinds of mimicry and memorization and
- pattern drills based on analogies.

The main basic principles of the Audio-lingual method were listed by Celce-Murcia and McIntosh (1979:18) as follows:

- a) Language is speech, not writing.
- b) A language is a set of habits.
- c) Teach the language, not the facts about the language.

- d) A language is what native speakers say, not what someone thinks they have to say
- e) Languages are different.

Krashen (1982:129) states that at this method “The lesson typically begins with a dialogue, which contains the structures and vocabulary of the lesson”. The student is expected to mimic the dialogue and sooner or later memorize it. There is no explicit effort to limit learning to rules that are learnable, convenient and not yet acquired, nor is there any attempt to encourage rule use only in certain situations.

New vocabulary and structural patterns are presented through dialogs. The dialogs are learned through imitation and repetition. Students’ successful responses are positively reinforced and grammar is induced from the examples given.

According to Freeman (2000:45) the goals of teachers who use the Audio-Lingual Method are that “Teachers want their students to be able to use the target language communicatively. In order to do this, they believe students need to overlearn the target language, to learn to use it automatically without stopping to think”.

Freeman (2000:47) in addition states that an audio-lingual test in nature would be discrete-point or each question on the test would focus on only one point of the language at a time. Students might be asked to make a distinction between words in a minimal pair, for example, or to provide an appropriate verb form in a sentence.

2.2.4 Total Physical Response

Total Physical Response is usually introduced in the student’s native language. After the introduction, rarely would the native language be used. Meaning is made clear through body movements.

According to Krashen (1982:140) Total Physical Response consists basically of obeying commands given by the instructor that involves an evident physical

response. Meaning in the target language can often be conveyed through actions. Memory is activated through learner response. The target language should be presented in chunks, not just word by word.

Freeman (2000:113) defines the aims of the teachers who use Total Physical Response as “they believe in the importance of having their students enjoy their experience in learning to communicate in a foreign language”, that is to say basing foreign language upon the way the student learn their native language.

Krashen (1982:140,141) defines Total Physical Response requirements for best possible input as logical because TPR meets this requirement. The Total Physical Response required of the student is, in effect, an expression of his understanding of the teacher’s expression.

Freeman (2000:113) states the first phase of a TPR as one of the modeling. Here the instructor issues commands to a few students, then performs the actions with them. In the second phase, these same students demonstrate that they can understand the commands by performing them alone.

Vocabulary and grammatical structures are emphasized over the language areas. These are surrounded within imperatives. The imperatives are single words and multi-word chunks. Students learn the grammar inductively, working out the correct form of the rule during the class activity.

Freeman (2000:115) states that at this method concerning the student errors “teachers should be tolerant of them and only correct major errors”. Even these should be corrected unremarkably. As students get more advanced, teachers can “fine-tune” –correct more unimportant errors.

According to Krashen (1982:141) Total Physical Response Method makes one very important contribution to reducing the student anxiety saying that “Students are not asked to produce in the second language until they themselves decide whether they are ready”.

Observing the students' actions does evaluation and teachers will know immediately whether or not the students understand. Formal evaluations can be conducted simply by commanding individual students to perform a series of actions. As students become more advanced, their performance of parody they have created can become the basis for evaluation.

2.2.5 The Silent Way

This method begins by using a set of colored rods and verbal commands. Although the name is somewhat confusing because the students do make oral statements and responses in the language they are learning (Celce-Murcia and McIntosh (1979:22). But the teacher, especially, speaks much less than in the Audio-lingual classroom and the students do not mimic and repeat aloud so frequently.

Freeman (2000:61) states that “the teacher should start with something the students previously know and construct from that one to the known. The students are helped only when necessary and their actions tell the teacher whether or not they have learned. The teacher works with the students while the students work on the language. This method is used to avoid the use of the language and create simple linguistic situations that remain under the complete control of the teacher.

According to Celce-Murcia (1979:22) the students “are motivated to think and say the appropriate sentences to go together with actions performed under the guidance of the teacher. The most significant point of this method is the devoted concentration with which the students watches the actions and listens to the utterances of the teacher and his or her colleague students.

Freeman (2000:61) defines silence as a tool. It helps to encourage autonomy, or the exercise of the initiative. This silence removes the teacher from the center of attention so he can listen and work with students. Again the same silence encourages group cooperation since here students learn from one another. The attention is a key

to learning and students must pay attention on what the teacher says in order not to miss what he says.

Freeman (2000:66) states that “all four skills are worked on from the beginning of the course, although there is a sequence in that students learn to read and write what they have previously created orally. The skills strengthen what students are learning and vocabulary is to some extent limited at first.

Evaluation of students is done during the learning, and here the teacher may never give a formal test.

2.2.6 Communicative Language Teaching

Freeman (2000:122) states that Communicative Language Teaching aims mostly to apply the theoretical point of view of the Communicative Approach by making communicative proficiency the goal of language and communication. The method stresses a means of organizing a language syllabus. The emphasis is on breaking down the global concept of language into units’ analysis in terms of communicative situations in which they are used.

According to Freeman (2000:125,126) whenever possible “authentic language-language as it is used in a real context should be introduced”. The target language is a vehicle for classroom communication, not just the object of study. The stress is on the process of communication rather than just mastery of language forms. Students must learn about structure and consistency, which bind the sentences together.

Freeman (2000:127) also states that at this method errors are tolerated and seen as a natural result of the development of communication skills. And the primary task of the teacher is to establish situation expected to encourage communication where he acts as a facilitator in setting up communicative activities and as an advisor during the activities.

The grammar and vocabulary that students learn follow from the function, situational context, and the roles of interlocutors.

The teacher evaluates not only the student's accuracy, but also their fluency. The student who has the most control of the structures and vocabulary is not always the best communicator. (Freeman 2000:132).

For more formal evaluation, a teacher is likely to use an integrative test which has a real communicative function.

2.3 Communicative Language Tests

Communicative language tests are intended to be measure of how the test takers are able to use language in real life situations. In testing productive skills, emphasis is placed on appropriateness rather than on to form grammatically correct sentences.

McNamara (2000:) states that communicative language tests ultimately came to have two characteristics:

1. They are performance tests, requiring assessment to be carried out when the learner or candidate is engaged in an extended act of communication, either receptive or productive, or both.
2. They pay attention to the social roles that candidates are likely to assume in real world settings, and they offer a means of specifying the demands of such roles in detail.

The communicative approach in English Language Teaching should be accompanied by the tests assessing systematically the individual's capacity to communicate. Savignon (1983:246,247) underlines the considerable needs for communicative language tests stating that:

“..... the most important implication of the concept of communicative competence is undoubtedly the need for tests that measure an ability to use language effectively to attain communicative goals... If we teach for communicative competence, we have to test for communicative competence so that we and our students know how well we are doing what we claim to be doing...”

Morrow (1979: 149) argues that traditional tests do not give any convincing proof of the candidate's ability to use the language in ordinary situations. In short, communicative tests are concerned primarily with how efficient language is used in communication. Success is evaluated in terms of both the effectiveness of the communication which takes place and the correctness of the grammar.

The design of any test is primarily based on a set of criteria. In any test which claims to assess communicative language ability, test constructors must identify the skills and performance conditions that are the most important components of language use in particular context. The identification and incorporation of these two features, skills and performance conditions, makes a test task applicable in real life so that the learner can function in normal conditions in his future target situation. Skehan (1988: 98).

The language of the test should be day-to-day discourse, not edited but presented with all its expected irregularities. It seems as Morrow (1979:152) maintains, that conventional tests ignore certain features of genuine language use, which are that language is interactive, unpredictable, purposive, authentic, contextualized, based in performance and assessed in terms of behavioral outcomes.

On the other hand, Harrison (1985:) also puts forward three distinguishing features which may be considered as fundamental for communicative language tests:

1. **A communicative test should assess language used for a purpose beyond itself.** An oral interview, for example, can be used to assess how well the learner can manipulate language in response to stereotyped questions, but a communicative question demands a response to

circumstance. This rules test out in which the learner is asked to display his language competence for no other reason than to have it assessed.

2. **A communicative test should depend on the bridging of an information gap.** It has to propose a language-using purpose which can be fulfilled by the communicative skill so far acquired by the learner. For example, one kind of role play may consist of reading out the alternative parts in a script, another could be planning a holiday on the basis of travel brochures. The former has no intrinsic value as communication process; the latter necessarily involves expression of information by one part essential to the exchange and understanding of it by the other. This approach rules out exchanges in which the questioner has no interest in the answer.
3. **Communicative test should represent an encounter.** The situation at the end of the test should be different from what it was at the beginning, and this means that there has to be some sequence within the test. It may require a simple transition from the unknown to the known.

However, a test type does not become communicative by applying it in a realistic context; it is communicative because of the use made of it, and if can not be used to represent a communicative purpose, it cannot be a communicative test. It should always be noted that the more the tests include what is considered as the best practice in the language classroom, the more the match between teaching, testing and reality enhanced.

2.4. Types of Vocabulary Knowledge

In this stage we will deal with receptive and productive vocabulary or with active and passive vocabulary and the vital difference between them, and their importance when teaching vocabulary.

2.4.1 Active Vocabulary

“The number of words we can recognize and understand is larger than the number of the words we use in our speech and writing” (Read, 2000:154). Words first are known receptively and later become available for productive use.

Ellis and Tomlinson (1980:65) state that a learner’s active vocabulary consists of the total number of lexical items which he/she can exactly use in speech and writing. Any item which becomes a learner’s active vocabulary must initially have been part of his/her passive vocabulary and the suggestion of this for teaching is important.

The use of active vocabulary involves wanting to express a meaning through speaking or writing and retrieving and producing the appropriate spoken or written word form.

Ellis and Sinclair (1989:63) define active vocabulary as a language item which the learner can use appropriately in speaking or writing. Shepherd (1987:3) also defines active vocabulary as “words you use when you speak or write”.

Nation (2001:28) states that knowing an active word involves:

- Being able to say it with correct pronunciation including stress
- Being able to write it correct spelling
- Being able to construct it using the right word parts in their appropriate forms
- Being able to produce the word to express the meaning
- Being able to produce the word in different contexts to express the range of meanings
- Being able to produce synonyms and opposites
- Being able to correct the word correctly in an original sentence
- Being able to produce words that commonly occur with it

- Being able to decide to use or not to use the word to suit the degree of formality of the situation.

Shepherd (1987:15) states that when you raise your vocabulary size of passive vocabulary, you also increase the possible size of your active vocabulary because the words you use when you write and speak are the words you know when you read or hear.

2.4.2 Passive Vocabulary

Nation (2001:25) defines using of passive vocabulary as “perceiving the form of a word while listening or reading and retrieving its meaning”. Ellis and Sinclair (1989:63) also define passive vocabulary as “language items that can be recognized and understood in the context of reading and listening”. Shepherd (1987:3) as well states that passive vocabulary is the words you know when you receive thoughts from others.

Passive vocabulary is also the basic vocabulary or the total stockroom of words you can use to understand the thoughts of others when you listen and read.

Ellis and Tomlinson (1980:65) state that a learner’s passive vocabulary “will always be larger than his active vocabulary as he will have heard or read items which he will never need to use, and he will be able to understand around many items which he can not use with any degree of accuracy”.

Nation (2000:26) states that knowing a passive word involves:

- Being able to recognize the word when it is heard
- Being familiar with its written form so that it is recognized when it is met in reading
- Recognizing that it is made up of parts

- Knowing what the word means in the particular context in which it has just occurred
- Knowing the concept behind the word which will allow the understanding in a variety of contexts
- Knowing that there are related words to that word
- Being able to recognize that the word has been used correctly in the sentence in which it occurs

2.5. Vocabulary Teaching Techniques

Teachers when choosing the materials for classroom use pay attention to students' needs and preferences, i.e., they try to compile with students' expectations in a way and to the curriculum design, or their preparation. The vocabulary introduced in the classrooms by teachers is conditioned by the books used, that cover certain areas. Thus, it is very important for teachers working in the field of teaching foreign languages to be aware of the variety of techniques that might help them to teach and to use vocabulary in every day life, to enlarge their vocabulary skills for everyday usage, etc.

Wallace (1982:9) states that, it is possible to have a good knowledge of how the system of a language works and thus far not be able to communicate in it; but if we have the vocabulary we need, it is generally possible to communicate.

Celce-Murcia and McIntosh (1979:242) also define vocabulary as an essential part in language teaching from the beginning stages.

Ellis and Tomlinson (1980:60) state that the aim of vocabulary teaching is to increase the students' vocabularies or to increase the number of lexical items which the students can understand and use.

Teachers should help students expand their vocabulary and also advance their vocabulary use. To make the learning of words easy, teachers should build up

techniques for different words. Therefore, Celce-Murcia and Rosenweig (1979:247) indicate that “it seems reasonable that different techniques are required for teaching different item types: e.g. idioms, sets of related words, abstract words”. For that reason, you will be recommended you should not rely on just one technique but better use more techniques so that different students can act in response or react to it in their own manner.

There are several aspects of words that need to be taken into account when teaching vocabulary. The list below is based on the work of Gairns and Redman (1986:)

- *Limits between conceptual meanings*: knowing not only what lexis refers to, but also where the boundaries are that separate it from words of related meaning (e.g. cup, mug, bowl).
- *Polysemy*: distinguishing between the various meaning of a single word form with several and closely related meanings (head: of a person, of a pin, of an organisation).
- *Homonymy*: distinguishing between the various meaning of a single word form which has several meanings which are NOT closely related (e.g. a file: used to put papers in or a tool).
- *Homophony*: understanding words that have the same pronunciation but different spellings and meanings (e.g. flour, flower).
- *Synonymy*: distinguishing between the different shades of meaning that synonymous words have (e.g. extend, increase, expand).
- *Affective meaning*: distinguishing between the attitudinal and emotional factors (denotation and connotation), which depend on the speakers attitude or the situation. Socio-cultural associations of lexical items are another important factor.

- *Style, register, dialect*: Being able to distinguish between different levels of formality, the effect of different contexts and topics, as well as differences in geographical variation.
- *Translation*: awareness of certain differences and similarities between the native and the foreign language (e.g. false cognates).
- *Chunks of language*: multi-word verbs, idioms, strong and weak collocations, lexical phrases.
- *Grammar of vocabulary*: learning the rules that enable students to build up different forms of the word or even different words from that word (e.g. sleep, slept, sleeping; able, unable; disability).
- *Pronunciation*: ability to recognise and reproduce items in speech.

The goals of vocabulary teaching must be more than simply covering a certain number of words on a word list. We must use teaching techniques that can help realise this comprehensive concept of what it means to know a lexical item. And we must also go beyond that, giving learner opportunities to use the items learnt and also helping them to use effective written storage systems.

In this study eight vocabulary teaching techniques will be viewed. The suggested techniques are as follows: games in teaching vocabulary, songs in teaching vocabulary, pictures in teaching vocabulary, realias in teaching vocabulary, translations and explanations, mime and gestures, guessing the meaning of a word by the context and guessing the meaning of a word by the form.

2.5.1. Using Games in Teaching Vocabulary

Learners of English have to deal with unfamiliar vocabulary during their language acquisition. In order to learn and retain new words, learners should participate in different task-based activities in their classroom. Such activities also

include vocabulary games which especially focus on helping learners develop and use words in different contexts by making the lessons enjoyable.

Some experts have figured out characteristics of games that make vocabulary learning more effectively. Lee (1995:35) lists several main advantages when games are used in the classroom, including "a welcome break from the usual routine of the language class", "motivating and challenging" "effort of learning", and "language practice in the various skills". Ersöz (2000) holds that games are highly appreciated thanks to their amusement and interest. Teachers can use games to help their students practice more their skills of communication.

Betteridge and Buckby (1983:1) state that games also help the teacher to create contexts in which the language is useful and meaningful. "The learners want to take part and in order to do so must understand what others are saying or have written, and they must speak or write in order to express their own point of view or give information.

Huang (1996: 1) comes to a conclusion that "learning through games could encourage the operation of certain psychological and intellectual factors which could facilitate communication heightened confidence, motivation and naturalness, reinforcing learning, improving intonation and building confidence".

Dawson (1984:53) associates games in people's mind with fun and relaxation, stating that "while the students are having fun and relaxing they can also be learning English". Thus, games will serve as a relaxation in a way, and enjoyment where students feel free to use their acquired vocabulary. In addition, using games stimulate students' interest, and it can be one of the best motivating techniques.

McMallum (1980: ix) lists the reasons of using games in language teaching as follows:

1. They focus student attention on specific structures, grammatical patterns, and vocabulary items.

2. They can function as reinforcement, review and enrichment.
3. They involve equal participation from both slow and fast learners.
4. They can be adjusted to suit the individual ages and language levels of the students in the class.
5. They contribute to an atmosphere of healthy competition, providing an outlet for the creative use of the natural language in a non-stressful situation.
6. They can be used in any language teaching situation and with any skill area whether reading, writing, speaking or listening.
7. They provide immediate feedback for the teacher.
8. They ensure maximum student participation for a minimum of teacher participation.

Wallace (1982:105) gives two reasons about using language games:

First, a growing emphasis on the importance of real communication, and the appropriate kind of positive affective atmosphere in the classroom and second, an increasing emphasis on the importance of real communication. If a game works properly, it very often supplies a real desire to communicate in the target language, even within the artificial limitations of the classroom.

The below suggested games can be used by the teachers in their classes.

2.5.1.1 Hidden Object

Ten or twenty small objects are shown to the students, who name them. They are then put into a bag. The teacher picks up an object but does not remove it from the bag. A member of each team in turn is asked to guess what it might be. The first person to guess successfully is given the object. The team which ends up with most objects wins.

2.5.1.2 A-A, B-B

This game is helpful for revision of vocabulary items according to categories. For example: color; making of sentences containing selected vocabulary items; remembering and repeating.

Step 1:

Select several sentence patterns suitable for using with vocabulary items that are to be revised.

Step 2:

Give the learners time to work in groups to bring together lists of words, all beginning with a letter, organized according to one or more categories.(colors, articles of clothing, shops, vehicles, animals, occupations, verbs of motion, etc.)

Step 3:

When the learners are ready, divide them into teams and give the opening of your selected sentence to a representative of one team. The team representative must repeat what you have said and complete the sentence using a word (or words) from the given category. remembering as best as s/he can the preparatory work just done in groups, plus any other ideas that may spring to mind.

Repetition of any key word within a sentence is against the rules, but repetition of what other players have said earlier in the context of other sentences, is allowed, indeed is to be encouraged since in that way it will enable the players to learn from listening to each other.

Objective of the game:

The object of the game is to see how long and elaborate sentence the teams' representative can develop. The longest sentence, containing no self-repetitions, and spoken without too much hesitation, wins.

2.5.2 Using Songs in Teaching Vocabulary

Songs can create an enjoyable, relaxing or particular atmosphere, to establish a mood or to suggest associations. Music or songs can also provide cover to encourage learners to speak, acting in the same way as music at a party to break down people's inhibitions about talking to one another.

Dawson (1984:54) states that songs provide good pronunciation practice and unconscious repetition of useful words and phrases. The music helps create a relaxed atmosphere in the classroom.

This song is presented by Fatma Kirlar (2001) and performed by The Beatles. The teaching points in this song are Simple Past Tense and Vocabulary.

YESTERDAY

Yesterday, all my troubles seemed so far away,
Now it looks as though they here to stay
Oh, I believe in yesterday

Suddenly, I'm not half the man I used to be,
There is a shadow hanging over me,
Oh, yesterday came suddenly

Why she had to go I don't know,
She wouldn't stay
I said something wrong now I long for
Yesterday

Yesterday, love was such an easy game to play,
 Now I need a place to hide away
 Oh, I believe in yesterday

Why she had to go I don't know,
 She wouldn't stay
 I said something wrong now I long for
 Yesterday

Yesterday, love was such an easy game to play,
 Now I need a place to hide away
 Oh, I believe in yesterday.

Before Listening

Activity 1: The teacher writes the title of the song on the board and then s/he has the students in pairs make guesses about what the song is and write them down on a piece of paper.

Activity 2: The teacher gives the students some words related with the song.

1. yesterday 2. trouble 3. game 4. love 5. to hide away 6. to believe in 7. place 8. to hang over 9. shadow

For Example: Marry saw a man yesterday and she fell in love with him suddenly. She believed that the man also loved her..... etc.

While Listening

Activity 3: In this section, the teacher firstly divides the students in four groups and each group's answers one passage of the song. Then, the teacher gives the students a worksheet and s/he wants them to fill in the blanks using the words that are given below. The song will be played twice.

WORKSHEET

Fill in the blanks with the words given below.

- | | | |
|-----------------|--------------|---------------|
| 1. far away | 2. long for | 3. hide away |
| 4. used to be | 5. had to | 6. believe in |
| 7. hanging over | 8. had to go | 9. easy game |

YESTERDAY

Yesterday, all my troubles seemed so,

Now it looks as though they here to stay

Oh, I yesterday

Suddenly, I'm not half the man I,

There is a shadow me,

Oh, yesterday came suddenly

Why she go I don't know,

She wouldn't stay

I said something wrong now I

Yesterday

Yesterday, love was such an easy game to play,

Now I need a place to

Oh, I yesterday.

Activity 4: At this time the teacher asks the students which word mostly repeated in the song. Sample answers are yesterdays and suddenly.

After Listening

In this section, the teacher asks some comprehension questions about the song. Below are the samples of questions.

1. How does the singer feel?
2. What happened to make him feel like this?
3. What do you think will happen next?
4. What is the relationship between the people in the song?

2.5.3. Guessing the Meaning from the Context

Words last in students' memories when their meanings are deduced from the context with help of other known words in the text. Hence, the meanings come by the way of thinking and guessing.

Guessing from the context is the most important of all sources of vocabulary learning. Wilf (1988:108) confirms this, stating that it is hard to comprehend the meaning of words when they stand alone. When the same words are used in phrases, sentences, or paragraphs, it is easier to understand their meaning. For example, when the word *ruler* stands alone it has two meanings. First a ruler is someone who governs. and second, a ruler is a marked strip of wood, metal or plastic used for

drawing lines. When the word ruler is used in a sentence, you can tell which meaning fits.

Caesar was the **ruler** of Persia.

Use a **ruler** to make lines on your paper.

Wilf (1988:109) suggests the following steps to help us find word meanings in context:

- Use what you know about other words in context to form a picture in your mind.
- See which of the word's meanings fits in that picture.
- Check to see if your meaning fits in context by using it in place of the unknown word.

Candlin and al (1988:63) also classify guessing vocabulary from the context as the most frequent way to discover the meaning of new words. In order to guess the vocabulary from the context we have to look for context clues. Context clues can provide the meaning of a new word. For example:

Synonym in apposition: Our uncle was a *nomad*, a permanent wanderer who never could stay in one place.

Antonym: While the aunt loved Marty deeply, she absolutely *despised* his twin brother Smarty.

Cause and Effect: By surrounding the protesters with armed policemen, and by arresting the leaders of the movement, the rebellion was effectively *quashed*.

Association between an object and its purpose: The scientist removed the *treatise* from the shelf and began to read.

Description: Tom received a new *roadster* for his birthday. It is a sports model, red with white interior and bucket seats, capable of reaching speeds of more than 150 mph.

Example: Mary can be quite *vulgar*; yesterday she blew her nose on the new linen tablecloth.

Eckstut and Sorensen (1992:71) state that “you can guess the meaning of a word if you look for clues which come before and/or after it and if you use the knowledge you have about the world around you”. For example, students can guess the meaning of the words in bold italic in the below interview when using the phrases in the box to complete the statements which follow.

- a. begin to love someone very much
- b. spend time together on a regular basis
- c. have a friendly relationship
- d. does not change easily and is long lasting
- e. there is something about them that you like and that pulls you toward them
- f. about love
- g. their relationship is ended
- h. asking her to get married
- i. have a good opinion of them
- j. are usually married, live together, or are on date

1. People who ***get along*** with each other _____.
2. If two people are a ***couple***, they _____.
3. If a woman thinks about a man's ***proposal***, she thinks about his _____.
4. When you ***fall in love***, you _____.
5. When two people are ***going out***, they are usually of the opposite sex and they _____.
6. When two people ***break up***, _____.
7. A ***romantic*** movie is a movie _____.

Nunan (1998:121) argues that instead of learning lists of vocabulary items the focus in class should be on encouraging students to develop strategies for inferring the meaning of new words from the context in which they occur. The suggestions are:

1. **Word elements such as prefixes, suffixes and roots.** The ability to recognize component parts of words, word families, and so on is probably the single most important skill a student of reading can have.
2. **Pictures, diagrams and charts.** These clues, so obvious to the native speaker, must often be pointed out to the EFL student. S/he may not connect the illustration with the item that is giving him/her difficulty.
3. **Clues for Definition.** The students must be taught to notice the many types of highly useful definition cues. Among these are:
 - a) Parentheses or footnotes, which are the most obvious definition clues.
 - b) Synonyms and Antonyms usually occur along with other clues: that is, is clauses, explanations in parentheses, and so on.
4. **Inference clues from discourse,** which are usually not limited to one sentence:
 - a) Example clues, where the meaning for the word can be inferred from an example, often use physical clues such as *i.e.*, *e.g.*, and *for example*
 - b) Summary Clues: from the sum of the information in a sentence or paragraph, the students can understand the word.
 - c) Experience clues: the reader can get a meaning from a word by recalling a similar situation s/he has experienced and making the appropriate inference.

5. **General Aids**, which usually do not help the student with specific meaning, narrow the possibilities. These include the function of the word in question.

Nation (2001:250) lists the most important ways in which the teachers can help students improve learning from context as follows:

1. helping them to find and choose reading and listening material of suitable complexity
2. encouraging them to read a lot and helping them gain a lot of understandable spoken enter
3. improving their reading skills so that they read fluently and with good comprehension
4. providing training in guessing from context

2.5.4. Using Pictures

Things that we see play an important part in affecting us and giving us information. Dawson (1984:27) states that an interesting picture makes a strong visual impact and increases a lot of useful oral work. They are very easy to be found from magazines and illustrated advertisements for special vocabulary from different areas such as: kitchens, cars, clothes, interiors, plants, etc. Ruth and Stuart (1986:103) define pictures as tools that can be used for presentation, practice, revision and testing.

Wright (1989:17) characterizes the roles of pictures in language learning as follow:

1. Pictures can motivate the student and make him/her want to pay attention and want to take part.

2. They contribute to the context in which the language is being used. Pictures bring the world into the classroom.
3. The pictures can be described in an objective way or interpreted or responded to subjectively.
4. Pictures can cue responses to questions or cue substitutions through controlled practice.
5. They can encourage and provide information to be referred to in conversation, discussion and storytelling.

In order to decide whether to use an activity or not, whether the activity is a straight tool or a game making use of pictures Wright (1989:3) suggests the below criteria:

Easy to Prepare. Is it easy for you to prepare? If it is difficult in relation to what you will get out of it, don't do it.

Easy to Organize. Is it easy to organize in the classroom? If it is difficult to organize in the classroom then don't do it.

Interesting. Is it interesting to the students? Is it interesting to you?

Meaningful and Authentic. Will the language and the way you want the students to use it be authentic and essential to the activity?

Sufficient Amount of Language. Will the activity give rise to a sufficient amount of language in order to give good reason for its addition in the language lesson?

Nation (2001:85) identifies the advantage of using pictures as "learners see an instance of the meaning and this is likely to be remembered". Pictures involve student interaction and are useful in teaching real vocabulary items.

Betteridge and Buckby (1983:17) recommend the “**Happy Twins**” activity suitable for intermediate level. The teacher collects 16-20 pictures of single objects or people from magazines and mounts them on pieces of cards which are the same size. Then put all the pictures into an envelope, together for the instructions of the game. The procedure is like this:

The aim is for two players, working together to collect all the pairs. The pictures should be mixed and placed in a pile face down. Each player takes it in turn to pick up a card without showing it to the other. Then each player describes his/her own card without letting the other see. If they decide they have a pair, they place them on the table. If the cards are a pair they put them on one side. If the cards are not a pair they put them back into the pack.

2.5.5. Guessing the Meaning by Form

Form-based words are important because they make known sometimes hidden patterns of spelling in words that students already know; for example, the verb root pronounced 'seev' is spelt *ceive* (*receive, deceive, conceive*), and always corresponds to *ception* in the corresponding noun (*reception, deception, conception*). This shows that many words are made by adding a letter or group of letters to the end of a base form word, or by adding a letter or group of letters to the beginning of a base word. (Shepherd, 1987:61, 71).

Example: misconception

mis concept ion

If the students know the meaning of the base form of the word **concept**, the meaning of the prefix **mis** (incorrect), the meaning of the suffix **-ion** (noun suffix) in this way they can find the meaning of the entire word. (a wrong idea you have about something).

An understanding of word families also allows either the form or the meaning of unfamiliar words to be guessed with some confidence. For example, we can guess

that someone using a skate-board is a *skate-boarder* engaged in *skate-boarding*, and if we see the word *unteachability* we can guess from knowledge of other word families that it means 'state (-ity) of not (*un-*) being able to be (-*abil-*) taught (*teach*)'.

2.5.6. Using Mime and Gestures

When we speak and listen to each other, we communicate not only by using and interpreting spoken words or sounds but also by such visual signs as gestures - a nod or shake of the head - and by such facial expressions as a smile or a raised eyebrow. We use such physical gestures and facial expressions naturally and intuitively to convey and respond to meaning. Such visual signs or clues can also be consciously and deliberately crafted and used to communicate effectively, with or without spoken or written words, in vocabulary teaching.

Mime is a medium for performing in role without using words and focusing on conveying richness of meaning through visual language alone. Since it does not involve language, it may seem curious that it should be supported as an aid to language teaching and it is easy.

Maley and Duff (1982:39) suggest an activity which can be applied to learners of pre-intermediate upwards. In this game the teacher asks the class about various situations in which they might wish to complain, and lists these on the board. Individual students then mime these situations, and the class has to guess which situation is being mimed. The class can be asked to suggest what the speaker could say in such a situation.

2.5.7. Using Realias

Bringing realia (authentic objects from a culture), to the classroom helps teachers in providing comprehensible input in a second language. Students should be

allowed to touch, smell, and taste, if possible, prior to being exposed to the lesson, for best understandable input.

Allen (1983:25) proposes that for showing the meaning of an English noun, use real the object whenever possible. Students are not used to see telephones, toys etc. in the classroom but their originality arouses attention.

Dawson (1984:31) suggests using realias as they are valuable because:

- They can be used during the first five minutes of a lesson, while latecomers are still arriving, to stimulate oral work.
- They are a very direct way of teaching vocabulary.

This is a

Where is it?

What is it used for?

- They can be handled by the student, which gives rise to a whole range of descriptive vocabulary.
- They can be used as a way of introducing a whole range of associated vocabulary.
- They can be used to build up dialogues (for example, with a ticket and a program)
- They can be used in communicative activity as a support.

Allen (1983:24, 25) states that real objects are better than pictures whenever we have them in the classroom. In order to show the meaning of personal objects such as eyeglasses, sunglasses, wallets, watches, handbags, umbrellas which often can be found in the classroom or, other objects which can be brought easily in the classroom such as can opener, a pair of scissors, a box of paper clips, a toothbrush, a toothpaste using the real object is the best way of teaching.

2.6. Vocabulary Testing Techniques

If we are going to teach vocabulary to our students, at some point along the way we will want to test them to see how they are progressing. A vocabulary test aims to measure the comprehension and production of both active and passive vocabulary.

Schmitt (2000:163) emphasizes that if vocabulary is regarded as an important part of language knowledge, then it obviously needs to be assessed in some way. Beginners' vocabulary may be about the only characteristic of language they know well enough to be tested.

In order to learn words more easily teachers should develop techniques for different words for learning. According to Celce-Mursia and Rosensweig (1979:247) different techniques are needed for teaching different item types: e.g, idioms, abstract words etc. Therefore, you will have not to rely on just on one technique; you might better apply two or more techniques responding to different learners so they can react to it in their own ways. And in order to test the knowledge of the students we might apply do different testing techniques.

Read (2000:7) outlines two contrasting perspectives on the role of vocabulary in language assessment. First stating that it is sensible to construct tests that show whether learners know the meaning and usage of sets of words and second, that vocabulary must be assessed in the context of a language-use task.

Madsen (1983:12) presents four general kinds of vocabulary tests. As for beginners limited response items are presented and here the students are required to do simple physical actions like pointing at something or a simple verbal answer such as "yes" or "no". The second, multiple choice completions are presented in which the students choose one of the items given to complete the sentence. Then, multiple choice paraphrase in which a sentence with one underlined word is given and the students choose which of the four options is the closest in meaning to the underlined item. Finally, simple completion item is presented in which the students write the missing part of words.

Heaton (1989:79) classifies multiple choice items as useful vocabulary items, matching items as useful items for testing vocabulary in context and the completion items for testing the ability to produce appropriate words.

Lado (1961:188) also categorizes multiple choice items as the most successive vocabulary test providing an appropriate context for the lexical problem in the lead. However, he suggests pictures as alternatives to test children, adult beginners and illiterates.

In addition, Hughes (1989:146) states that a clear knowledge of vocabulary is crucial to the development and demonstration of linguistic skills, characterizing this as a main reason for testing vocabulary.

Schmitt and McCarthy (1997:303) place the vocabulary in language tests as primary goals of language testing, because the vocabulary knowledge is accepted as a basic component of foreign language proficiency and it is natural to expect assessing.

Bachman and Palmer (1996:8) state that being competent in language testing will help readers to avoid some of the misconceptions and to develop a set of reasonable expectations for any given language test. Language teachers need to be able to either make informed judgments in selecting appropriate language tests or to plan, construct, and develop appropriate tests of their own.

At this study several vocabulary testing techniques will be viewed. The major techniques that are going to be discussed are: multiple choice items, completion items, limited response items, matching items, items involving synonyms, definitions, associated words, integrative testing of vocabulary with other language skills, characteristics of a good vocabulary test, and things to be done after tests application.

2.6.1. Subjective Language Testing

Traditionally, teachers tend to feel that objective tests are the best way in evaluating the success of their students for they are easier to score. However, evaluating complex learning outcomes requires a much more complex and productive kind of testing. At this very point the essay type questions and essay tests present themselves as completely functional and helpful devices in testing.

Essay type testing is a widely used testing method in traditional education systems. Sometimes they are called as “Composition Type Tests” as Turgut (1995:47) points out, and they are generally performed by dictating or giving written questions to the students and asking for written answers, which is a very popular testing type among the teachers of almost every subject. In this kind of tests, students are required to think about the answers, plan to write, organize, write and then check what has been written before the end of the time given.

Turgut (1995:48) lists the main characteristics of an essay type test as follows:

- In essay type tests the learner should think and write their answers themselves. It's known that in some types of objective tests students are required to decide whether a given sentence is true or false or they are required to choose the answer among given answers. In the light of this data it can be said that essay questions assess more complex learning outcomes and skills.
- In essay type tests, students have to give their answers in a written form. Because of this, in every essay type test, good timing, pacing, suitable handwriting and the like qualities of the students interfere with the evaluation of the knowledge.
- In essay tests, it's not always possible to limit the questions. Every learner has the freedom to write his answers as he wishes. In most cases, even limiting the answers cannot solve the problem because it's not usual to

come across with students who tend to broaden the subject if they don't know the exact answer.

- It's not always possible to categorize the answers of questions as wholly right or wrong. Mostly, answers fall into the groups as partly acceptable, partly relevant and so on. Therefore, it's thought to be hard to ensure the validity and reliability of such tests.
- It's the teacher's task to decide whether the given answer is right or wrong. In other words the teacher is the only responsible person to assign a score for the given answers.
- A relatively long time is needed for the answering process. Most of the test time needs to be paid for the answering; therefore, the number of questions to be asked is limited.
- The scoring of the essay question type tests is much more difficult than the objective tests. This difficulty arises from the fact that the answers cannot be classified as wholly wrong or right and the responsibility is on the shoulders of the teacher.
- The time for preparing essay questions is relatively short but it takes a long time to score the answers. For this reason, such questions are not appropriate for large classes and for teachers whose teaching responsibilities are too much.
- It's difficult to determine the difficulty levels of the questions objectively. Because of this reason, the test statistics cannot be easily applied to this type of tests, which is a handicap for determining the reliability and validity of a test.

2.6.2. Objective Language Testing

The objective type item tests are an economical way of providing information from a student as it takes less time to answer an objective item. Read (2000:75) defines objective tests as “tests in which the learning material is divided into small units, each of which can be assessed by means of a tests item with a single correct answer that can be specified in advance”.

Chase (1999:33) defines objective tests as objective because they are scored objectively, including multiple choice, true-false, and matching items which means that all items have one right answer and no judgment- calls are involved in scoring these tests.

Objective item tests can be marked more easily and more accurately, and they may create an encouragement for students to build up a wide base of knowledge, skills and abilities.

Trice (2000:95) also states that in objective items no subjective judgment is required to score them. He regards true false items, multiple choice items, matching items and fill in the blank items as objective items.

Abbot and Wingard (1981:188) even define objective testing as “capable of being marked with 100% reliability”.

These tests are objective in the sense that they can be scored without requiring any judgment by the scorer as to whether an answer is correct or not. Read (2000:76) sees vocabulary becoming a component of objective language tests because:

- Words could be treated as independent linguistic units with a meaning expressed by a synonym, a short defining phrase or a translation equivalent.
- The list of the most frequent English words provided a large stock of vocabulary items that could be conveniently sampled to select the target words for a test.

- Objective vocabulary tests seemed to be valid indicators of language ability in a broad sense. Furthermore, the strong relationship between vocabulary and general intelligence is one of the strongest findings in the history of intelligence testing.

Taylor (2003:19) states that objective type tests may be used to supplement or compare results from standardized tests or where the standardized tests are not adequate for reasons of content, difficulty, scope, or sensitive cultural materials.

Constructing objective type tests requires some steps to be followed, they are as follows:

1. Write specifications for the test
2. Decision on the type of the objective test items to be used
3. Arrange the items in order of difficulty from easy to hard
4. Prepare directions for administering the test
5. Teachers should be skilled in using and communicating assessment results when making decisions about individual student' instructional programs relevant to their school and society
6. Teachers should develop skillfully valid student grading procedures, which use results from student assessments
7. To recognize wrong, dishonest, and if not inappropriate assessment methods and use of assessment information.

Taylor (2003:20, 21)

2.6.2.1. Multiple Choice Items

The multiple choice type of item has achieved its success in vocabulary tests. Although there are variations of format and style, in general the item consists of a stem containing the problem, one option representing the best response, and two to

five options representing the distractors to attract the student who doesn't know the best answer.

Mitchel (1992:5) defines multiple choice items as “choosing a single answer from among a small number of prepared possible answers”.

Gronlund (1998:53, 54) classifies multiple choice items as the most widely used and highly regarded of the selection type items. They consist of a **stem**, which presents a problem situation, and several **options**, which provide the possible solutions to the problem. The stem may be a question or an incomplete statement. The options include the correct answer and several plausible wrong answers named **distracters**. The role of the distracters is to distract those students who are uncertain of the answer.

Chase (1999:114) also defines multiple choice items as the most popular selection type item, suitable for use in assessing outcomes in almost all subject matters. It has two parts, the stem- defining the problem and the options- the possible responses to the problem.

Example:

1. Jane doesn't like to live in the country. She prefers _____ life.
a) subtle b) powerful c) tropical *d) urban

Taylor (2003:23) identifies multiple choice tests as one of the most flexible of objective items that may be used to assess a variety of information, such as comprehension, understanding concepts, applying principles, interpreting data, and assessing knowledge in most subject areas.

Lado (1961:188) defines multiple choice items success as the most impressive in vocabulary test. A better type of multiple choice items to test vocabulary is a multiple choice item that gives an appropriate context for the lexical problem in the lead.

Example: What we saw moving in the distance was a **seal**.

- a) stamp
- b) sentry
- c) *marine animal*
- d) wild pig

Madsen (1983:16) suggests some steps to be taken when writing multiple choice items:

1. Select the words to be tested
2. Get the right kind of sentence to put each word in (stem)
3. Choose several wrong words to put the right word with (distractors)
4. Prepare clear and simple instructions.

When the suitable words are selected, the next step is preparing the context for them. For the beginners more than one sentence is needed to help clarify the meaning. In order to check the meaning of the word “**brush**” you can prepare a two line mini dialogue.

Example:

“I want to paint too”

“All right use that _____ over there.”

*a) brush b) pencil c) broom d) spoon

Heaton (1988:54) also suggests some steps to be followed for writing multiple choice items, they are defined like:

- If the problem area to be tested is positioned in the options the stem should be kept simple. If the problem area is included in the stem, the alternatives themselves should contain only those vocabulary items which the test takers can understand.

- Each option should belong to the same word class as the word in the stem. If this step is followed there will be less danger of context providing important grammatical clues for the test takers.
- The correct option and the distractors should be approximately the same level of difficulty, because if the correct option is more difficult than the distractors the student will get the right answer by elimination.
- The item writer should limit the options the same general area of activity where possible.
- All the options should be approximately the same length.

Gredler (1999:92) suggests using of multiple choice items for assessing the recall of knowledge and the purpose of knowledge, together concept learning and rule learning.

A good multiple choice item needs time for constructing. Trice (2000:99-103) suggests some guidelines for constructing better multiple choice items.

- **Distractors should be reasonable.** The first rule of creating multiple choice item is that the distractors should be possible answers.
- **For beginning students three choices are enough.** Sometimes children need to be introduced to the idea of multiple choice tests as early as first grade using pictures as answers.
- **Don't connect items.** In tests where three or four items are related, if you miss the first item, you will miss all of the next ones.
- **Stop and think long and hard before using options such as "all of the above", "none of the above", "both a and b".** A question like this will confuse elementary and most middle school students.

- **Don't use patterns of answers.** Many students believe that they can discover the correct answers to multiple choice tests by looking for patterns.
- **Play attention to class patterns.** If you ask a question and most of the students miss it, there are three possibilities:
 - a) It was a hard question, although a good one.
 - b) You did not teach the material well enough.
 - c) There was something wrong with them.
- **Teach students how to approach multiple choice items.** If one of your objectives is for students to become better test takers, you should help them develop skills in approaching different kinds of test items.

The following rules for item writing are proposed as guides for the preparation of multiple choice items:

1. **Design each item to measure an important learning outcome.** The problem situation around which an item is to be built should be important and should be related to the intended learning outcome to be measured.
2. **Present a single clearly formulated problem in the stem of the item.** The task set out in the stem of the item should be so clear that a student can understand it without reading the alternatives.
3. **State the stem of the item type in simple, clear language.** The problem in the stem of a multiple choice item should be stated as precisely as possible and should be free of unnecessarily complex wording and sentence structure.
4. **Put as much of the wording as possible in the stem of the item.** Avoid repeating the same material in each of the alternatives.
5. **State the stem of the item in positive form, wherever possible.** A positively phrased test item tends to measure more important learning outcomes than a negatively stated item.

6. **Emphasize negative wording whenever it is used in the stem of an item.** When negative wording is used, it should be emphasized by being underlined or capitalized and by being placed near the end of the statement.
7. **Make certain that the intended answer is correct or clearly best.** There should be only one correct answer in the item.
8. **Make all alternatives grammatically consistent with the stem of the item.** The correct answer is usually carefully phrased so that it is grammatically consistent with the stem.
9. **Avoid verbal clues that might enable students to select the correct answer or to eliminate an incorrect analysis.** The most frequently found verbal clues in multiple choice items are:
 - a) *Similarity of wording in both the stem and the correct answer.*
 - b) *Stating the correct answer in textbook language*
 - c) *Stating the correct answer in greater detail.*
 - d) *Including absolute terms in the distracters (always, never, all)*
 - e) *Including two responses that are all comprehensive*
 - f) *Including two responses that have the same meaning.*
10. **Make the distracters plausible and attractive to the uninformed.** The distracters should be so attractive to the students who lack the knowledge called for by the item that they select one of the distracters in the preference to the correct answer.
11. **Vary the relative length of the correct answer to eliminate length as a clue.** This happens because of the need to qualify statements to make them clearly correct.
12. **Avoid using the alternative “all of the above”, and use “none of the above” with great caution.** The addition of “all of the above” makes it possible to answer the item on the basis of incomplete information.

13. **Vary the position of the correct answer in a random manner.** The correct answer should appear in each alternative position about the same number of times, but its placement should not follow a pattern that may be obvious to the student.
14. **Control the difficulty of the item either by varying the problem in the stem or by changing the alternatives.** Increase the item difficulty by increasing the level of the knowledge called for by making the problem more complex.
15. **Make certain each item is independent of the other items in the test.** Occasionally information given in the stem of one item will help the students answer another item.
16. **Use an efficient item format.** The alternatives should be listed on separate lines, under one another.
17. **Follow the normal rules of the grammar.** If the stem is in question form, begin each alternative with a capital letter and end with a period or other appropriate punctuation mark
18. **Break any of these above mentioned rules if it will improve the effectiveness of the item.**

Gronlund (1998: 60-77)

Taylor (2003:24) states that constructing multiple choice items is a complex process that involves understanding of and competencies in the subject area being measured or assessed. The approaches to be followed while constructing are:

1. Differ the position of the alphabet representing the correct answer. Place the right answer in different positions.
2. Don't contain answers that are so discrete to the question that the answer can without difficulty be known.

3. Don't list wrong answers that appear true enough to be the right answer.
4. Don't give clues that will give away the correct answer, such as having the right answer being consistency long or short, or having one choice or alphabet letter being continuously correct.

Gredler (1999:94, 95) points out that the first thing to be considered in developing multiple choice questions is to review the objectives that deal with key concepts and principles and then develop examples and situations that require the application of the defining criteria for the concepts and the exacting rules. Then, the stem should cause a problem to be dealt by the student and should not be confused with unrelated information. Finally, the given alternatives should be brief and nearly the same length and equally plausible.

2.6.2.1.1 Advantages of Multiple Choice Items

The multiple choice items are the most widely used and highly regarded of the selection type tests. They can be designed to measure a variety of learning outcomes such as language skills and elements providing the highest quality items.

Khodadady (1999:109-113) lists the advantages of multiple choice items as follows:

1. ***Multiple choice items are scored objectively.*** In answering cloze tests, test takers supply single word or short phrases, but in answering open ended questions, they have to compose complete sentences, paragraphs or texts.
2. ***Multiple choice items are administered easily.*** These items are not only objective in terms of scoring but are administered easily to a large number of test takers in one single testing session.
3. ***Multiple choice item tests are the most effective measures of linguistic competence, achievement and program effectiveness.*** The most direct means of evaluating linguistic competence exists in multiple choice items

that require the student to make a comparative judgment of the grammatical acceptability of the options given.

4. *Multiple choice items are reliable.*
5. *Multiple choice item tests are valid.*

Madsen (1983:23) states that multiple choice items help students see the full meaning of words by providing natural contexts. They put off word list memorization and scoring is easy and reliable.

Taylor (2003:23) states that there are several advantages of using multiple choice items:

1. Answers are objective and easily scored.
2. Items may be written to measure critical thinking skills, such as analysis, synthesis, and evaluation.
3. Choices are given to assist in minimizing guessing.
4. Items may be written to measure knowledge, comprehension, and application.

On the other hand, Chase (1999:123) lists the strengths of multiple choice tests like:

- Multiple choice tests can be written at all levels of the classification.
- Multiple choice tests can be used in assessing outcomes in almost all subject matter taught in schools.
- Multiple choice tests sample a large number of information points in a short time.
- Multiple choice items provide information that helps diagnose learning errors.
- Multiple choice items reduce guessing.
- Multiple choice tests are easy to score.

Gredler (1999:95) states that multiple choice items can assess students' skills identifying examples of key concepts in the lesson and in applying significant values. They may also be used to evaluate knowledge, mainly when the item asks the student to select the correct mixture of sentences.

2.6.2.1.2 Disadvantages of Multiple Choice Items

Although multiple choice items are one of the most widely used methods of assessment they do have some limitations such as constructing that requires time and practice; it's hard to find plausible distractors.

Read (2000:77, 78) summarizes the limitations of multiple choice vocabulary items as follows:

- They are difficult to construct, and require difficult field-testing analysis.
- The learner may know another meaning for the word, but not the one sought.
- The learner may choose the right word by a process of elimination, and has in any case a 25 per cent chance of guessing the correct answer in a four alternative format.
- Items may test students' knowledge of distractors rather than their ability to identify an exact meaning of the target word.
- The learner may miss an item either for lack of knowledge of words or lack of understanding of syntax in the distractors.
- This format permits only a very limited sampling of the learner's total vocabulary.

Gredler (1999:95) states that constructing multiple choice items requires practice and the best time for the teacher to get ideas for questions is during

preparation time for teaching the lesson, when various approaches to the subject matter are being explored.

Madsen (1983:23) also states that it is rather difficult to prepare good sentence contexts that clearly show the meaning of the word being tested. In multiple choice items it is easy for students to cheat by copying what others have circled.

Khodadady (1999:114-123) criticizes multiple choice tests as follows:

- Writing multiple choice is very difficult
- Writing multiple choice items is time consuming
- Multiple choice item tests are not communicative
- Multiple choice items do not show understanding or misunderstanding of the text
- Multiple choice item tests put words in the test takers' mouth
- Multiple choice item tests measure only recognition knowledge
- Guessing may have a considerable but unknowable effect on test scores
- Multiple choice item tests strictly restrict what can be tested
- Backwash may be harmful
- Cheating may be facilitated
- Superficial mastery of the subject matter may be encouraged

Chase (1999:125-127) states that the most common disadvantages of multiple choice items are:

- It is difficult to write multiple choice items at higher order process level.
- Distractors are not equally plausible.
- The reading demand of multiple choice items is often high.
- Guessing does have some influence on scores.

2.6.2.2 Completion Items

Completion items require very limited writing (a word or phrase) and they can be scored somewhat objectively. Completing these items requires recalling specific facts and the ability to comprehend overall relationships.

Read (2000:173) defines completion items as consisting of a sentence from which the target word has been deleted and replaced by a blank; the function of the sentence is to provide a context for the word and perhaps to cue a particular use of it.

Chase (1999:145) divides completion items into two formats: The first one uses statements from which one or more key terms have been dropped and a blank inserted to take their place. The second item may be a question, the answer to which is a single word or phrase. Since students have to supply the answers, completion items are identified as supply type tests. The options are not provided and students have to give their own responses.

Taylor (2003:25) suggests some steps to be followed when constructing completion items:

1. Do not copy information directly from the text; say differently the information to reduce the dependence on rote memory. Questions should be designed to encourage critical thinking skills.
2. Limit the number of blanks required to complete. Including many blanks confuses students.
3. If words rather than phrases are deleted, scoring of the completion items will be more objective.
4. The blanks included in completion items should be the same size, in order to prevent clues about the length of the correct word.
5. When there is more than one correct answer, provide for them in the correct answer.

6. Watch against giving clues to students by making sure completions do not depend upon grammatical form or textbook expressions.

In addition, Madsen (1983:16) states that this type items make the student depend on context clues and sentence meaning, constructed by deleting a word from a sentence. The suggested steps when writing completion items are:

- Select the words to be tested
- Get the right kind of sentence to put each word in.
- Prepare clear and simple instructions

Heaton (1990:86, 87) also ranges completion items into one- word completion answers to the completion of sentences. One word completion items are also defined as more objective. In other words, there is either only one correct answer for each blank or a limited number of correct answers. In contrast, items requiring finishing a sentence by writing several words are more subjective and more difficult to score.

Chase (1999:146) states that completion items focus largely on assessing recall rather than recognition. This item type gives a few clues, but for the most part the students most depend on bringing to mind the topic. However, there are some subjects that teachers want students to bring to mind without prompting. Among these items are translations of vocabulary where completion items are valuable for assessing unprompted knowledge of these items.

Heaton (1988:62, 63) recommends the following two types of completion items to be used for the testing of vocabulary. Tests which present such items in a context are usually preferable to those which rely on single words or on definitions.

Type 1

Read through the following passage containing a number of incomplete words. Write each completed word on your answer sheet at the side of the appropriate number. (Each dash represents one letter.)

Snakes are one of the (1) d-m-n- - t groups of (2) r-pt- - - -: there are at least 2000 different (3) sp-c- - s of snakes (4) sc-t- - - d over a wide area of the earth.

Type 2

(a) Complete each blank with the most appropriate word to replace each number in the text.

Rosnah: What's the (1) today? (1)
Mohamed: It's the seventh. (2)
Rosnah: At what (2) does the concert start?
Mohamed: Seven o'clock, I think. Just a moment.
I made a note of it in my (3). (3)

Type 3

(b) Complete the following paragraph on problems caused by weightlessness by writing ONE word in place of each blank.

Increasing is now being one of the effects of weightlessness on man. For, scientists are the role of gravity on the way cells function.

In order to improve the validity of the completion items and avoid accidental errors the following procedures should be followed when constructing completion items:

- If at all possible, completion items should be answerable with a single word. The subsequent example illustrates this issue.

Not so good: *The purpose of the carburetor in a gasoline engine is to _____?*

Better: On a gasoline engine, what is the name of the part that vaporizes the gasoline? _____ .

With a single word answer there is no judgment call required for scoring. In the first item, different students will use different phrasing which reduces the reliability of the test score.

- Avoid lifting statements directly from the textbook. This may encourage rote memory which is considered wrong by the educators.
- Make the intent of the statement clear so that one and only one answer will be correct.

Example:

Poor: Abraham Lincoln was born in _____ .

Better: Abraham Lincoln was born in the year _____ .

In the first item several answers are true, and students who give logical answers may say "I didn't know what you wanted there".

But in the second item the purpose of the question is known and the students are supposed to write the year when Lincoln was born.

- The blanks in a completion item should preferably be placed at the end of the statement rather than at the beginning. If the blank is at the beginning of the statement it tends to leave the student without direct orientation.

Example:

Poor: _____ is the name of the town was Dwight Eisenhower is buried.

Better: What is the name of the town in which Dwight Eisenhower is buried?

_____.

At the first example none knows what direction the question will take. But if the given eliminated word is at the end of the sentence as in the second example the subject and the predicate of the sentence is given to the reader before the answer is called for.

- Do not give clues by making some blanks longer or shorter than others. Even you don't intend to give clues by the length of the blank; some students will interpret longer blanks as meaning that a longer answer is required.

Chase (1999:147, 14)

2.6.2.1 Advantages of Completion Items

1. *Completion items are easy to write.* You can access many objectives for a unit of instruction with the completion test, requiring only a word or short sentence for each item
2. *Guessing is less possible than in other types of tests.* There is only one correct answer for a well written completion item. The student is not given options among which to guess. Students hate to leave blanks on their tests, so that the test taker may write an answer of which he or she is not certain. but this is not the same as taking a chance on one of several given options.
3. *They are adaptable to many types of subject matter.*
4. *Completion test scores are quite reliable.* In order to be reliable the items must be well constructed and the scoring done in a systematic way.

2.6.2.2 Disadvantages of Completion Items

1. *It is difficult to apply completion items Higher Order Process.* The limitation of answers to a single word (or a few words) makes it difficult, but not impossible, to use this item analysis, synthesis or evaluation. Higher Order Processes usually require some elaboration or explanation. In tests that require short answers this is hard too. However, assessing at Higher Order level can be done better with other kinds of tests items.
2. *Scoring is monotonous and contaminated by spelling errors and exchange wording.* Some students will put limitations on their answers, while others will elaborate. Judgments must be made as to the adequacy of many answers. And shall I give limited credit if the answer appears to be there but if misspelled? If so, how much credit shall I give? What criteria shall I use to make this decision?
3. *Some topics must be left out of the test because they require longer responses.* Topics that require more than a word or phrase are best tested using the essay format.

Chase (1999:149, 150)

2.6.2.3. Items Involving Synonyms

Hughes (1989:148) states that making students guess the meaning of vocabulary items is something which we would probably wish to encourage. However, writing successful items has its own difficulties. Writing items that involve synonyms is one of the ways to involve a number of different operations.

Example:

Choose the alternative (A, B, C, D) which is closest in meaning to the word on the left of the page.

gleam a) gather b) shine c) welcome d) clean

The writer of this item has probably chosen the first alternative because of the word *gleam*. The fourth may have been chosen because of the similarity of its sound to that of *gleam*.

However, the item could have a common word as the stem with four less frequent words as options:

- shine* a) malm b) gleam c) loam d) snarl

Heaton (1988:61) also suggests involving synonyms in testing words.

Example:

(a) *Write in each space the best word to replace the words underlined in each sentence.*

- | | |
|---|-------------|
| 1. Tom went <u>at once</u> to the docto's. | Immediately |
| 2. <u>All of a sudden</u> there was a loud cry. | |
| 3. I <u>came across</u> an interesting book. | |
| 4. The boat is over fourteen feet in | |

(b) *In each space write one word that means almost the same as the word on the left. The word you write must rhyme with the word on the right.*

Example: early soon moon

1. purchase die
2. miserable bad

2.6.2.4 Matching Items

The basic matching task requires learners to make a connection between target words and their synonyms or definitions. By itself, it is a recognition rather than a bringing to mind task, focusing on the basic word meaning. The standard practice is to group a number of words together with a large number of definitions.

Read (2000:171)

Example:

Next to each word, write the number of its meaning.

region	___	1. position in relation to others
status	___	2. gas that we breathe
cell	___	3. part of a country
oxygen	___	4. smallest part of living things
skeleton	___	5. useful liquid
		6. set of bones in the body
		7. exciting event

Gredler (1999:88) states that matching items present several problems accompanied by a single list of possible answers. The item typically consists of two parallel columns of information and the task of the student is to identify the pairs that are related on the basis of the association described in the instructions.

Chase (1999:139) defines matching items as items consisted of a column of terms on the left side of the work sheet called *stimuli*, and one more column in the right called *responses*, related to the stimulus.

When constructing matching items, the teacher matches a list of words, names or phrases against another list. Matching items may be used to test the relationships between statements. Constructing matching items also requires the following:

- Limiting the number of items in the list; five or seven items are recommended.
- Choosing information from one academic area, so that item in Column 1 has several matches in Column 2.
- Arranging names in alphabetical order and dates and numbers in sequence order to save the student time in completing the test.
- Try to abstain from giving clues.

Taylor (2003:24, 25)

Trice (2000:103-106) on the other hand suggests the following rules for making up a successful matching items.

1. Use matching for at least four or five items that go together.
2. Do not have a one to one correspondence between stems and answers. If you give matching items so students can eliminate answers each time they are used, you will allow too much for what students know.
3. Do not use more than 15 items in a matching bunch. Using large numbers of items together does not measure ideas and knowledge, but how well a student takes tests.
4. Make sure all your answers are of the same class. Matching items work better when all of the answers are similar: all people, all books, all adjectives and in this way context clues are eliminated.

Heaton (1988:59) also recommends that it is much more efficient to test words from the same word class: nouns only, adverbs or verbs.

Example:

From the list of words given, choose the one which is most suitable for each blank. Write only the letter of the correct word after each number on your answer sheet.

- | | | |
|---------------|--------------|------------|
| A. completely | C. busily | E. quickly |
| B. politely | D. carefully | F. angrily |

‘Write (1) ‘the teacher shouted (2), but don’t waste time. You must get used to working (3) Please sir, a student said (4) “I’ve finished”. ‘No, you haven’t’, answered the teacher. ‘You haven’t (5) finished until you’ve ruled a line at the end. ‘Meanwhile, the boy sitting up next to him was (6) engaged in drawing a map.

Read (2000:172,173) also proposes some characteristics of the design of matching items considering them as important:

- Add one or two extra definitions to prevent situation where the student knows four of the target words and can then get the fifth definition correct by eliminating, without knowing what the word means.
- Supposing the focus of the test is on knowledge of the target words, the definitions should be easy for the learners to understand.
- If the purpose of the test is just to assess knowledge of word meaning, then all of the target items should belong to the same word class because the students may be able to match the words and definitions on the basis of form as well as meaning.

2.6.2.4.1 Advantages of Matching Items

Gredler (1999:91) as the main advantage of matching items defines their power in assessing concept learning outcomes. They can require that the student make careful judgments in applying the criteria for concept membership.

Gronlund (1998:85) lists the advantages of matching items as:

- A compact and efficient form is provided where the same set of responses fit a series of item stems.
- Reading and answering time of matching items is short.
- This item type is easily constructed if converted from multiple choice items having a common set of alternatives.
- Scoring is easy, objective and reliable.

Chase (1999:143, 144) looks closer to the advantages of matching items illustrating that:

- *Matching exercises offer a short procedure for assessing learning of a set of related ideas, facts and similar content.* This type of test is made up of single words or short phrases in one column which are to be matched with single words or short phrases in a second column.
- *Matching calls for a minimum of reading.* This test item can provide you with a set of related facts quickly and requires little reading on the student's part.
- *Matching exercises are relatively easy to construct.* Once the list of related concepts with which the exercise is to be made is ready, the test writing time is short.
- *Scoring is easy and reliable.* Scoring matching items is quick and simple, especially if students are asked to write the letter of the response on a line before the appropriate stimulus.

2.6.2.4.2 Disadvantages of Matching Items

Read (2000:173) makes a criticism of matching items saying that it presents each target word in isolation and the decision about presenting words in isolation or context should be made by reference to the purpose of the assessment.

Taylor (2003:24) recommends us to consider factors in constructing matching items since they may measure memory rather than comprehension.

On the other hand, Gronlund (1998:85) also lists the limitations of matching items like:

1. This item type is largely restricted to simple knowledge outcomes based on associations.
2. It is difficult to construct items that contain a sufficient number of homogenous answers.

3. Defenselessness to irrelevant clues is greater than in other item types.

Chase (1999:144) elaborates the main disadvantages of matching items as follows:

- *Matching tests depend upon having a number of related stimuli for an exercise.* This happens as a result of the type of subject matter in which are making an assessment and the skill of the testwriter and not of the procedure itself.
- *Matching is primarily suited to the foundational content level.* It is easier and less time consuming to make matching items that are at the factual level than at higher order processing levels.
- *Matching is not suited to wide sampling of a unit's content.* A wide sampling of the unit of instruction will probably require use of multiple choice or true false items.

2.6.2.5 Definitions

Harris (1969:51, 52) defines this type of tests as “classic” type of vocabulary item, consisting of a test word followed by several possible definitions.

Example:

nap

- A. a brief sleep
- B. a happy song
- C. a sharp rock
- D. a short meeting

A variation of the previous overturns the order and places the definition first.

Example:

a brief, light sleep

- A. nap
- B. yawn
- C. stroll
- D. hug

The above mentioned both definition item types have the advantage of being 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.

In contrast, Heaton (1988:62) states that these item types are of very little use. They test writing ability to a knowledge of word meaning. In addition, it is too hard even for native speakers to produce sentences showing the meaning of words. A student may be familiar with the meaning of the word and may use it correctly, without being able to express this meaning clearly in a sentence.

Example:

- (a) *Use each of the following words in a sentence so as to show the meaning of the word*

economy politics industrious

- (b) *Explain the meaning of each of the underlined words in the following phrases.*

an archaic word a fortuitous event

Hughes (1989:150) also states that not all items can be identified exceptionally from a definition.

2.6.2.6. Associated Words

Heaton (1988:58) states that difficulties arising from the testing of collocations are avoided by the testing of word sets. In such tests the student's knowledge with a variety of associations is measured.

Example:

Type 1: Recognition

Read each of the following lists of four words. One word does not belong in each list.

son	happy	arrive
father	married	depart
boy	engaged	go away
brother	single	leave

Type 2: Production

Each group of words is related to a particular subject. Write down the particular subject which is connected with each group of words.

hand	theatre	volume	nursery
wrist	sister	track	lift
dial	bed	head	slope
face	ward	spool	snow
(= watch)	(=hospital)	(tape recorder)	(=skiing)

Read (2000:180) defines associated words as task which involves presenting subjects with a set of stimulus words one by one and asking them to say the first related word that comes into their head. A developed word association task that would require the learners to select response rather than supplying their own is shown like:

edit

arithmetic	film	pole	publishing
revise	risk	surface	text

team

alternative	chalk	ear	group
orbit	scientists	sport	together

There are three main types of relationship between target word and associate:

1. Paradigmatic: The two words are synonyms or at least similar in meaning, perhaps with one being more general than the other.
2. Syntagmatic: The two words often occur together in a phrase, that is, they collocate.
3. Analytic: the associate represents one aspect or component of the target word and is likely to form part of its dictionary definition.

Read (2000:181)

2.7 Selecting Words to Test

The selection of vocabulary test words is relatively easy in achievement tests, because these can be drawn from the particular textbooks that have been used in class.

Gairns and Redman (1986:54) suggest four main sources of selecting words:

1. Through the course book. This will include the written and spoken texts, activities for the presentation and practice of grammatical structures, testing exercise, and so on.
2. Through supplementary materials provided by the educational institution or selected by the teacher himself or herself.

3. Through the students. A great amount of surprising, and random items will certainly surface from student enquiries, queries, and errors.
4. Through specific vocabulary activities designed by the teacher for his particular group of students.

Harris (1969:48) states that the first decision to be made is whether to test the student's active or passive vocabulary-that is, the words they should be using in their speech and writing or those they will need only to comprehend, especially in their reading.

On the other hand, Madsen (1983:12, 13) proposes the following way to choose words you want to include in your test:

- Somehow we need to find out which words our students need to know. This problem can be solved by a course syllabus that tells what words must be learned.
- Another way is to record the words that students use wrongly. These may be test items.
- Finally, do not overlook the words and phrases needed to run the class, such as "Take your seat" or "the assignment for tomorrow". These are useful test items at the beginning level.

Even though the dictionary may be used in the selection of test words, usually it is more suitable to use word lists based on frequency counts of lexical items occurring in actual samples of the language.

Harris (1969:49, 50) suggests to be prepared for word counts' several weaknesses:

1. Word counts are usually based on the written language only, for this reason, many words that are really common in the oral language will receive low occurrence scoring in the word lists.

2. The word lists classify words according to relative frequency rather than total difficulty, and the two are by no means always alike.
3. Word frequency in English does not serve as a good guide to the probable difficulty of lexical items for which there are similar forms in the foreign student's native language.
4. Some of the word lists do not make a distinction among the various meanings of a word.
5. Unless the word lists are based on very recent surveys of frequency, they are likely to contain items whose status now is quite different from what it was at the time the data were collected.
6. Some word lists are based on a sample of written materials quite unlike those which the typical foreign learner of English is likely to have read.

2.8. Integrative Testing of Vocabulary with other Language Skills

Comprehensive measures are particularly suitable for assessment procedures in which vocabulary is embedded as one component of the measurement of a larger construct, such as communicative competence in speaking, writing, reading or listening comprehension.

2.8.1. Testing Vocabulary with Reading

It is a fact that there is a rule stating "Teach what you test, test what you teach". Therefore, it is a good idea to make the students aware of what we are testing in a reading test. Traditionally, reading classes are conducted by giving a text to the students to read loudly once or twice, and then answer some comprehension questions. This kind of a teaching is neither adequate nor useful. (Pumfrey, 1989:132).

Choosing a text for reading tests is as important as preparing the questions because the text that is used to test reading will, surely has an impact on the results.

Traditionally, reading tests are thought to be difficult and boring. In fact, this prejudice stems from the fact that most teachers insist on using lengthy specific texts in reading tests. (Heaton, 1975:123). However, if the text type is varied then it can be observed that students feel more confident, and of course, motivated in reading tests. In other words, variety should be provided in a reading test.

Valette (1977:132) points out that “in preparing discrete tests based on vocabulary, the teacher should take care to see that the comprehension problem lies in only one place, either in the stem or in the options”. If the difficult vocabulary item is in the stem, all students should understand the meaning of the options.

Madsen (1983:76, 77) recommends informal techniques of evaluation such as exercises and individual practice. Early testing is often limited to reading subskills, the most common being vocabulary. The most typical of this skill is the “same-different” technique.

Example:

sad	sat	S	<u>D</u>
red	red	<u>S</u>	D
meat	meet	S	<u>D</u>

Aldersen (2000:207) states that “gap- filling tests are somewhat different in that the tests constructor does not use a pseudo- random procedure to identify words for deletion: she/ he decides which words to delete, but tries not to leave fewer than five or six words between gaps”.

Example:

The fact is that one cloze test can be very different from another cloze test based on the same text. 1)..... pseudo-random construction procedure guarantees that 2) test writer does not really know 3) is being tested: he/she simply 4) that if enough gaps are 5), a variety

of different skills 6) aspects of language use will 7) involved, but inevitably this is 8)

Madsen (1983:93) suggests using a multiple choice cloze format so that students might rely on the recognition of the needed word. Here the correct underlined word is to be circled.

Is your home safe? Do it / *you / them / each protect your family? Are you safe of / for / with / *from the fire or falling? Do you guard against watching / knowing / *eating / sending unsafe food or drinking poison? It perhaps / *is / always / so important to keep your home safe *and / since / because / for develop good safety habits.

2.8.2. Testing Vocabulary with Writing

Writing can be divided into more specific “discrete” elements such as grammar, vocabulary, spelling, punctuation, capitalization etc. The only way to test these formal elements is objective tests. However, it should be mentioned that the results don’t prove how good a writer is or what a student can do in writing. But testing writing with ignoring these components will be unreliable.

Read (2000: 197) states that "one way in which we can assess the written production of learners is by calculating various statistics that reflect their use of vocabulary".

The ability to write in a foreign language presupposes knowledge of the lexical units of the language. A lexical unit is a word or group of words possessing a specific meaning.

Valette (1977:223) states that vocabulary test should be constructed around lexical units rather than words. Consider the following example:

Joe belongs to the human *race*.

Joe went to the dog *race*.

Joe and Sam *race* each other.

Joe and Sam *race* their turtles.

Written vocabulary tests evaluate the student's ability to recall and to produce lexical units in the new language. Vocabulary items may be developed for use in conjunction with visual, printed, or oral clues.

2.8.3. Testing Vocabulary with Speaking

Testing the ability to speak is a most important aspect of language testing. However, at all stages beyond the elementary levels of mimicry and repetition it is an extremely difficult skill to test, because it is too complex skill to permit any reliable analysis.

Valette (1977:132) states that "in testing vocabulary in the classroom, the teacher should always encourage the students to speak in natural utterances". Visual cues are very helpful in testing spoken vocabulary, especially at the elementary and intermediate levels.

The visual cue for a spoken vocabulary test can also be provided by slides of the target country and of its contributions to world civilization (paintings, sculpture, drama). The slide is projected on the screen and the teacher uses a long pointer to indicate specific aspects of the photograph. This type of vocabulary test has the additional advantages of increasing the student's familiarity with various aspects of the target language.

A speaking test at intermediate level will have to consider the following criteria:

- **Accuracy:** Pronunciation still obviously influenced by L1 though clearly intelligible. Grammatical/ lexical accuracy is generally high, though some errors which do not destroy communication are acceptable.

- **Appropriacy:** Use of language generally appropriate to function. The overall intention of the speaker is always clear.
- **Range:** A fair range of language is available to the candidate. He is able to express himself without clearly having to search for words.
- **Flexibility:** Is able to take the initiative in conversation and adapt to new topics or changes of directions- though neither of these may be consistently manifested.
- **Size:** Most contributors may be short, but some evidence of ability to produce more complex utterances and to develop these into discourse should be manifested.

2.8.4. Testing Vocabulary with Listening

Tests are always made for a purpose, and the purpose has considerable influence on the construct we want to measure. There are a number of purposes for testing listening.

- **General language proficiency:** More than 45 % of our total time communicating is spent on listening, which suggests the great importance of listening skills in overall language ability.
- **Representing oral skills:** Listening is not only included in language tests for its own sake, it is also used as a substitute for other skills.
- **Assessing achievement:** There are teachers who feel that if they are teaching listening, they ought to assess it. This is basic achievement testing, and tests are commonly used to provide grades, give credit for course completion or to determine whether the students are ready to proceed to the next level of instruction.
- **Diagnostic testing:** The other main purpose for testing is diagnosis, as one of the most imperative aim of education to find out where student

knowledge is deficient so that teaching, can more effectively target their needs.

Underhill (1997:32)

Heaton (1988:67) suggests the following activity, where the testees are required to choose the correct definition for the word they hear. Each of the items is given on tape while four definitions are printed on the booklet.

1. *Spoken*: cot - cot

- Written*: A) stopped and held
 B) a baby's bad
 C) pulled by horses
 D) a small pet animal covered with fur

2. *Spoken*: threw – threw

- Written*: A) made something move through air
 B) not false
 C) some but not many
 D) made a picture or diagram on paper

3. *Spoken*: bud – bud

- Written*: A) part of a tree or a flower
 B) a creature with wings
 C) something to sleep on
 D) not good

Read (2000:196) states that a vocabulary measure may provide an indication of how easy it will be for the test taker to understand a spoken text when it is used for assessment purposes. In the case of listening, it is not just the relative frequency of the content words that affects comprehension but also how concentrated they are in the text.

Listening texts must be chosen with the test specifications in mind. If we are interested in how candidates can cope with language intended for native speakers, then ideally we should use samples of authentic speech. In fact these can usually be readily found, with a little effort. Possible sources are: radio broadcasts, teaching materials and the recordings of native speakers.

Buck (2001:71) suggests filling in blanks item to test listening stating that “if we want to prevent students filling in blanks without actually understanding the meaning, we can ask them to fill in blanks on a summary of the passage”. The procedure is as follows: the students are given a summary of the passage they are going to hear, in which some of the important content words have been replaced by blanks. After looking at the summary for a while, students listen to the original passage. They should be discouraged from writing while listening since their task is to use their understanding of the passage to fill in the blanks.

2.9. Characteristics of a Good Language Test

Testing is an important and indispensable subject in teaching and learning process. It has a wide scope including uses of testing; types of testing; construction; administration; evaluation and interpretation of tests scores and testing certain skills.

However, these areas mentioned above do not mean anything without the requirements of tests. In short, requirements of tests constitute a vital aspect of testing. Without fulfilling the requirements of tests, a test is not a test and hence no value. All good tests possess three qualities: *reliability*, *validity* and *practicality*. That is to say, any test that we use must be dependable in the evidence it provides, appropriate in terms of our objectives and applicable to our particular situation”. Harris (1969:13)

2.9.1. Reliability

Even though most experts of the subject consider validity as being the most important of the three elements, it is more suitable to begin with reliability due to its feature of being a prerequisite for a valid test. For a test to be valid, first of all a test should be reliable as a measuring instrument and it is really difficult to comprehend the concept of validity without a basic understanding of reliability.

Harris (1969:13)

Harrison (1983:10) defines reliability as a tests' consistency, or the student's score should be the same (or as nearly as possible the same) whether the student takes one version of a test or another. Reliability also means the consistency with which a test measures the same thing all the time.

Reliability refers to the consistency of measurement- that is, how consistent test scores or other evaluation results are from one measurement to another. In other words, a test cannot measure anything well unless it measures consistently. In order to have confidence in a measuring instrument, we need to be assured that approximately the same results are obtained on different occasions.

The concept of reliability can be appreciated better after having a look at its features expressed by Gronlund and Linn (1990:78- 79) as follows:

- a) *Reliability refers to the results obtained with an evaluation instrument and not to the instrument itself.* In other words, it is more appropriate to speak the reliability of the "test scores" or of the "measurement" than of the "test" or the "instrument".
- b) *An estimate of reliability always refers to a particular type of consistency.* Test scores are not reliable in general. They are reliable over different raters, and so on. It is possible for the test scores to be consistent in one of these respects and not in another.

- c) *Reliability is a necessary but not a sufficient condition for validity.* A test that produces totally inconsistent results cannot possibly provide valid information about the performance being measured.
- d) *Reliability is primarily statistical.* The logical analysis of a test will provide little evidence concerning the reliability of the scores. One cannot find out whether test scores are consistent or not without statistical measures.

In objective vocabulary tests, such as the tests consisting of multiple choice, matching and true false questions, one can expect high scorer reliability. On the contrary, subjective tests tend to have less scorer reliability; however, if scoring directions are clear and specific as to the exact details the rater should attend to, then such scoring can become reasonably consistent and dependable of course as long as the rater is adherent to these clear and specific instructions. Brown, (1990:254)

Weir (1990:32) takes into account three aspects of reliability. The first concerns the consistency of scoring among different markers, e.g., when marking a test of written expression. The second concern is how to enhance the agreement between markers by establishing, and maintaining loyalty to, explicit guidelines for the conduct of this marking.

The third aspect is that of parallel- forms reliability, the requirements of which have to be borne in mind when future alternative forms of a test have to be planned.

2.9.1.1 Factors Affecting Reliability Measures

- a) **Test items which are not related to determined outcomes lead to low reliability.** The questions to be asked in the test should match what the teachers have done in the class. A table of specifications can help to do so.

- b) **Poorly- constructed test questions have an unfavorable effect on reliability.** Each question format has its own item rules, so the testers have to follow these rules to write well- constructed defensible test items.
- c) **Ambiguity in the test items makes the test less reliable.** Everybody should understand the same thing in the same question. Ambiguity as a result, lowers reliability of the test. Marshall et al., 1972:5
- d) **Complex and unclear sentence structure and vocabulary used in the instructions and questions influence reliability adversely.**
- e) **Poor instructions lower the reliability of the test.** The first thing the testers should do is to give a good instruction.
- f) **The test should start with the easiest items and go on with the difficult ones.** Placing difficult items early in the test causes the students to spend too much on these and prevent them from reaching items that they can easily answer and affect their motivation.
- g) **The short tests affect reliability in a negative way.** The longer the test is, the higher its reliability will be.
- h) **Inappropriate level of difficulty has a negative influence on reliability.** If a test is too difficult or too easy, we suffer from low reliability.
- i) **Overtesting the certain areas or objectives at the expense of other areas or objectives has a negative effect on reliability.**

2.9.1.2 Methods of Estimating Reliability

a) Test-retest Method

In this method, the same test is administered twice in the same group, with a time interval. Through over-repeated testing, the scores should be more or less the same. If the time interval is short the student will remember the questions and answers of the previous test and there will be high correlation and the same score tendency. If time interval is long, students will change in time.

b) **Equivalent-forms Method**

In order to eliminate remembering effect, parallel forms of the test is given to the same group with or without time interval and the results are correlated. If the correlation between the two tests is high, then the tests will have high reliability. However, there is a problem of test equivalence. Such tests should be incidental in the nature of their sampling, difficulty, length, instructions and so on.

c) **Split-half Method**

This method is based on the principle that an accurate measuring instrument is divided into two equal parts and the scores belonging to the first part should correspond exactly to the scores of the second part. Generally, odd numbered items are correlated with the even numbered items.

2.9.2. **Validity**

Validity refers to whether a test serves its purpose or not. In other words, a test is said to be valid to the extent that it measures what it is supposed to measure. (Heaton, 1975:159). For instance the following test item is invalid if we want to measure writing ability: "Is photography an art or science? Discuss it". This item is likely to be invalid because it demands some knowledge on photography and will consequently favor certain students.

Harrison (1983:11) also defines validity of a test the extent to which the test measures what it is intended to measure.

Validity can be clarified better by pointing out its nature as stated by Gronlund and Linn (1990: 45-50)

- Validity refers to the appropriateness of the interpretations of the results of or evaluation instrument for a given group of individuals, and not the instrument itself.

- Validity is a matter of degree; it does not exist on all-or-none basis. Validity is best considered in terms of categories that specify degree, such as high validity, reasonable validity, and low validity.
- Validity is always specific to some particular use or interpretation. No test is valid for all purposes. For instance, an arithmetic test may have a high degree of validity for indicating computational skill, and essentially no validity for predicting success in art or music.

Weir (1993:19) states that test validity presupposes that the writer can be explicit about what is to be tested and takes steps to make sure the test reflect realistic use of the particular ability to be measured. To the degree that is possible a test should limit itself to measuring only what it is intended to test and not unrelated or unintended abilities.

The primary concern in test development and use is demonstrating not only that test scores are reliable, but that interpretations and uses make the test scores are valid. (Bachman. 1990:236)

The interpretations of test scores are likely to have greater validity when we have a fuller understanding of:

- The appearance of the test which refers to **face validity**.
- The test content and the specifications it was derived from which refers to **content validity**.
- The relation of the test scores to other significant measure which refers to **criterion validity**.
- The nature of the characteristics being measured which refers to **construct validity**.

Henning (1987:91-93) lists the considerations that may bring to reduction of the validity as follow:

- a) *Invalid Application of Tests.* The most apparent threat to validity arises from misapplication of test. A test that is highly valid for measuring general language proficiency in a foreign language may not be valid as a measure of a specific classroom achievement objective.
- b) *Inappropriate Selection of Content.* Frequently tests are constructed with a kind of built-in invalidity. This commonly occurs when items do not match the objectives or the content of the instruction. Also this may happen when the test items are not comprehensive in the sense of reflecting all the major points of the instructional program.
- c) *Imperfect Cooperation of the Examinee.* If the examinee does not approach the testing situation in the expected manner the results may prove to be invalid.
- d) *Inappropriate Referent or Norming Population.* The items that are found to be suitable, with high discriminability, for one group of examinees may function quite differently from a qualitatively different underlying population.
- e) *Poor Criterion Selection.* If the criterion measure itself has low reliability and validity as a measure of the target competence, then validity coefficients obtained by this procedure will tend to miscalculate true validity.
- f) *Sample Truncation.* Any irregularity in the testing situation that reduces test reliability will be found to produce a corresponding reduction in test validity.
- g) *Use of Invalid Constructs.* In the situation where the constructors are not valid to begin with, it is not possible that the tests will be valid

2.9.3. Practicality

Practical matters also require consideration. How much data does the test provide on the candidates for the amount of time spent on designing it, administering it and marking it and processing the results? In other words how cost effective is the test? (Weir, 1993:19)

Harrison (1983:12) defines administrative as the main question of practicality. A test must be well organized in advance. How long will the test take? What special arrangements have to be made? Is any equipment needed? How long will it take to get the marking done, and how many people will be involved? How will test materials be reproduced in quantity, and at what cost? Shortly, the test should be as economical as possible in time and in cost.

Gronlund and Linn (1990:102) simplify ease of administration into:

- a) The directions should be simple and clear
- b) The subtests should be relatively few.
- c) The time allowed for the test should be appropriate for the given tasks in the test.
- d) The tester should make certain that the necessary equipment will be available for the test.
- e) There must be high standardization in the test administration.
- f) The test sheets should be printed or typewritten and appear neat, tidy esthetically pleasing without full of misspelling, omissions and corrections

The examinees may be unfamiliar with the test format or the way the tester wants to measure the target ability. Take the example of giving cloze tests to the students who have never encountered test of this type and format before. Their performance will certainly suffer due to the lack of familiarity with the test, so their results will not be a reflection of their actual ability or competence. Consequently,

the testers have to take this factor into account while choosing an appropriate test. (Henning, 1987:11)

From the very beginning, any test advocated should satisfy societal and institutional demands. Many excellent tests have been abandoned simply because they were not found acceptable in the eyes of teachers, parents or administration. Thus one should involve the general trends of the community in the decision-making process at this point.

2.10. Things to be done after Test Application: Item Analysis

When large numbers of students did poorly in a test, we must ask ourselves questions like:

1. Where the objectives above the level of the children's ability to understand them? For example, you may have assumed that your students knew basic information which was essential to understanding the material in the unit which was essential to understanding the material in the unit in which, actually they were unfamiliar.
2. Did something disrupt the process of instruction? Events outside of a teacher's control can sometimes disrupt the most carefully laid plans.

If the answer to these two questions are no, then the teacher may want to examine the test itself. Maybe there is something wrong with the items on the test.

(Trice, 2000:112)

In addition, Chase (1995:156) states that item analysis can be done if you first collect the highest scorers on the total test, then collect the lowest scorers and follow this procedure:

1. Find the item difficulty for item 1. If the number of students who pass item 1 in the high group added to the number who pass it in the low group turns out to be nearly all of the students, this means the item is easy. If the number in the high group added to the number in the low group who pass the item is relatively small, the item is difficult.
2. Go to item 1 and see how well it sorts out the students in the highest scoring group from the lowest scoring group. This will indicate the item discrimination.

If an item can be answered correctly by all the students both before and after instruction, the item obviously is not measuring instructional effects. In the same way, if an item is answered incorrectly by all students both before and after instruction, the item is not serving its intended function. (Gronlund, 1998:127)

Weak items should be identified and removed from the test. However, weak items cannot be identified without administration of the test.

Chase (1999:155) states that item analysis of a test can help teachers find answers when your test results are surprising, when the items you thought will be answered by your good students and all of them miss these items, something is wrong with this item.

2.10.1 Difficulty Level

Weir (1993:22, 23) states that if there is a variety of items testing a particular ability, try and put easier items first, as this will encourage the students to try their hardest and show their best. But, if you begin your test with the most difficult items the weakest will soon give up.

Calculating item difficulty will give us the percentage of the high and low scoring students who got the item correct. Chase (1999:157) suggests a procedure which will work for a class of 20 to 40 students. The suggested figure looks like:

Item	Right hi	Right lo	R + R hi lo	(R + R)/total hi lo	R - R hi lo	(R - R)/.5 total hi lo
1.						
2.						
3.						
4.						
5.						
6.						

Figure 1 A worksheet for Recording Item Analysis Data

1. After you have scored your test, sort out the 10 highest scorers on the total test. Then sort out ten lowest. Put aside the rest of the papers.
2. Go to item 1 on the “high” papers. Put into a table all the “lows” who got the item correct.
3. Compute the item difficulty index from item 1. Do this by adding together the number right form from the “highs” and the number right from the “lows”. Then divide this sum by the total number of pupils in the two groups. (10+10 or 20 pupils). This produces a proportion ranging, typically from 0.00 to 1.00. The formula looks like this:

Rights hi + Rights lo

Difficulty= -----

Total number of pupils in high plus low groups

2.10.2 Discrimination Level

One of the important features of a test is its capacity to discriminate among the testees and to reflect the differences in their performance. The testers should select or develop the tests after they take the abilities and other characteristics of the

examinees into account. Inspection of the difficulty level of reading passages, the length and complexity of problem statements, the familiarity of vocabulary, and so on can reveal whether a test is appropriate for the testees. (Heaton, 1975:165)

Even in achievement test one might wish to include some of the more difficult elements from the syllabus which will be achievable perhaps by only the best students in the class. The number of these should be limited so as not to demotivate the others, possibly include them at the end of a test.

Henning (1987:51) states that another important characteristic of a test item is how well it discriminates between weak and strong students.

Some items on a test are going to be harder than others. There is no theory on criterion referenced classroom test about what percentage should pass an item. Maybe giving students an occasional low grade will motivate them to study harder.

The next step is to see whether the students who scored in the top of the test did better than those who scored in the bottom of the test. If the better students did worse on an item than those on the bottom, we should be very concerned. (Trice, 2000:113)

Usually, we compare the percentage of students passing the item in the top third of the class, with those passing the item in the bottom third of the class.

Henning (1987:52) defines as more common technique for the calculation of item discriminability is to calculate the correlation between item responses and total scores for any given test. Since item scores in this instance are binary, zero or one, and total scores are continuous, the resultant correlation coefficient is called a point biserial correlation. The formula for biserial correlation is like this:

$$r_{pbi} = \frac{X_p - X_q}{S_x} \sqrt{pq}$$

r_{pbi} = the point biserial correlation

X_p = the mean total score for examinees who pass the item

X_q = the mean total score for examinees who fail the item

S_x = the standard deviation of test scores



CHAPTER III

DATA ANALYSIS AND INTERPRETATION OF THE RESULTS

3.0. Introduction

A questionnaire (See Appendix) was administered to teachers teaching English as a Foreign Language at Fourth year high school in different parts of Macedonia such as: Pance Poposki High School, Kiril Pejcinovic High School, Zlate Malakoski High School, Nikola Shtejn High School, Kliment Ohridski High School, Gjoce Stojceski High School, Koco Racin High School and Zef Lush Marku High School.

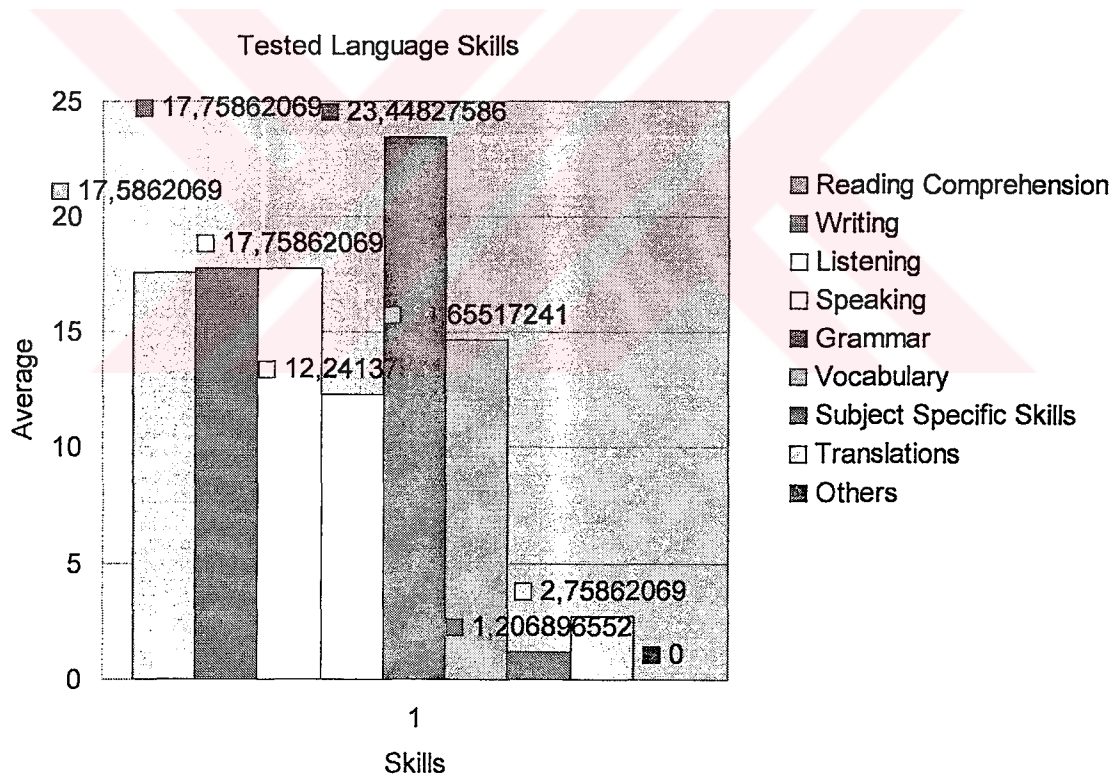
Tested Language Skills

Table 2.

	Number	Out of	%
Reading Comprehension	28	30	
Writing	30	30	100 %
Listening	27	30	
Speaking	27	30	
Grammar	30	30	100 %
Vocabulary	29	30	
Subject Specific Skills	4	30	
Translations	11	30	
Others	0	30	0 %

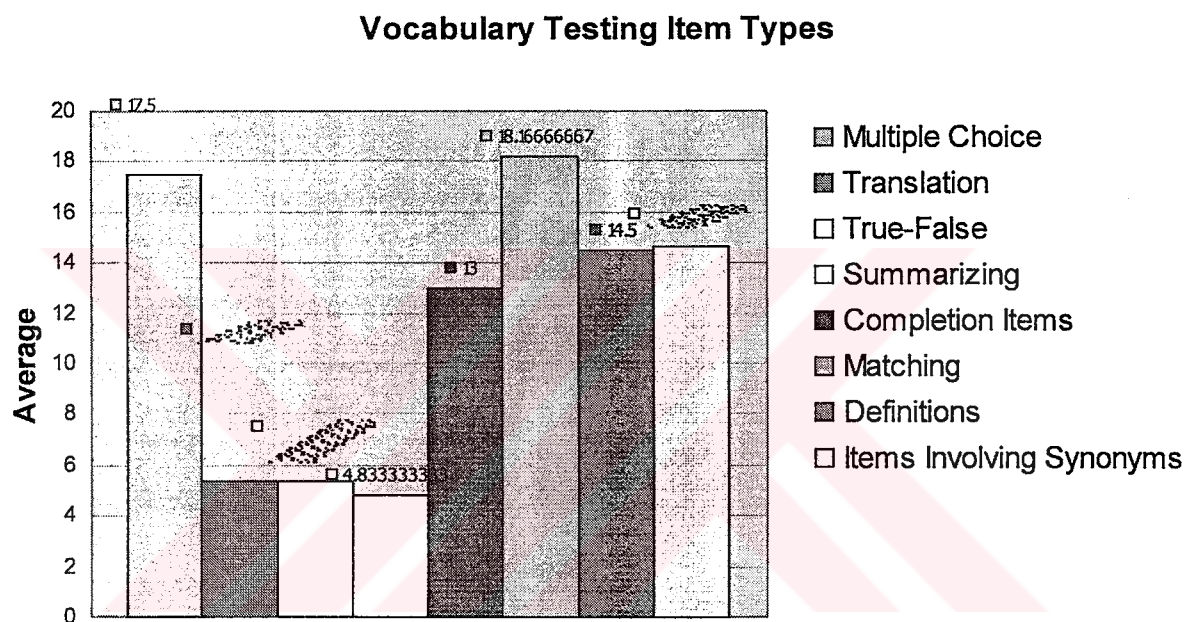
According to the responses given to item one 28 out of 30 teachers test reading comprehension. Hundred percent of the teachers test writing. 27 of the teachers test listening and only 3 of them don't test listening. Speaking is also tested by 27 of participants in this questionnaire. Grammar is tested hundred percent. Vocabulary is tested almost by all of the teachers, that is to say only one of the participants in this questionnaire don not think testing vocabulary is necessary. Subject specific skills are tested only by four teachers out of 30. Eleven teachers use translations in foreign language testing.

Graph 1.



The Item Types Used for Vocabulary Testing

Graph 2.

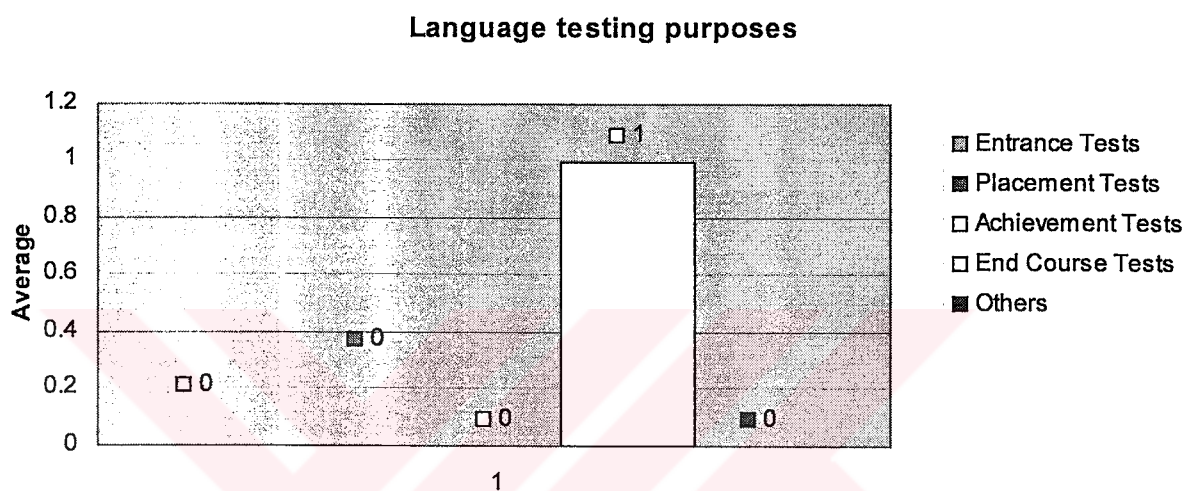


In this item teachers were asked which item types they use for vocabulary testing. According to the results 18.6 % of the teachers use matching items to test vocabulary. 17.5 % of the participants use multiple choice type tests. 14.6 % of the teachers use items involving synonyms. 14.5 % of the participants use definitions in testing vocabulary. 13 Percent of them use completion items. 5.3 Percent of them use translations and 4.8 of the teachers use summarizing.

From the obtained results we can say that teachers who participated mostly use matching items to test the vocabulary knowledge of their students. Only 4.8 % of them use summarizing as a testing technique.

The purpose of using language test

Graph 3.

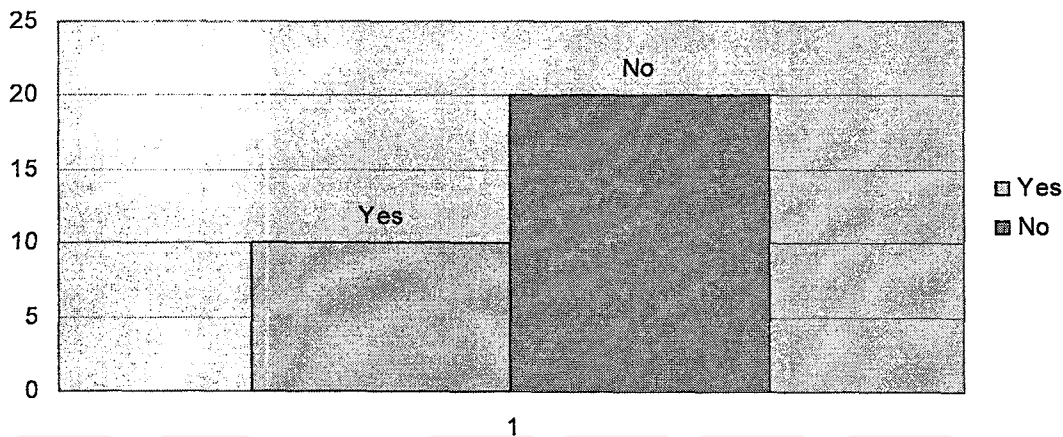


In this item as we can see from the above graph most of the participants use language tests as end course tests.

Teachers that took Language Testing Subject

It is an interesting fact that only nine of the participants took a subject on language testing and assessment during their college studies. Teachers that took such subject were graduated from universities out of Macedonia, whereas, teachers graduated from universities in Macedonia did not take such a subject.

Percentage of Teachers that took subject on Language Testing and Assessment



Attending Seminars or Conferences on Language Testing

This item was answered like this. The number of teachers that have attended seminars and conferences on language testing was 21 and the teacher who did not attend at all any kind of seminar or conference was 9.

Sharing ideas with colleagues and co-operating about preparing tests

Only two teachers of the participants do not co-operate and share ideas with their colleagues in preparing language tests, whereas, 28 of them mostly consult and share experiences with their colleagues.

Writing tricky items in multiple choice items

The number of teachers that use tricky items in multiple choice items is 20, and ten of the participants do not use them

Pairing words of opposite meaning with the right answer.

Eight teachers do not prefer to pair words of opposite meaning with the right answer, and 22 of them, pair words with the right answer.

Does the stem give the students a problem before considering answers

Fifteen of the teachers consider that stem should give the students a problem, and fifteen of them do not think so.

Should all distractors be equally plausible?

Once more, half of the participants or 15 teachers think that distractors should be equally plausible and 15 of them think the opposite.

Using specific determiners in True- False Items

According to the responses 15 of the teachers use specific determiners in their True- False items and the other 15 do not use.

Using “large”, “a long time ago” and “often” in True- False Items

Most of the teachers do not use such words in true false items, only nine of participants responded positively to this question.

Using homogenous material in Matching Items

The given answers show that only 8 teachers do not use homogenous material in their matching items the other 22 answered that they use such material.

Taking statements from the textbooks for Completion Items.

Concerning this question only one teacher does not take statements from the textbooks, the other 29 prefer taking statements from the textbooks.

Grammatical consistency of the alternatives with the stem of the item.

According to the responses, 19 of the teachers think that alternatives should have a grammatical consistency with the stem, and 11 of the participants do not think that this issue may be important for a multiple choice test.

Using “all of the above” and “none of the above” in multiple choice

The number of teachers that use the alternative “all of the above” is 19, the others use the alternative “none of the above”

Some characteristics to be followed when evaluating items.

	Number of Teachers	Out of
Does each item measure an important learning outcome included in the tests specifications?	21	30
Is each item type appropriate for the particular learning outcome to be measured?	19	30
Does each item present a clearly formulated task?	28	30
Is the item stated in simple, clear language?	28	30
Is the item free from extraneous clues?	12	30
Is the difficulty of the item appropriate?	25	30
Is each test item independent and are all the items as a group free from overlapping?	8	30
Do the items to be included in the test provide adequate coverage test specification?	18	30
All above	9	30
None	2	30

According to the given responses 21 of the teachers participating when evaluating items think that each items measures an important learning outcome.19 of the teachers think that each item type used for learning outcomes are appropriate.28 of the teachers consider each item present clearly formulated task. 28 of the teachers state the item clear and in simple language.12 of the teachers state that the items are free from extraneous clues.25 of the teachers think that the item difficulty is appropriate. 8 of the teachers point out that each item is independent and free from partly covering. 18 of teachers state that the items included in the test provide enough coverage. 9 of the teachers follow all these mentioned rules when evaluating items. 2 Teachers do not consider any of these rules.

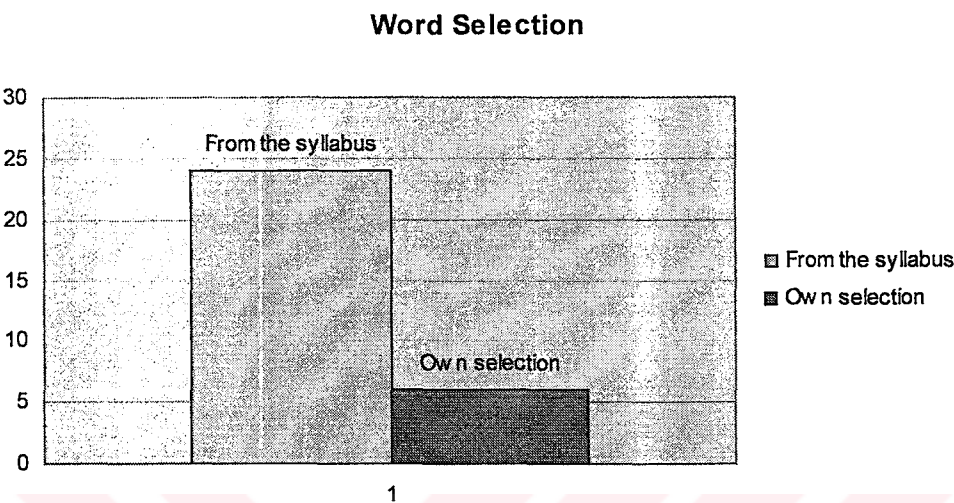
The considerations teachers are concerned with when bearing in mind the validity of the assessment results.

Table 4

	Number of Teachers	Out of
Content representativenes	11	30
Criterion relationships	11	30
Construct evidence	7	30
Consequence of using	0	30
None	1	30

11 teachers are concerned with content representativeness when considering the validity of the assessment results.11 teachers also are concerned with criterion relationships. 7 teachers are concerned with construct evidence. No teacher is concerned with the consequence of using and 2 teachers are not concerned with any of the given considerations.

The way teachers select words to test



24 of the participating teachers select the words to teach and test from the syllabus 6 of the teachers make their own selection when teaching/testing vocabulary.

How often teachers calculate the arithmetic mean of test scores.

Table 4

	Number of teachers	Out of
Never	2	30
Seldom	3	30
Sometimes	7	30
Usually	10	30
Always	8	30

The results show that 2 teachers never calculate the mathematics mean of the test scores. Three teachers seldom calculate the mathematics mean. Seven teachers sometimes calculate the mathematics mean of the test scores. Ten teachers usually

calculate the mathematics mean, and eight teachers always keep calculating the mathematics mean of the test scores.

Can mathematics mean of the scores help teachers.

Only three teachers answered that calculating the mathematics mean of scores can not help the teachers. 27 teachers agree on that this calculating can help teachers.

Do you compare the test results with colleagues?

Table 5

	Number of teachers	Out of
Never	5	30
Seldom	7	30
Sometimes	10	30
Usually	3	30
Always	5	30

5 Teachers never compare the test results with their colleagues. 7 Teachers seldom compare the test results. 10 teachers sometimes compare the test results with the colleagues. 3 Teachers usually compare the test results with colleagues. 5 Teachers always compare the test results with their colleagues.

Do you consider test results for constructing further tests?

24 Teachers participating in this questionnaire consider test results for further tests. Six teachers do not consider test results for constructing further tests.

Objective tests are easier and less time consuming to answer than subjective tests.

Only three teachers do not agree with the definition “objective tests are easier and less time consuming”. Twenty seven teachers support the definition.

The degree of objective tests to measure success.

Table 5

	Number of teachers	Out of
A great deal	3	30
Quite a lot	9	30
A lot	10	30
A little	5	30
Very little	2	30

Three teachers state that objective tests measure the success a great deal. Nine teachers state that objective tests measure the success quite a lot. Ten teachers state that objective tests measure a lot of the success. Five teachers state that such type tests measure a little of the success. Two teachers state that objective tests measure very little success.

The tests are in accordance with the course objectives.

Only two teachers think tests are not in accordance with the course objectives. Teachers state that tests are in accordance with course objectives.

Giving clear and right instructions.

Every teacher that conducted this questionnaire state that they give clear and right instructions.

Given sufficient number of tests.

20 Teachers give sufficient number of tests to identify students’ success. 10 Teachers state that the given number of tests is not enough sufficient to identify the success of the students.

Other purposes of using tests.

Table 6

	Number of teachers	Out of
To identify students’ learning difficulties	26	30
To motivate learners for better learning	0	30
To look up for a new teaching approach	4	30
Others		30

26 Teachers apart from testing the students’ success use tests to identify learning difficulties. Four teachers except testing the success, use tests to look up for a new teaching approach.

Satisfaction with the tests.

Table 6

	Number of teachers	Out of
Absolutely	2	30
Most likely	15	30
Possibly	11	30
Not likely	1	30
Not at all	1	30

Only two teachers in most cases are absolutely satisfied with tests.15 Teachers are most likely satisfied with tests.11 Teachers are possibly satisfied with tests. One teacher is not likely satisfied with tests. One teacher is not at all satisfied with tests.

3.1. Interpretation of the Questionnaire Administered to the Teachers

According to the data gathered from teachers who conducted the questionnaire the interpretation can be stated as follows:

Tests are said to be appropriate for students in terms of their knowledge, expectations and skills. For teachers tests are fitting with the aims and practices of the language teaching. Language is tested in the way it is taught. The majority of the teachers (29 out of 30) give an important place to vocabulary testing seeing it as a crucial element in language learning and testing.

Clearly knowledge of vocabulary is essential to the development and demonstration of linguistic skills. But that does not mean that it should be tested separately.

The majority of the teachers see matching items as the main vocabulary testing item type. Chase (1999:143, 144) lists the advantages of matching items as follows:

- *Matching exercises offer a short procedure for assessing learning of a set of related ideas, facts and similar content.* This type of test is made up of single words or short phrases in one column which are to be matched with single words or short phrases in a second column.
- *Matching calls for a minimum of reading.* This test item can provide you with a set of related facts quickly and requires little reading on the student's part.
- *Matching exercises are relatively easy to construct.* Once the list of related concepts with which the exercise is to be made is ready, the test writing time is short.
- *Scoring is easy and reliable.* Scoring matching items is quick and simple, especially if students are asked to write the letter of the response on a line before the appropriate stimulus.

On the other hand, the multiple choice type of item has achieved its success in vocabulary tests. Chase (1999:114) classifies multiple choice items as the most popular selection type item, suitable for use in assessing outcomes in almost all subject matters. It has two parts, the stem- defining the problem and the options- the possible responses to the problem.

Lado (1961:188) also defines multiple choice items success as the most impressive in vocabulary test. A better type of multiple choice item to test vocabulary is a multiple choice item that gives an appropriate context for the lexical problem in the lead.

The third most widely used item type of the teachers conducting this questionnaire is items involving synonyms. Heaton (1988:61) suggests involving

synonyms in testing words. Writing items that involve synonyms is one of the ways to involve a number of different operations.

Definitions according to the results from the questionnaire are ranked as the fourth mostly used item types. Harris (1969:51, 52) defines this type of tests as "classic" type of vocabulary item, consisting of a test word followed by several possible definitions.

In contrast, Heaton (1988:62) states that these item types are of very little use. It is too hard even for native speakers to produce sentences showing the meaning of words.

Completion items consist of a sentence from which the target word has been deleted and replaced by a blank, the function of the sentence is to provide a context for the word and perhaps to cue a particular use of it. This type of item is ranked as the fifth most used item type by the teachers.

The least used item types are summarizing and translations. Using of these types of items is not recommended since they tend to reduce the validity of a vocabulary test.

In order to teach effectively teachers must identify problems and specific weaknesses of their students. To do so, teachers should know what they are testing. End courses measure what students have achieved during the course. The most important reason of end course tests is to look back what students have achieved or how well students have mastered the language areas and skills that have been taught. (Heaton, 1989:9)

End course test are used to see how well students have learnt what the course set out to teach them. They are formal examinations given at the end of the school year or at the end of the course. They are like progress tests but are usually designed to cover a longer period of learning. In fact, they should attempt to cover as much of the syllabus as possible.

The second purpose why teachers conducting the questionnaire use tests is to sort students into groups according to their language ability at the beginning of a course. The placement tests should be as general as possible and should concentrate on testing a wide and representative range of ability.

The third ranked purposes of using tests are achievement tests but, unfortunately they are not used sufficiently. An achievement test is also like a progress test but is usually designed to cover a longer period of learning than a progress test. The purpose of achievement tests is to find an answer to the question 'Has the student learnt what he/she is supposed to have learnt according to the teacher's teaching?'.

Language testing and assessment has an important role in education, but most of the teachers did not take a subject regarding language testing. This is not their false but it isn't included in educational programs in Macedonian Universities.

Through regular, informal tests, teachers obtain feedback for themselves and the students on the efficiency of the teaching and learning process. Teachers should try to evaluate all the skills, in isolation and by means of mixed skill tests.

Considering the fact that most of the teachers did not took a subject on language testing and assessment I decided to ask them whether they are attending seminars or conferences on this subject. Most of them did this as a reasonable answer and need of each teacher.

Most of he teachers marked the "yes" option in co-operating with colleagues for bringing out their comments on tests. Lowe (1975:323) supports this idea stating the fact that the reliability and fairness of tests are directly related to the total number of different pieces of information from different teachers.

Half of the teachers think that the stem of the item should present a problem and half of them were against this. Linn& Gronlund (1995:185) state that the stem should present a problem. Often the stems of test items placed in multiple choice form are incomplete statements that make little sense until all of the alternatives have

been read. These are not multiple choice items but rather a collection of true false statements placed in multiple choice form. Formulating a problem in the stem not only improves the stem of the item, but also has a desirable effect on the alternatives.

The distracters should be plausible since the purpose of a distracter is to distract the uninformed from the correct answer. 15 of the teachers agree with this statement and 15 of them don't think that distractors should be plausible.

The main purpose of making plausible distractors is to eliminate the implausible answers and guessing among the remainder. If all of the alternatives are homogenous with regard to the knowledge being measured, the distractors are more likely to function as intended.

Half of the teachers answered the question of using specific determiners such as "all", "none", or "no" positively which means that they use such words in true false items.

Chase (1999:132) states that when constructing true false items the first guideline is to avoid such words. Specific determiners are words that indicate situations that are without exceptions. They provide the clue that the item is false.

Most of the teachers (21) do not use terms as "large", "a long time ago", and "often" in true false items which is considered as wrong because such terms are slippery and people do not agree on what they mean. They also lower the validity of the tests.

The question on using homogeneous material in matching items was answered positively by 22 teachers which show that they have better knowledge on matching items or maybe they use the most this item type. Linn & Gronlund (1995:168) define homogeneity as the most important rule of constructing matching exercises.

In order to construct completion items almost every teacher (29 out of 30) prefers taking statements from the books. This encourages rote memory, an activity that educators try to avoid. Memorization does not imply understanding.

The alternatives should be grammatically consistent with the stem of the item. While, 19 teachers agree with this statement, 11 of them don't.

The main purpose of this rule is not to present only complete grammar usage, however, its main purpose is to prevent irrelevant clues from creeping in.

When evaluating items 21 teachers stated that each item measure an important learning outcome included in the tests specifications. The problem situation around which an item is to be built should be important and should be related to the intended learning outcome to be measured.

"The stem of the item should be stated in simple clear language" statement was answered positively by almost all the teachers which indicate that the teachers are aware of the fact that anyone who possesses the knowledge to be measured by a test item should be able to select the correct answers.

Most of the teachers choose words to test from the syllabus. This way is supported by (Madsen, 1983:12, 13) also, stating that the need to find what words to test is solved by the course syllabuses that tell us what words students need to know. Another way is to record the words that students misuse.

Although the teachers believe in the usefulness of the tests scores only 8 teachers "always" calculate the mean of the test scores regularly. Those teachers sometimes compare the test results of their classes with the results of other classes. As a result the teachers cannot evaluate their students' performance in relation to other classes.

Test results in a way are a kind of guidelines to constructing forthcoming tests. They show the needs of the tests, and points that need better revision.

Chase (1999:33) states that objective tests are objective because they are scored objectively, including multiple choice, true-false, and matching items which means that all items have one right answer and no judgment-calls are involved in scoring these tests.

Objective item tests can be marked more easily and more accurately, and they may create an encouragement for students to build up a wide base of knowledge, skills and abilities. The teachers are aware of the above facts on objective tests, since 27 seven of them agreed with the statement that objective are easier and less time consuming.

According to the majority of the teachers the results of the 29th question the test results are mostly used to find out students' learning difficulties. This means that they are going to reinforce the parts that students did not do well in the test.

In spite of certain weaknesses in test constructing, half of the teachers view present tests as most likely satisfactory.



CHAPTER IV

SUGGESTED VOCABULARY TESTS

4.0 Samples of Vocabulary Tests for the Fourth Year Secondary School Students

Teachers that were included in this study use the New Headway Upper-Intermediate coursebook from Liz and John Soars published on 1998 by Cambridge University Press. This book provides around 120 hours of language learning and language skills are integrated through tasks.

This coursebook is used for six years by the teachers and none of them has a test bank which can be used in their achievement tests and they always prepare new tests.

Most of the teachers do achievement tests after covering three units. Considering this fact, the suggested vocabulary test words will provide the basis for the balance of vocabulary contained in the syllabus, and the topic areas will point to the main fields to be covered.

The problem of choosing which word to test is solved by the course syllabus since it tells us what words must be learned.

The item type 1, tests the words from unit 1 to unit 3. The item type 2, tests the words from unit 4 to unit 6. The item type 3, tests words from a text included in unit 9. The text is given as a reading passage where new words are to be tested.

Type 1. Fill in the blanks with appropriate words from the box

sensible	proud	illegal	drugs	lifestyle	pitiable	immature	risk
mood	consume	sober	disloyal	pressure	advice	remote	

1. A cup of black coffee might _____ you up.
2. I think your suggestion is very _____, you judge things better than I do.
3. I'd been stuck in the traffic for three hours, which put me in a really bad _____.
4. The last government taxes will affect the _____ of many people.
5. There was some _____ that the fire would break out again.
6. Mehmet felt he had been _____ to his friends.
7. The refugees at the Stankovec Refugee Camp were living in _____ conditions.
8. They want to ask your _____ on how to take care of the baby.
9. She married much too young- she was too _____.
10. My friend Jane is penniless, but she is too _____ to ask for help.
11. For a long time he has been trying to put _____ on us to go to Tetova.
12. The union leaders have to beware of getting too _____ from their members.
13. Marijuana is _____ in many countries.
14. He was arrested for selling _____.
15. There are many countries that _____ more than they produce.

Type 2. Match the words on the left with the definitions on the right.

__compromise	1. a body organ
__armour	2. ashamed or uncomfortable in a situation
__depression	3. to gain money from something
__pottery	4. to work for someone for no money
__dull	5. a strong layer for protection
__profit	6. amount of money paid to the government
__slave	7. getting better
__fodder	8. metal container for food
__wealth	9. agreement between two people
__decrease	10. reducing the amount of something
__tin	11. food for farm animals
__embarrassment	12. amount of money you possess
__mend	13. a factory where baked clay objects are made
__tax	14. not interesting or exciting
__gland	15. feeling of sadness
	16. the amount of money you spend
	17. arresting someone for something

Type 3. Reading Comprehension

My mate Stuart is a funny bloke. He is an insomniac, he is dyslexic, and he is an atheist. He is single, unemployed, and lives all alone in a small basement flat without even a pet for company. Also he is vegetarian and teetotal, and he does not touch caffeine. He is pretty anti-social, actually.

I went round to see him last Sunday. As I walked up the drive, his dog started barking. His wife answered the door, and we went upstairs to their living room. He was in a bad mood because he had overslept that morning and had been late for church. He had a bit of a hangover, as well. Over a cup of coffee, he told me about the wild party that they had had at his house the night before. They had had a barbecue

in the garden with hamburgers and hot dogs. One of his favourite pastimes is crosswords, and he spends all his lunch break at work doing them.

"So how are you Stuart?" i asked him the other day.

"OK, mate, OK. How about you?"

Anyway, as I said, Stuart is an insomniac, dyslexic, atheist. So the joke is that he lies awake all night wondering about the existence of dog. Get it?

1. In its context, the word "bloke" means most nearly

- A) friend
- B) guy
- C) man
- D) person

2. The concept of "insomniac" refers to

- A) someone who is sick
- B) someone who can not sleep
- C) someone who can not eat
- D) someone who is hungry

3. The word "dyslexic" refers to

- A) a person who is not able to read
- B) a person who can read
- C) a person that reads a lot
- D) a person that reads books

4. Which of the following comes closest to the meaning of the word "atheist"

- A) a person with certain duties
- B) the belief that God does not exist
- C) someone who often does sports
- D) a man looking for a job

5. The word “teetotal” means

- A) a person in diet
- B) a person that does not smoke
- C) a person that never drinks alcohol
- D) a person that does not drink coffee

6. In this text the word “barking” refers to

- A) to run
- B) to attack
- C) to pursue
- D) to make sound

7. The explanation of the word “mood” would be

- A) the way you feel
- B) the way you behave
- C) the way you speak
- D) the way you do something

8. The expression “overslept” means

- A) to wake up early in the morning
- B) to sleep longer than you intended
- C) to have difficulties while sleeping
- D) to stay in bed after you wake up

9. The word “hangover” refers to

- A) a suffering from a disease
- B) a serious sickness
- C) a physical or mental pain you feel
- D) a headache and feeling of sickness

10. The word “barbecue” means

- A) an ordinary meal
- B) an outdoor party
- C) a celebration of something
- D) a plain dinner

Type 4. Write a sentence using the words given below.

- 1.....
.....(generous)
- 2.....
.....(deposit)
- 3.....
..... (annum)
- 4.....
..... (match)
- 5.....
..... (income)
- 6.....
.....(permanent)
- 7.....
..... (fascinating)
- 8.....
..... (dishwasher)
- 9.....
..... (unpleasant)
- 10.....
..... (compromise)

Type 5. Multiple Choice Items

Choose the correct response to fill the gap

1. The _____ is upstairs in the bathroom.
a) sink b) washbasin c) landing d) hall
2. "I tried to phone you, but your line was _____."
a) occupied b) taken c) engaged d) connected
3. My uncle's farm has a lot of _____.
a) gardens b) grasses c) fields d) lands
4. My parents often _____ money from the bank.
a) save b) lend c) cash d) borrow
5. Members of Parliament in Macedonia are _____ by the people.
a) elected b) represented c) voted d) nominated
6. Sorry, I can't repay you this week. I'm completely _____.
a) dead b) broken c) broke d) destroyed
7. This test is so difficult I feel absolutely _____.
a) sharpened b) shattered c) scattered d) scrambled
8. We'll be glad to put you _____ for the weekend.
a) down b) up c) in d) away
9. The teacher told us the very _____ tale of how he got stuck in a lift.
a) hilarious b) gorgeous c) joyous d) ridiculous
10. It seemed _____ to move to a bigger house since we became a crowded family.
a) conventional b) easily c) sensible d) complete

Type 6. Definitions

Chose the correct definition for the given words

1. extravagant

- a) to spend more money than you can afford
- b) costs more money than is reasonable
- c) something of no real value or use
- d) to spend more money than you have in your account

2. loan

- a) a sum of money you give to hire something
- b) a sum of money you get
- c) a sum of money you borrow
- d) a sum of money that is spent on something

3. junk

- a) waste from a kitchen
- b) an amount of old and useless things
- c) unwanted things or materials
- d) dirty or untidy things

4. tramp

- a) a person asking for food or money from people
- b) a person who can not be trusted
- c) a person who deceives you to get your money
- d) a person with no home, job and no money

5. conscious

- a) awake or you realize what is happening
- b) worry about what is happening
- c) know what is happening
- d) not concerned about what is happening

6. relevant

- a) not interested about the topic
- b) give attention to the topic
- c) connected with what is talked
- d) similarity between two things

7. mature

- a) to reach the complete development
- b) to be experienced at something
- c) to get the necessary knowledge
- d) to know something very good

8. permit

- a) to do something no longer
- b) to allow something
- c) to reduce the amount of something
- d) to stop something going further

9. decrease

- a) to grow up
- b) to become smaller
- c) to collapse
- d) to enlarge

10. refuse

- a) agree to do something
- b) take no notice of something
- c) partly agree to do something
- d) not do something at all

Type 7. Completion Items

Read through the following passage containing a number of incomplete words. Each dash represents one letter.

One of my (1) f-v-u- -t- parts of London is Soho, which is right in the centre, and includes Piccadilly Circus, and Leicester Square. One of the main reasons I like it is that it is always (2) li-e- - and colorful, with people dashing around going about their business, which is mainly (3) h- - es- but not always. The place is a bit of a mess, and the buildings aren't the most beautiful in London, but the streets are always interesting, with surprises around every corner. The name is derived from a (4) hu- -in- call, 'so- ho', that huntsmen were heard to cry as they chased deer in royal parklands. It has been a (5) c-s-o- -l-t-- area since the first immigrants, who were French Huguenots, arrived in 1680's. More French arrived escaping the revolution during the late 18th century. Soho is packed with (6) -o- -i-en-al food shops and restaurants. Soho has a (7) r-p-t-t- -n for attracting artists, writers, poets, and people in the media. There are also street markets, (8) a- -e-t-s-n- agencies, clothes shops, music (9) p-b-i-h-rs, and recording studios. Piccadilly Circus is like a (10) m-g-e- for young people from all over the world. It is said if you wait long enough at Piccadilly Circus, you will meet everyone you have ever known!

Type 8. Synonyms

Choose the alternative (a, b, c, d) which is closest in meaning to the italic word on the left of the page.

- | | | | | |
|--------------------------|--------------|----------------|----------------|----------------|
| 1. <i>admit</i> | a) allow | b) accept | c) expel | d) oppose |
| 2. <i>survive</i> | a) exist | b) die | c) endure | d) improve |
| 3. <i>accomplishment</i> | a) failure | b) operation | c) accompany | d) achievement |
| 4. <i>consume</i> | a) to use | b) to spend | c) to diminish | d) to destroy |
| 5. <i>bet</i> | a) wager | b) prediction | c) assert | d) claim |
| 6. <i>compromise</i> | a) affection | b) rumor | c) contract | d) agreement |
| 7. <i>faith</i> | a) trust | b) belief | c) confidence | d) hope |
| 8. <i>clever</i> | a) wise | b) congenial | c) cute | d) fortunate |
| 9. <i>brave</i> | a) coward | b) courageous | c) fierce | d) lethal |
| 10. <i>polite</i> | a) courteous | b) considerate | c) respectful | d) kind |

Type 9. Associated Words

Read each of the following lists of five words. Erase the word that does not belong to the group.

- | | | | |
|----------|---------|------------|-----------|
| video | kind | great | happiness |
| computer | loyal | marvellous | hilarious |
| radio | honest | sound | cold |
| speed | polite | boring | starving |
| recorder | popular | annoying | terrified |

CHAPTER V

CONCLUSION AND SUGGESTIONS

5.1 Conclusion

Evaluation has an indispensable part in language teaching process but it is not restricted to the context of education; it is a part of our everyday lives. However, in the educational context the results of a test may determine whether a learner moves into a higher level, class or repeats the course or year.

However, language testing does not include only the testing of language skills, it requires assessing of the words the students have learnt since the vocabulary knowledge is accepted as a main component of second/foreign language proficiency. Vocabulary is central to language and of critical importance to the typical language learner.

For conclusion it can be said that if we are going to teach vocabulary to our students, at some point along the way we will want to test them to see how they are progressing. A vocabulary test aims to measure the comprehension and production of both active and passive vocabulary.

From different points of view, vocabulary can be seen as a priority in 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.

Read (2000:7) regards two contrasting perspectives on the role of vocabulary in language assessment. First stating that it is sensible to construct tests that show whether learners know the meaning and usage of sets of words and second, that vocabulary must be assessed in the context of a language-use task.

The place of vocabulary in language tests is as a primary goal of language testing, because the vocabulary knowledge is accepted as a basic component of foreign language proficiency and it is natural to expect assessing.

Since vocabulary knowledge is accepted as basic structure of foreign language proficiency, one of the primary goals of language testing will be to assess whether the learner knows the meaning of the words they need to communicate effectively in the foreign language. Madsen (1983:12) supports that “the aim of a vocabulary test is to measure the comprehension and production of words used in speaking and writing”.

Any advisability of testing grammar applies equally to the testing of vocabulary”.

A clear knowledge of vocabulary is crucial to the development and demonstration of linguistic skills, characterizing this as a main reason for testing vocabulary.

A test of vocabulary measures students’ knowledge of the meaning of certain words as well as the patterns and collocations in which they occur. A well-constructed vocabulary test will provide the students with an opportunity to show their ability to perform certain tasks in the language. Such a test may test the words they should be able to use in speaking and writing or the words they should be able to recognize and understand when they are listening to someone or when they are reading.

The misconceptions about the development and use of vocabulary tests, and unrealistic expectations about what vocabulary tests can do and what they should look alike, have caused teachers becoming not enough competent in vocabulary testing.

Testing vocabulary is similar to other testing in other areas of language testing and the same criteria of reliability and validity needs to be considered when designing vocabulary tests.

While testing vocabulary we have to consider the test item type which should depend upon the below criteria:

1. Is the knowledge required to answer the item correctly similar to the knowledge that you want to test?
2. Is it easy to make enough items to test all the vocabulary you want to test?
3. Will the items be easy to mark?
4. Will answering the item provide useful repetition of the vocabulary and perhaps extend learner's knowledge?

Read (2000:76) even classifies vocabulary becoming a component of objective tests where no subjective judgment is required to score them because:

- Words could be treated as independent linguistic units with a meaning expressed by a synonym, a short defining phrase or a translation equivalent.
- The list of the most frequent English words provided a large stock of vocabulary items that could be conveniently sampled to select the target words for a test.
- Objective vocabulary tests seemed to be valid indicators of language ability in a broad sense. Furthermore, the strong relationship between vocabulary and general intelligence is one of the strongest findings in the history of intelligence testing.

In addition, Bachman and Palmer (1996:8) state that being competent in language testing will help students to avoid some of the misconceptions and to develop a set of reasonable expectations for any given language test. Language teachers need to be able to either make informed judgments in selecting appropriate language tests or to plan, construct, and develop appropriate tests of their own.

5.2. Suggestions

Language testing and assessment has an important role in education, but most of the teachers did not take a subject regarding language testing.

Tests provide the teacher with some indication of how well certain material has been taught by the teacher and assimilated by the students. In other words, tests are measures of feedback to both the students and teachers regarding learning successes and failures.

The same feedback provides to students reinforcement of successful learning and identifies the specific learning errors that need correction.

The quality of a test depends on its degree of relevance, representativeness, and reliability. These major qualities enable teachers to determine the precise value of tests and prove some tests better than others. After deciding what he is going to test, a teacher should think of how he will assess this aspect of language. After this comes the selection of the type of test. Here the teacher must be sure that the test suits his situation and gives good discrimination among students.

Most of the teachers have so far failed to realize that a logical explanation for their students' poor test scores are essentially ascribable to wrong test items that have been administered.

Their tests provide us with a good account of the students' ability to manipulate only the grammar of the target language and some comprehension questions to answer. Although instruction has been oriented toward helping students use the language genuinely, testing has remained static; it only measures whether or not the test takers have been linguistically accurate. A language test should be dynamic, reflecting students' communicative needs rather than being a body of passive items. Thus, language tests sacrifice communicative fluency and as well the vocabulary in favor of grammatical accuracy. Such a mismatch should be eradicated to allow reliability, validity, and authenticity to take place in the process of assessment.

Teachers should decide whether to test the student's active or passive vocabulary-that is. the words they should be using in their speech and writing or those they will need only to comprehend, especially in their reading.

In order to do this the main source will be the course syllabus which tells us what words students need to know, through supplementary materials selected by the teacher, or record the words that students use wrongly and make them as test items.

Teachers should be careful when they test vocabulary by using multiple choice items. It should give an appropriate context for the lexical problem in the stem and if the problem area is included in the stem, the alternatives themselves should contain only those vocabulary items which the test takers can understand.

The distracters should be so attractive to the students who lack the knowledge called for by the item that they select one of the distracters in the preference to the correct answer.

When using completion items teachers should keep in mind that copying information directly from the textbook may cause memorizing of the same. The purpose of completion items is to make the student bring to mind the word that is deleted and encourage critical thinking.

Teachers should add one or two extra definitions when constructing matching items in order to prevent situation where the student knows five out of six target words and can then get the sixth definition correct by eliminating, without knowing what the word means. If the aim of the test is on knowledge of the target words, the definitions should be in accordance with level knowledge of the learners so it is easy for them to understand.

When preparing reading tests based on vocabulary teachers should check whether the problem lies in only one place, either in the stem or in the options. Valette (1977:132) states that "If the difficult vocabulary item is in the stem, all the students should understand the meaning of the options".

Written vocabulary test should be constructed around lexical units rather than words and they should evaluate the student's ability to recall and produce lexical units in the target language.

Teachers should always encourage the students to speak in natural utterances when they test vocabulary with speaking.

They should also choose the listening texts with the test specifications in mind. If they are interested in how candidates can cope with language intended for native speakers, then ideally they should use samples of authentic speech.

Knowing the crucial role of the assessment teachers involved in foreign language learning should attend seminars and other topic related professional development sessions to share their ideas and needs in language testing. This will help them develop the competency of constructing better language tests.

Finally, teachers must believe in the usefulness of, importance of and the need for tests in the teaching process and help those colleagues involved in test preparation to find ways to increase the power of language tests.

BIBLIOGRAPHY

- ABBOT Gerry & WINGARD Peter, 1981. *The Teaching of English as an International Language*. Biddles Ltd.
- ALDERSON, J. Charles. 2000. *Assessing Reading*. New York. Cambridge University Press.
- ALLEN, French Virginia. 1983. *Techniques in Teaching Vocabulary*. Oxford University Press.
- BACHMAN, F. Lyle. (1990). *Fundamental Considerations in Language Testing*. Oxford: Oxford University Press.
- BACHMAN, F. Lyle and PALMER, S. Adrian. (1996). *Language Testing in Practice*. Oxford: Oxford University Press.
- BETTERIDGE, David and BUCKBY Michael. (1983). *Games for Language Learning*. Cambridge: Cambridge University Press.
- BRADFIELD, James M and MOREDOK, H. Stewart. (1957). *Measurement and Evaluation in Education*. Macmillan.
- BROWN, H. Douglas. (1994). *Principles of Language Learning and Teaching*. Prentice Hall Regents.
- BUCK, Gary. 2001. *Assessing Listening*. New York. Cambridge University Press.
- CANDLIN, C. N. (Ed.). (1988). *Vocabulary and Language Teaching*. Longman Group Limited.
- CARROLL, B. J. (1980). *Testing Communicative Performance*. Oxford: Pergamon Press.

CELCE-MURCIA, Marianne & ROSENWEIG, Fred. (1979) *Teaching English as a Second or Foreign Language*. Rowley Massachusetts: Newbury House Publishers Inc.

CELCE-MURCIA, Marianne & McIntosh Lois. (1979). *Teaching English as a Second or Foreign Language*. Rowley Massachusetts: Newbury House Publishers.

CHASE, I. Clinton. (1999). *Contemporary Assessment for Educators*. New York: Longman

DAWSON, Colin. (1984). *Teaching English as a Foreign Language*. Edinburgh: Thomas Nelson and Sons LTD.

ECKSTUT, Samuela and KAREN Sorensen. (1992). *What's in a Word?: Reading and Vocabulary Building*. New York: Longman Publishing Group

ELLIS, Rod and TOMLINSON Brian, 1980. *Teaching Secondary English*. Longman Group Limited

ERSÖZ, A. (June, 2000). *Six games for EFL/ESL classroom*. *The Internet TESL Journal*, Vol. VI, No. 6

FINOCCHIARO, M. & Brumfit, C. 1983. *The Functional-Notional Approach*. New York: Oxford University Press.

GAIRNS, Ruth & REDMAN, Stuart. (1986). *Working With Words*. Cambridge: Cambridge University Press.

GREDLER, E. Margaret. (1999). *Classroom Assessment and Learning*. New York: Longman

GRONLUND, E. Norman. (1998). *Assessment for Student Achievement*.
Boston: Allyn& Bacon

HARMER, Jeremy. (2001). *The Practice of English Language Teaching*.
Harlow: Longman.

HARRIS, P. D. (1969). *Testing English as a Second Language*. McGraw Hill.

HARRIS, P. David.(1969). *Testing English as a Second Language*. New York
McGraw- Hill, Inc.

HEATON J.B.1984.*Language Testing*. Great Britain, Modern English Publications.

HEATON, J. B. 1988. *Writing English Language Tests*. New York. Longman.

HEATON, J. B. 1989. *Classroom Testing*. Longman.

HEATON, J.B. (1990). *Classroom testing*. New York: Longman Inc.

HEATON, J. B. (1990). *Classroom Testing*. New York: Longman

HEATON, J. B. (1990). *Writing English Language Tests*. New York: Longman

HENNING, Grant (1987). *A Guide to Language Testing: Development, Evaluation
Research*. Massachusetts: Newbury House Publishers

HUANG, F. (December, 1996). *The role of play in the accelerated learning
ambience in the teaching of foreign languages*. *Teaching Review*, No.2
Australia: The University of Newcastle. Retrieved from:
<http://www.cityu.edu.hk/lib/bamboo/abstract/doc121.htm>

KIRLAR, Fatma.(2001).*Teaching Vocabulary Through Songs*. {On- line}
Available: <http://www.geocities.com/stuncel2001/lyrics.html>

KITAO, S. K. and KITAO, K. (2001) *Testing Reading*. {On- line} Available:
<http://ilc2.doshsha.ac.jp/users/kkitao/library/article/test/reading.htm>

(December 15)

LEE, S. K. (January - March, 1995). *Creative games for the language class*. *Forum*.
 Vol. 33, No. 1, 35.

KHODADADY, Ibrahim. (1999). *Multiple Choice Items in Testing: Practice and Theory*. Tehran: Rahnama Publications.

KUBISZYN, Tom and BORICH Gary.1993.*Educational Testing and Measurement*
 New York. Collins College Publishers.

LADO, Robert.1961.*Language testing*. Longman.

LARSEN-Freeman, Diane.2000.*Techniques and Principles in Language Teaching*.
 Oxford. Oxford University Press.

MADSEN, S. Harold.1983. *Techniques in Testing*. Oxford University Press.

MALEY, Alan and DUFF, Alan. (1982). *Drama Techniques in Language Teaching*. Cambridge: Cambridge University Press.

MARSHALL, John Clark and HALES, Loyde Wesley. (1972). *Essentials of Testing*. Addison-Wesley.

McMALLUM, P. George. (1980). *101 Word Games*. New York: Oxford University
 Press <http://www.geocities.com/stuncel2001/lyrics.htm>.

McNAMARA, Tim.2000. *Language Testing*. Oxford, Oxford University Press.

McNAMARA, Tim. (2000). *Language Testing*. Hong Kong: Oxford University
 Press

MITHCHEL, Ruth. (1992). *Testing for Learning. How New Approaches to Evaluations can Improve American Schools*. Free Press.

MOOROW, Keith.(1982). *Communicative Language Testing*. London: Oxford.

NATION, I.S.P. 2001. *Learning Vocabulary in Another Language*. Cambridge: Cambridge University Press.

NUNAN, David. (1998).*Language Teaching Methodology: a textbook for teachers*. New York: Phoenix ELT.

PUMFREY, P. D. (1989). *Measuring Reading Abilities*. London: Hodder and Stoughton.

READ, John. (2000). *Assessing Vocabulary*. Cambridge: Cambridge University Press.

RICHARDS, C. Jack. (1976). *The Role of Vocabulary Teaching*. TESOL Quarterly, Volume 10, No.1, pp.77-88.

SCHMITT, Norbert & McCarthy Michael.1997. *Vocabulary: Description, Acquisition and Pedagogy*. Cambridge University Press.

SCHMITT, Norbert. 2000. *Vocabulary in Language Teaching*. Cambridge: Cambridge University Press.

SHEPHERD, James F. (1987). *College Vocabulary Skills*. Boston: Houghton Mifflin Company.

STERN, H.H. 1983. *Fundamental Concepts of Language Learning*. Hong Kong: Oxford University Press

- TAYLOR, R. George. (2003). *Informal Classroom Assessment Strategies for Teachers*. Maryland: The Scarecrow Press, Inc.
- TRICE. D. Ashton. (2000). *A Handbook of Classroom Assessment*. New York: Addison Wesley Longman, Inc
- TURGUT, M. F. (1995). *Eğitimde Ölçme ve Değerlendirme Metotları* [Measurement And Evaluation Methods in Education] Ankara: Yargıcı
- UNDERHILL, Nick. (1987). *Testing Spoken Language*. Cambridge: Cambridge University Press
- VALETTE, M. Rebecca. 1977. *Modern Language Testing*. London. Brace Jovanovich, Inc.
- WALLACE, Michael. (1982). *Teaching Vocabulary*. Oxford: Heinemann Educational Books
- WEIR, C. J. (1990). *Communicative Language Testing*. London. Oxford.
- WEIR, J. Cyril. (1990). *Communicative Language Testing*. Prentice Hall International
- WEIR, J. Cyril. (1993). *Understanding and Developing Language Tests*. England: Prentice Hall
- WILF, Selma. (1988). *Basic Skills for Effective Reading*. Englewood Cliffs: Prentice-Hall, Inc
- WRIGHT, Andrew. (1989). *Pictures for Language Learning*. Cambridge: Cambridge University Press

A decorative graphic consisting of several parallel diagonal stripes in a light pink color, creating a stylized 'X' or 'K' shape that frames the central text.

APPENDICES

APPENDIX A

Questionnaire for Secondary School Teachers

Dear colleague,

The purpose of this questionnaire is to find out the ways that teachers test vocabulary in Macedonian Secondary Schools.

In order to do so, I prepared the following questionnaire and ask you to take few minutes in order to fill it in. Your support will help me to prepare a report on language testing and develop a substantial proposal for an improvement of the situation all over Macedonian State High Schools.

Thank you very much for your support.

Yours sincerely,

Rufat OSMANI

1. What skills/ elements do you test?

Reading Comprehension	Writing	Listening	Speaking	Grammar	Vocabulary	Subject Specific Skills	Translations	Others (please specify)

Please indicate the approximate percentage

2. Which item types do you use for vocabulary testing?

Multiple Choice	Translating	True-False	Summarizing	Completion Items	Matching	Definitions	Items Involving Synonyms

Please indicate the approximate percentage

3.For which purposes do you use language tests?

Entrance Tests	Placement Tests	Achievement Tests	End Course Tests	Others (please specify)

4. During your college studies, did you take any subject on Language Testing and Assessment?
YES ()
NO ()
5. Have you ever attended any seminar or conference on Language Testing and Assessment?
YES ()
NO ()
6. Do you co-operate with colleagues in making the tests, or do you ask them to bring out their comments on given tests?
YES ()
NO ()
7. Do you write tricky item that use close spelling or sound contrasts in you multiple choice type tests?
YES ()
NO ()
8. Do you pair words of opposite meaning with the right answer?
YES ()
NO ()
9. Do the stem of a particular item gives the students the problem before they consider the optional responses?
YES ()
NO ()
10. While constructing multiple choice items do you think all distractors should be plausible equally?
YES ()
NO ()
11. Do you include specific determiners (words such as “all”, “none” or “no”) in your True- False Items?
YES ()
NO ()
12. Do you use words such as “large”, “a long time ago” and “often” in your True-False Items?
YES ()
NO ()
13. Do you prefer to use homogenous material in your matching items?
YES ()
NO ()

14. Do you take statements from the textbook to write completion items?
 YES ()
 NO ()
15. Do you think all of the alternatives should be grammatically consistent with the stem of the item?
 YES ()
 NO ()
16. Which of the two given alternatives do you use in multiple choice items?
 all of the above ()
 none of the above ()
17. While item evaluations which of the following do you consider?
- Does each item measure an important learning outcome included in the tests specifications? ()
 - Is each item type appropriate for the particular learning outcome to be measured? (....)
 - Does each item present a clearly formulated task? ()
 - Is the item stated in simple, clear language? ()
 - Is the item free from extraneous clues? ()
 - Is the difficulty of the item appropriate? ()
 - Is each test item independent and are the items as a group free from overlapping? ()
 - Do the items to be included in the test provide adequate coverage test specification? ()
 - All above ()
 - None ()
18. What kinds of the below considerations you are concerned with when considering the validity of assessment results?
- Content representativeness
 - Criterion relationships
 - Construct evidence
 - Consequences of using
 - None
19. How do you select words to teach/test?
- From the syllabus ()
- Own selection ()

20. Do you ever calculate the mathematics mean of the test scores?

- Never ()
- Seldom ()
- Sometimes ()
- Usually ()
- Always ()

21. Do you think this calculating mathematics mean of test scores can help to the teachers?

YES ()

NO ()

22. Do you ever compare the results of your test with other teachers?

- Never ()
- Seldom ()
- Sometimes ()
- Usually ()
- Always ()

23. Do you take into consideration the test results in the constructing of further tests?

YES ()

NO ()

24. Do you think that objective tests are easier and less time consuming to answer than subjective tests?

YES ()

NO ()

25. What do you think the extent of objective tests to measure the success would be?

- A great deal ()
- Quite a lot ()
- A lot ()
- A little ()
- Very little ()

26. Do you think the tests are in accordance with the course objectives?

YES ()

NO ()

27. Do you give clear and right instructions in your tests?

YES ()

NO ()

28. Is the given number of test/s at your classes sufficient enough to identify students' success?

YES ()

NO ()

29. Apart from testing the students' success, what other purposes for do you use tests?

- To identify students' learning difficulties ()
- To motivate students for better learning ()
- To look up for a new teaching approach ()
- Others (please state).....

30. Are you satisfied with the tests in most cases?

- absolutely ()
- most likely ()
- possibly ()
- not likely ()
- not at all ()

APPENDIX B**ANSWER KEY FOR THE SUGGESTED TESTS****Type 1. Fill in the blanks**

1. sober
2. sensible
3. mood
4. lifestyle
5. risk
6. disloyal
7. pitiable
8. advice
9. immature
10. proud
11. pressure
12. remote
13. illegal
14. drugs
15. consume

Type 2. Matching

1. gland
2. embarrassment
3. profit
4. slave
5. armour
6. tax
7. mend
8. tin
9. compromise
10. decrease
11. fodder
12. wealthy
13. pottery
14. dull
15. depression

Type 3. Reading Comprehension

1. c) man
2. b) insomniac
3. a) a person who is not able to read
4. b) someone who beliefs that God does not exist
5. c) a person that never drinks alcohol
6. d) to make sound
7. a) the way you feel
8. b) to sleep longer than you intended
9. d) a headache and feeling of sickness
10. b) an outdoor party

Type 4. There may be various answers to these questions.

Type 5. Multiple Choice

1. a) sink
2. c) engaged
3. c) fields
4. d) borrow
5. a) elected
6. c) broke
7. b) shattered
8. b) up
9. a) hilarious
10. c) sensible

Type 6. Defintions

1. a) to spend more money than you can afford
2. c) a sum of money you borrow
3. b) an amount of old and useless things
4. d) a person with no home, job and no money
5. a) awake or you realize what is happening
6. c) connected with what is talked
7. a) to reach the complete development
8. b) to allow something
9. b) to become smaller
10. d) not do something at all

Type 7. Completion Items

1. favourite
2. lively
3. honest
4. hunting
5. cosmopolitan
6. continental
7. reputation
8. advertising
9. publishers
10. magnet

Type 8. Synonyms

1. b) accept
2. a) exist
3. d) achievement
4. a) to use
5. a) wager
6. d) agreement
7. b) belief
8. a) wise
9. b) courageous
10. a) courteous

Type 9. Associated Words

speed

popular

sound

happiness