

## **Introduction**

Testing foreign language learners has a major importance in the teaching processes. It is the only way to recognise the testees' level in order to decide their future in the English language. To hold this testing, test designer must recognise its theoretical aspects in terms of its narrow and broad definitions, its common approaches, its major traits, its design and the different testees' levels. In this chapter, we will try to clarify these aspects. So, we will know the dimensions of foreign language testing generally all over the world, in order to find out at the end of this dissertation the features that are followed in Biskra University

## **I.1. Definitions of Testing English Language Learning**

Foreign language learning tests have two senses: broad and narrow; however, both complete each other. Both are ways to evaluate the testing process globally and specifically.

### **I.1.1. its broader sense**

In its broader sense, testing is a useful way to evaluate the educational operation and the qualifications of its members. It is to assess its passive members who are the foreign language learners in terms of their mastery of the foreign language program and the ability to use it by means of marking. It also includes its active members, who are teachers, in terms of their effectiveness by presenting this program. Furthermore; it is basically an assessment and measurement of foreign language program itself Finch, Sampson, Miller (2000: 1) describe, in details, this sense as follows,

*“ .....is an essential part of every language program and of most language classroom...being used to measure language aptitude, proficiency, placement, diagnosis, progress and achievement, and providing feedback for the program evaluator(s), washback information for the teachers and students, and motivational washforward implications for all concerned. »*

### **I.1.2. its narrower sense**

Whereas; in its narrower sense, the term “*testing*” has various explanations and interpretations, it is used by teachers and testers to mention their use of certain procedures for measuring the learners’ language abilities. The latter are expressed through appropriate structure of language to convey an appropriate meaning using two

forms of output: written or spoken. Tests give chance to the testees to check their knowledge of English structure and vocabulary as well as their ability to convey meanings using these linguistic features by means of the communication skills: writing, reading, speaking and listening (Paul Rea-Dickens2000: 378).

These definitions are influenced by the current communicative approach to language testing that is based on communicative competence. It is also a way for decision-making to place learners in the appropriate position in order to improve their levels, and to remedy their deficiencies in language ability. Skehan (1998:153, cited in Finch, Sampson, and Miller 2000:1) states that it is a 'systematic method of eliciting performance' that can be used as procedure to decide the testees' levels.

In testing foreign language, a tester evaluates one's testees' use of language structure as well as the meaning they want to convey through this structure. And these are expressed using their four skills: listening, reading, writing and speaking; each alone or in an integrative way. At the end, their evaluation is generated in a certain marking criteria.

## **I.2.Approaches to Testing English Language Learning**

Applied linguistics and educational measurement co-operated to provide certain appropriate approaches to test language learning. The first one has introduced an understanding of the nature of the presented language, and the second has provided the testers with certain appropriate formats for their test design. As a result, there were two approaches to testing language: structuralist and communicative.

### **I.2.1. Structuralist Approach to Testing English Language Learning**

It was a natural consequence of linguistics developments at that time (i.e. in the late 1960's and the early 1970's), of course, in terms of **structuralism** that was based on the structure of the language rather than its meaning and the ability to use it.

Concerning testing foreign language, this approach is based on three main linguistic elements: *form, lexis and phonology*. The knowledge about language was identified by these linguistic elements, and so did its test design. In addition, there were two main frameworks used as bases of test design for this approach: Lado(1961) and Carroll(1961). Both of them view language as ‘ a set of separate parts’ (Rea-Dickens: 378).

This approach was regarded as objective for it is based on objective test formats using objective item types like: MCQ's that are considered as *discrete-point item* (ibid: 379). And this can provide reliable tests as result, not like the case of *open-ended item* (ibidem) like discussion in Essay writing.

In the late 1960's and the early 1970's, the scholars have realised the irrelevance of the traditional approach to language testing that is based on: *form, lexis and phonology*. They found that there must be an approach to convey the main purpose of

teaching and testing language i.e. communication. Then, it was the emergence of communicative approach.

### **I.2.2. Communicative Approach to Testing Foreign Language Learning**

Testing foreign language learning was being for so long developed. It was, in the past, based on the learners' knowledge of the target language and it involves both grammar and vocabulary. However; nowadays the view towards testing foreign language learning has been changed to testing learners' communicative ability to use this knowledge appropriately.

Dell Hymes, in 1972, introduced a new notion for language testing: communicative competence. It is through which one must consider the grammatically correct sentences that are communicatively appropriate, as shown by S.K.Kitao and K.Kitao (1996:1):

*“The basic idea of communicative competence remains the ability to use language appropriately, both receptively and productively, in real situation.”*

The concept was adopted by so many scholars, in the field of language teaching and testing, the models of Canale and Swain (1980) and Bachman (1990) are best instances on that. Thus, communicative *language ability* became an academic and a specific purpose of foreign language testing, and a principal one for communicative approach to both language teaching and testing.

### **I.2.2.1. Communicative Language Ability**

Communicative language ability is now the main goal of foreign language learning. It denotes the learners' ability to communicate using a spoken or written language in terms of their productive and receptive skills; however, this concept has received many definitions that determine its nature and limits. It means the knowledge about language and the ability to use it when communicating in various settings and situations (Hedge 2000: 45). It is also considered as the learners' abilities to use forms of the target language in their appropriate contexts. And it is their abilities to communicate with an audience (readers/hearers) using spoken or written forms of language in real situations.

Communicative competence (another term) is based on the four skills of language, and they are regarded as its parts: listening, speaking, reading and writing. They interact in communicating using the target language. Rea-Dickens (2000: 379) denoted the interaction between these parts,

*“These features, regarded as parts of communicative language ability, reflected the characterisation of language use as interaction-based, and stimulated by a specific communicative purpose on the part of the speaker or writer to convey a particular message to a pre-determined audience.”*

For example, an interviewee listens to his interviewer's questions and comments, and then he replies. In the same time, they take down some notes, and read some quotations and comments from their notebooks.

### **I.2.2.2. The Components of Communicative Language Ability**

#### ***I.2.2.2.1. Linguistic Competence***

It is based on the learners' knowledge about the language form [spelling, pronunciation, vocabulary, word formation, grammatical structure and sentence structure]. It is considered by researchers as an important part in communicating operation like suggested by Hedge (2000: 47):

*“... Linguistic competence involves knowledge of spelling, pronunciation, vocabulary, word formation, grammatical structure, sentence structure and linguistic semantics.”*

It represents the language structure that is a basic component in communicative language testing; moreover, it is the first to be tested.

#### ***I.2.2.2.2. Pragmatic Competence***

It is related to the appropriateness of the language use, especially in terms of its social context. Or it is the relevance of an utterance or a discourse to social rules of the native speakers of the target language that have their own traditions and ways of life and their own social relations. All of them are expressed in their language. Hymes (1972, cited in Hedge 2000: 50) describes them,

*“...to know when to speak, when not, what to talk about with whom, when, where and in what manner”*

In this context, one must consider in one's evaluation the appropriateness of a sentence to the social conventions of the English-speaking countries.

#### ***1.2.2.2.3. Discourse Competence***

It is the ability to use certain language strategies and signals, like to start a discourse, how to develop a topic through logical steps and how to reformulate a discourse by finding out its causes, consequences and its structure, and compare it with other heard or written discourses. It was called like this by a group of scholars as: Canale and Swain (1980), Faerch and Phillipson (1984) and was also called *textual competence* by Bachman (1990, cited in Hedge 2000: 50). Beale (2004) defines it as a way for ,

*“Shaping language and communicating purposefully in different genres (text types), using cohesion (structural linking) and coherence (meaningful relationship).”*

#### ***1.2.2.2.4. Strategic Competence***

It is the learners’ ability to hold a conversation without hesitating or mistaking. Through this conversation, the learners can express their intentions or opinions easily, confidently and successfully. This competence is presented in the learners “communication strategies” (Hedge 2000), and it is based on the previous competencies. Canale and Swain (1980, cited in Hedge 2000: 25) described it as:

*« How to cope in an authentic communicative situation and how to keep the communicative channel open »*

It is primarily used in oral tests where the testees must talk freely without cease in order to check their abilities to hold long conversation without making many mistakes.



#### ***1.2.2.2.5. Fluency***

It is another important component in the communicative language ability. And it is related to speaking skill. It accounts for the learners' ability to link the parts of speech together in a certain logical and grammatical order easily and confidently without hesitation or lapses. Furthermore, it is the learners' abilities to respond to direct or indirect questions rapidly and easily. Thus, fluent learners can hold conversations, exchange questions and answers, pronounce well and respect the stress and intonation. Hedge (2000: 54) considered it as 'language production' in a spoken form, and it is the ability to use the 'units of speech' easily appropriately and without hesitating. And it is the rapid use of language structure to convey certain meanings appropriately.

#### **1.2.2.3. Testing Communicative competence**

According to Canale and Swain (revised in 1983 by Canale) testing communicative competence is based on testing the learners' abilities to use the target language through the four skills of communication that are listening, speaking, reading and writing. All together are to be tested on the bases of the components of communicative competence.

These skills are used as signs for the level of communicative language ability of language learners. And one may test each skill alone or in an integrative way like the case of receptive skills (reading/listening), and productive ones (writing/speaking). Weir (1990:7 cited in Alderson & Bachman: 9) denotes that when we test communicative language ability, we take samples of testees' performance. These latter are held under certain constraints and specific contexts in order to check their communicative capacity and language ability. Testing communicative competence

needs real contexts for its audience to guarantee its communication trait. And it is usually focused on two or more skills at once.

#### ***1.2.2.3.1. Testing productive skills***

Testing productive skills includes both *speaking* and *writing* skills. It focuses on the appropriateness and relevance of the written or spoken utterance or discourse to the rules of grammar, society and culture of the target language (i.e. communication appropriateness). Thus, the relevance of the produced language to the communicative competencies in real-like situations is the focus of Communicative language tests of productive skills. It focuses, first of all, on the appropriateness; then comes ability to form grammatically correct sentences S.K.Kitao and K.Kitao (1996:1). They are mostly involved with one of the receptive skills, or both of them, in their testing.

##### **1.2.2.3.1.1. The Speaking skill**

Many tasks can be used for such a skill that can take different shapes: individual, pair and group. Its assessment is based on the grammar use, vocabulary use, organisation, content, pronunciation and fluency. The evaluation will reflect the testees' ability to manipulate the target language appropriately. Harmer (2001: 269) considers it as:

*“The ability to speak fluently presupposes not only knowledge of language features, but also the ability to process information and language ‘on the spot’”*

To facilitate holding this test, one needs to record the testees' speech through a tape-recorder or a camera.

#### **I.2.2.3.1.2. The writing skill**

Writing skill can be done using many tasks like writing letters, paragraphs, essays, news reports, stories, filling gaps, correcting mistakes....etc. In such a skill, many elements must be involved as grammatical ability, lexical ability, mechanical ability ‘punctuation, spelling...’, stylistic skills, organisational skills and judgment of appropriacy (S.K.Kitao and K.Kitao1996 :1)

Perhaps the major problem, in this skill, is its evaluating, especially in terms of objectivity. Thus, testers must put a general scale that covers a set of writing aspects like organisation, grammar, vocabulary, content and medium. Despite of that, testers can face many problems like finding good grammar use and poor vocabulary use (ibid: 5).

#### **I.2.2.3.2. Testing Receptive Skills**

It includes both *listening* and *reading* skills. It focuses on the learners’ abilities to comprehend the communicative intention of either a speaker or a writer. In this case, the learners play the role of receivers of speech using their listening skill, or a piece of writing using their reading skill. Harmer (2001: 199) affirms that,

*“Receptive skills are the ways in which people extract meaning from the discourse they see or hear.”*

Moreover, in these two skills, the testees are to extract ideas, by understanding general or specific meaning.

#### **I.2.2.3.2.1. The Listening skill**

Listening skill can be held in either two major ways: first, it can be done through taped materials (video/audio tapes). Second, it can be held through “live listening” (Harmer 2001). The testees, here, are going to be tested from the lowest parts of the language (discriminating between sounds, stress and intonation) to the complicated ones (comprehending sentences and discourses). In addition, we can include various questions in such skill like: multiple choices, short-answer questions, sentence completion, table completion, classification, matching.

The evaluation will be either on a written or a spoken output. That is, the testers can give certain written questions to be answered through their listening to recordings (video/audio). Or through live listening to oral lectures, story-telling by the testers. This latter can hold a conversation with another tester, or if possible, a native speaker, and let the students listen to it to intervene or to be asked some written or spoken questions about its topic. Otherwise, they can make its output in a spoken way.

There are many ways to test listening, however; it cannot be tested alone, but with other skills, specially speaking and writing. Furthermore, it needs to be in real context.

#### **I.2.2.3.2.2. The Reading skill**

In this skill, students can be given a piece of writing « poetry, extracts from books, newspapers or magazines ». Here, the tester can present activities through which the testees are going to deal with the structure of the text« grammar, vocabulary ». This can be called \*testing intensive reading\* (Harmer 2001). Otherwise, they can be tested through their comprehension of the main ideas, and the tasks, here, will be on

manipulating the ideas of the text by analysing and judging them. This can be called \*testing extensive reading\* (Harmer 2001).

Of course, the output can be taken in a written form in terms of Question/Answer, multiple choice, True/False, matching....etc, especially for intensive reading skill, whereas; in extensive reading, testees need to manipulate its ideas through writing an essay in most cases.

Its evaluation is based on their understanding and their use of English in terms of grammar, vocabulary, content and organisation.

#### ***1.2.2.3.3. Testing Integrative Skills***

However, testing communicative competence includes all the four skills without separation and in real context since the learners, when communicating, use different skills. For example, they can talk about a written lecture or a heard discourse, and criticise it orally, or in a written form. Both productive and receptive skills are integrated in communicative testing. These skills are not tested in isolation (S.K.Kitao and K.Kitao 1996: 1), they illustrate it by the following example that testees are to listen to a lecture, then they use information from written lecture. And when in a group they discuss, they use both listening and speaking skills. When they read a book for pleasure re-tell it to a friend.

Tasks of communicative language testing focus on both form and meaning in terms of language production or reception in real context. Littlewood (1999:89) comments on that, and states that,

*“In communicative activities, the production of linguistic forms becomes subordinate to higher-level decisions, related to the communication of meanings. The learner is thus expected to increase his*

*skill in starting from an intended meaning, selecting suitable language forms from his total repertoire, and producing them fluently. The criterion for success is whether the meaning is conveyed effectively.”*

Communicative approach to language testing accounts for the use of more realistic tasks to check the learners’ abilities to be good communicators when they deal with native speakers of the target language as an instance. Finally, communicative language test must be an extension to communicative language program; thus, we can evaluate its success or failure through this test and its takers.

### **I.3. The Essential Traits of A Good Test**

To gain a good test, its designer must follow the coming-up traits to guarantee its success at the end that can make it characterised by the two main traits of any good test afterwards: validity and reliability. These traits are as follow:

#### **I.3.1. Purpose of a Good Test**

Normally, each language test has a certain purpose determined by the teachers (testers). The objective of any test procedure is to evaluate the level the testees have reached through their teaching of language program, basing on certain approach they choose, consequently; they consider the approaches' procedures to test the target language learning. For example, if the testers want to measure the learners' ability to use English in real situation, they must use communicative tasks in which they involve the four skills of communication (reading, writing, speaking and listening), thus they can reach their purposes. But if they use only one or two skills, they will not arrive at their goal at all. Tester must know what he wants to measure a skill or an area of knowledge and for each there is its method (S.K.Kitao and K.Kitao1996). Thus, one must be aware of any procedure they take that has to be suitable for arriving at one's purpose at the end.

#### **I.3.2. Audience of a good test**

Audience, here, means the stakeholders (Alderson 1999) **testees**. When considering them, one must consider their characteristics and levels which influence the testing processes and its outcome. These characteristics are of psychological effects as summarised by Huit (1995, cited in McIlrath and Huit, 1995: 9),

*“...Study habit; Learning Style; Age ; Sex/Gender; Race/Ethnicity; Motivation; and Moral, Socioemotional Cognitive, and Character Development all become important in the relationship of classroom processes/behaviour and school achievement”*

Furthermore; tester must not forget the testees’ levels and sub-levels: beginners, intermediate, advanced.

Therefore, the tester must include various test items in order to give chance to all the testees with different abilities, for example, a tester makes a test that involves choices like essay writing, multiple choice, ....etc in the same test. The same as noticed by Alderson (1996: 278). So that testers have to consider their testees’ backgrounds before designing any test.

### **I.3.3. Clarity of a good test**

Clear and understandable test determines the success or failure of the testees and the test itself. If the learners do not understand the test, the testers cannot assess them in a fair way. Incorrect answers may be due to the misunderstanding of the test; not to the ignorance of the answer. The test items can be clear to test designer, but ambiguous for one’s testees (K.Kitao and S.K.Kitao 1996) .i.e. the clarity and understanding of the test items (question) make sense of the test and its purpose.

### **I.3.4. Length and variety of items**

A long test with various test items can give chance to the testees to guarantee their success. If testers make a long test full of various items that can satisfy the testees’ abilities, they will help them to succeed with less wrong answers. However, a long test



with the same types of test items may cause failure since it does not cover all the testees' abilities. Long tests are good, especially if it is divided into sections with different and various items (K.Kitao and S.K.Kitao 1996).

Some tests are long, but of the same item type that cannot be suitable for other testees of different learning styles [abstract/concrete perceivers, or active/abstract processors]. Hence, a long test with various items that respect these differences, even in the same level, can be fair for all.

### **I.3.5. Objectivity of a good test**

Objectivity in making a test is another essential trait in measuring the testees that means to put objective scale for the different aspects of the test items. For example, in marking learners' speaking skill in real situation, testers must not take into account only the learners' dare to speak English, but they must divide the scale on their fluency of speech, pronunciation, organisation, vocabulary use, grammar use...etc.

### **I.3.6. Practicality of a good test**

It is related to the testers' ability to control the test by the appropriate means that can make the testing process well. For example, tester must assess the testees' speaking skill with their peers, of course, by recording it through tape-recorder. Though this cannot be possible in measuring the group activities that needs video-tape recorder. Practical test is cost-effective and easy to administer in terms of testees number, test timing, testing, marking human and physical (Beale 2004).

Providing a good climate for holding a test guarantees good outcomes for the test and its "stakeholders" (Alderson 1996: 278). Practicality is essential, for it is the

administrative area that includes setting, marking, equipments, arrangements...etc  
(Jordan 2000: 89)

### **I.3.7. Reliability of a good test**

It is one of the characteristics of a good test that means consistent scores of the same learners through repeating the same test each time. Repeated measurements with consistent test score during a short time prove the test's reliability (Good and Brophy 1990: 686).

Reliability is dependent on: first, the clarity of the test instructions. For example, if the learners do not understand the question, they may answer incorrectly, and the testers could not rely on these scores to consider their tests as reliable. Second, they rely on the objectivity of marking because subjective judgments on the same learner that performs the same way each time may get different scores that make the test unreliable.

### **I.3.8. Validity of a good test**

The valid test is the test that arrives at the tester's goal. Whether tests actually measure what they are supposed to measure (Thrasher 1984, cited in Hedge 2000), and to make a test a meaningful tool assessment, it must have recognisable logic (Beale 2004). Hedge, N. (1997) states that,

*“It is the extent to which a test procedures is an adequate basis for decision-making”*

For example, if a tester wants to measure the learners' ability to communicate with their peers, and they use a grammar knowledge test, then one's test will be invalid since it

does not involve the receptive and productive skills that are needed to test communicative language ability.

Test validity has various types, and each refers to one aspect of foreign language learning. And perhaps the two important types for a tester are content and criterion:

#### ***1.3.8.1. Content validity of a good test***

It is based on the language program that is designed to be taught to the learners. Or, it is the extent to which the test items include the specific language program to be tested. Thus, the test content must be limited to the knowledge or skills that were presented in the foreign language program. It is the extent to which the test content covers the tested curriculum (Good and Broophy 1990: 686). For an example, if a tester taught one's learners the uses of rising and falling tones, then he/she measures them in the language laboratory by asking them to pronounce certain statements or dialogues using these tones. In this case we say it is content valid. Kaur (2000: 7) describes it as follows,

*“Content validity (whether the test is a representative sample of the content of whatever the test was designed to measure)”*

#### ***1.3.8.2. Criterion-related validity of a good test***

It is very important in foreign language testing since it really reflects the actual communicative language ability. It is more related to the scoring of the test than its content. It reflects the degree to which the testees can use the language actually and successfully in real situations. And this can be reached through repeating the same test items periodically. It is the extent to which the test predict performance of testees [behaviours] (Good and Broophy 1990: 690)

It is also divided into two other kinds: predictive and concurrent validity:

#### **I.3.8.2.1. Predictive validity of a good test**

It is related to test item, and it is the level to which the item can predict the testees' performance in the future and in real situation. In other words, it is related to the question of the test (test items), or it is the extent to which this question can predict the testees' ability to use the response in real situations successfully. Cunningham (2002: 15) suggests that,

*“The predictive validity of a test is the degree to which accurately consistently predicts the testees' future performance or behavior.”*

For example, asking them to use the rising/falling tones in a certain dialogue, their responses will predict their success in using them if they go to Britain.

#### **I.3.8.2.2. Concurrent validity of a good test**

It is the repetition of nearly the same results of the different tests items, in which all have the same purpose. Cunningham (2002: 15) also states that the agreement of results from one test format with other formats means that it is concurrent valid. i.e. the basis and the purpose of both the foreign language program and its ability are well-organised and well-studied.

Validity is regarded by scholars as a difficult task to reach when constructing and holding a test. Kelly (1955:77, cited in Finch, Sampson and Miller 2000:4 ) points out that it is extremely difficult to design a test that is truly valid. One can say, here, that any test must be controlled on behalf of the these features because they are the essentials of its success or failure; otherwise, the testers are going to fail in their main tasks that is to evaluate their testees appropriately.

## **I.4.Procedures for Test Design**

To prepare a test, tester has to consider the following features as bases for designing it: its form, its types, its item types, and its evaluation.

### **I.4.1. Forms of English Language Testing**

English language testing is held in two forms: written and oral as language is presented orally or in a written form. For communicative language testing, the two forms are involved together in testing the four communication skills: writing, reading, speaking, and listening.

#### ***I.4.1.1. Written Form of Test***

In this form, the testees express their comprehension of listening to a recorded or live speech, or reading to a piece of writing like an article, a book, a written lecture....etc. Testees, here, will express specific or general ideas about the written or spoken input through responding to certain written or spoken questions. Writing skill, here, is the only way of responding (output) in which they use their own grammar, vocabulary, style, organisation and judgements. It is also called ***pencil-and-paper tests***. Rea-Dickens (2000: 386) describes this form of test as a formal one that is held under written stimulus and response.

Their evaluation basically focuses on the comprehension of the input, and the manipulation of its ideas, then comes the evaluation of grammatical, lexical, stylistic and organisational abilities.

***Example of written tests:***

Written form of test can take different shapes. The testers can ask their testees to fill in certain gaps in a passage, to correct it , to choose the appropriate response, or to write a paragraph or an essay. Of course, basing on the input, they receive before the test.

**Example one:**

*In your opinion, what are the most important advantages and drawbacks of the three different answers given by the applied linguists to the question; what's language? (See the appendix02)*

Testing foreign language was for so long based on this form of test; however, this test is now considered insufficient and needs to be promoted and completed by involving an “oral test” into the general foreign language testing. Thus, the evaluation can be fair and valuable.

***1.4.1.2. Oral test***

Oral test was first developed by Educational Testing Services in the late 1970's. It was introduced to the traditional foreign language testing that was based on the written language. Consequently, it was put to measure the learners' ability to communicate orally. At first, the evaluation was based on the linguistic competence of the learners that includes: pronunciation, grammar, and fluency, but later on it was revised to be based on the communicative language ability. Underhill (1987:7) defines it as:

*“... A repeatable procedure which a learner speaks and is assessed on the basis of what he says. It can be used alone or combined with tests of other skills”*

This test is held in the following ways, as designed by Underhill (1987: 28-31)

1. *Learner-interviewer/assessor*
2. *Learner-interlocutor*
3. *learner-learner*
4. *learner-group*

**1. Learner-interviewer/assessor:** So many tasks can be held under this way:

**Oral report:**

In which the testee is to make an oral presentation of an important topic, it is considered as an authentic and communicative activity both for professional and academic purposes (Underhill 1987: 47).

**Reading aloud:**

That is a way to test separately pronunciation from the content of the speech (K.S.Kitao and K.Kitao 1996: 1).

**Picture story:**

The testees are to be asked to describe a group of pictures that form a story.

**Re-tell:**

The testees can also be asked to listen to a heard or written story. Then they repeat it in their own words in order to check their understanding of the told story

**2. Learner-interlocuter/learner-learner:** Two testees or a testee and the tester hold the test through:

**Interviews:**

It is a direct and face-to-face exchange between learner and interviewer (Underhill 1987: 54) in which they hold a list of oral question to be answered.

**Role play:**

It is considered as the most realistic activity. Testees, here, are to use certain functions that are needed in real communication (K.S.Kitao and K.Kitao 1996:3).

### **3. learner-group:**

#### **Discussion and conversation:**

It is where testees are asked to discuss about a topic of common interest. And it is considered as the most natural thing in the world (Underhill 1987: 45).

The choice of these tasks relies on the purpose of the test and its surrounding factors. Nelson (1998: 112) mentions that,

*“There are many elicitation categories; their inclusion for testing purposes may be based on external factors such as overall class size or time available, or internal ones based on students’ training, comfort levels, discourse models and so on.”*

This test is a form of response to stimulus that can be heard or written, in which they use reading skills, and basically their listening. And its output, which is oral (spoken), is evaluated in terms of the testees’ communicative language ability. Chen and Pooler (1995: 4) suggest that the evaluators have to take into account certain factors to determine scores for each test item. These factors are communicative competencies: functional, socio linguistic, discourse, and linguistic.

Oral test was developed to assess learners’ ability to express themselves orally. And it was set to promote the foreign language testing. This test was a natural result of current thinking that was resulted from the development of communicative approach. This approach emphasises on the concept that language can be held orally or in written form. Thus, oral test was added to the written test to convey the goal of communication. Kent (2001: 1) also states that it is necessary to assess testees’ speaking and listening skills through pre-planned communicative context. In oral test, listening is the essential



skill that takes part with speaking skill for the speaker mostly listens to some one else , and then replies orally.

## **I.4.2. Types of English Language Tests**

There are four main types of test which differ in terms of their objectives. They may be used before, during and after the presentation of foreign language program, however; all of them complete each other.

### **I.4.2.1. Placement test of English language learning**

It is a test used to assess English language learners in order to “place” them in the appropriate position (school/college).The materials to be tested, here, are: grammar, vocabulary, receptive and productive skills. And it is used to decide the appropriate program for the learners. Therefore, it is global since it involves both the learner and the program. It is widely used for academic purposes, especially for placing certain learners in their appropriate colleges or classes.

### **I.4.2.2.Diagnostic test of English language learning**

This test, in foreign language learning, is used to recognise the learners’ deficiencies in terms of their receptive and productive skills as well as their gaps in the knowledge of grammar, vocabulary, pronunciation....etc. This test is to remedy such gaps in order to run the language program as good as possible. It normally takes place before the beginning of the program. Beale (2004) defines it as a way to find out more detailed information about the testees’ ‘strengths and weaknesses’.

It is a procedure used to remedy the learners' gaps in the target language. In addition, it is broadly held inside the classroom for specific purpose related to teaching processes along the year.

#### **I.4.2.3 Achievement test of English language learning**

It occurs during the presentation of the foreign language program. The material, to be tested, is the content of this program. It measures the extent to which the learners' progress has reached as far as the language program.

It is an academic test that is widely used in most colleges. It takes the program's content as the source for designing its test items, whereas; language structure (grammar, vocabulary...etc), receptive/productive skills are regarded as secondary. Furthermore; it is mostly used with written tests where reading and writing are usually involved. It helps to recognise the testees' standard reached by testees comparing to their peers of the same level (Jordan 2000: 88).

#### **I.4.2.4. Proficiency test of English language learning**

In this type of tests, learners are to be measured in terms of their ability to use the target language, but not their achievement of the program. Davies and West (1984 cited in Hutchinson and Waters (1987) state that,

*“Proficiency testing designed to assess whether candidates will be able to perform the language tasks required of them.”*

This test is largely involved in communicative language testing. It assesses the learners' ability to use language in real situation, and receptive and productive skills are its items.

Despite of their variety, they all complete each other since they all test the same material which is foreign language learning, so each tests one of its aspects. Furthermore, they can all be involved in the assessment of the education process during the course of the year.

### **I.4.3.Types of Test Items**

The test will be good if it contains various test items types. In the following passages, we will explain the different test items types we have in foreign language tests; and which are of two types: direct and indirect.

#### **I.4.3.1. Direct Test Items Type**

It is designed by communicative language test. In this type, the test designer puts certain procedures to test the four skills: listening, speaking, reading and writing. These test items must be held in real-like situations; in which the tester measures the learners' skills receptively, productively or integratively.

Test designer can take one of the two fields, designed by Harmer (1998: 325-326), in order to provide a real context for the testees to guarantee the purpose of the communicative language test:

##### ***I.4.3.1.1. A level playing field***

This is used widely in making essays that can be either through direct question, for example, [*what are the consequences of civil war on the American social life ?*]. It can also be done in another way that makes a question clearer and clearer. For example, [*slavery in the south ,and freedom in the north of the same country led to a struggle*

*between the brothers, then ended with bloody conflict which resulted positive and negative consequences on the American social life in both sides. What are they?]*

In this task, the testers are to test their testees' comprehension of a certain presented lecture as well as their writing. Moreover, it is widely in written test.

#### ***1.4.3.1.2. Real-situation interaction***

In this field, testees' communicative skills are tested in real context (both receptive and productive skills) that they may face in future-like situations when outside the classroom, or in the country of the target language. Traditionally, testing productive skills were more general and unreal. But the development of English language learning due to the development of communicative approach helped the development of its testing as well. Nowadays, it became more and more realistic, and so do the receptive skills which should be more realistic without the use of some indirect test items (MCQ).

#### ***Some Examples of Direct Test Items:***

##### ***Writing: (K.S. Kitao and K. Kitao 1996)***

*-Gap filling.*

*-Form completion.*

*-Taking correction.*

*-Letter writing.*

*-Essay writing.*

##### ***Speaking:***

Learner-teacher (Swan and Walter 1999: 43)

*-look at the group of pictures, and tell the teacher the story.*

Learner-learner: (ibid: 29)

*-Work in pairs, you and your partner are members of the same family, or husband and wife, or flat mates or room-mates in college. Talk about how you both spent your day ...*

Learner-group: (ibid: 27)

- *Work in group of five or six. Tell other students about a bad day in your life. You can react to other students' stories, too. Examples:*

*'Oh dear!'            'That's awful'*  
*'Poor you!'         'How terrible!'*

**Reading:** (Harmer 2001:327)

- *Multiple choice questions to test comprehension of a text*
- *Transferring written information to chart, graphs, maps, etc.*
- *Matching jumbled headings with paragraphs*
- *Inserting sentences provided by the examiner in the correct place in the text*

**Listening:** (K.S.Kitao and K.Kitao1996)

Using audio aids:

- *Discriminating phoneme*
- *Discriminating stress and intonation*
- *Understanding sentence and dialogues*

Using visual aids: (Harmer 2001: 327)

- *Completing charts with facts and figures from a listening text*
- *Identifying whether speakers are enthusiastic, encouraging, in disagreement, or amused*
- *Following directions on a map and identifying the correct house or place*

Using live listening (K.S.Kitao and K.Kitao1996):

- *"For students who will be using in schools where it is the medium of instruction, there will be situations where they need to listen to lectures or talks in English and take notes and/or answer questions on the lecture or talk"*

**Integrative skills:**

1. *Speaking/listening:*

1. *Reading/writing*

2. *Listening/writing*

3. *Reading/speaking*

These test items are of a major purpose i.e. the learners' abilities to communicate using the four skills (speaking, writing, listening and reading). These skills cannot be only tested separately, but collectively as well.

#### **I.4.3.2. Indirect Test Items Type**

It is usually adopted in the above items, and in testing grammar and vocabulary. Moreover, it was used by the structuralist approach to testing the learners' knowledge about the target language. They are used to check the testees' language beyond their receptive and productive skills. Harmer (2001: 322) stated them as a way to measure testees' knowledge through their receptive and productive skills.

There are various indirect items that are considered more objective.

#### ***Some Used Indirect Test Items:***

##### **I.4.3.2.1. Multiple choice questions (MCQ) test item**

It is a widespread instrument for many years. It is widely used in testing vocabulary and grammar. Cohen and Wallack describes it as the most common test item and the most objective one, for it gives a number of alternative choices that testees are to choose the correct one. In such an item, the testers can give more than 4 choices as possible answers to a question.

#### **Examples of MCQ test items**

Example one:

*This is a real diamond, not \_\_\_\_\_*

1. *a compromise*
2. *a conversion*
3. *an alienation*

4. *an imitation*

Example two: (Irvine-Niakaris, C 1997)

A: "did you go to class yesterday?" B: "Yes, and you \_\_\_\_\_ gone, too!"

1. *Shouldn't have*
2. *Would have*
3. *Wouldn't have*
4. *Should have*

It is advantageous in one hand, but in another is considered as invalid and unreliable of its high scores. Harmer (2001: 323) mentions that:

*"MCQ's are still widely used, but though they score highly in terms of practicality and scorer reliability, their 'validity' and overall 'reliability' are suspect"*

This item was used largely in the structuralist approach to language testing, and that considered till now as the most objective for its high reliability.

#### **I.4.3.2.2. Cloze procedure test item**

In such an item, the tester can present a passage that is full of blanks which are deliberately deleted to check the testees' ability in terms of grammar, vocabulary and reading comprehension; thus, it can be considered as 'an integrative testing item'. Harmer (2001: 323) considers them as 'ideal' direct items; however, 'integrative'. It can be prepared, but it cannot give a true evaluation about the testees' 'overall' knowledge. It is so useful in achievement, proficiency and placement test, however; it is basically used to assess vocabulary use in its context.

**Examples of cloze procedure test item:**

Example one: (*ibid*: 324)

*They sat on a bench attached 1 \_\_\_ a picnic table. Below them they 2 \_\_\_ see the river gurgling between overgrown 3 \_\_\_. The sky was diamond blue, with 4 \_\_\_ white clouds dancing in the freshening 5 \_\_\_. They could hear the call of 6 \_\_\_ and the buzzing of countless insects. 7 \_\_\_ were completely alone.*

Example two: (*Ur1996*: 13)

*The family are all fine, though Leo had a bad bout of flu last week. he spent most of it lying on the sofa watching \_\_\_\_\_ when he wasn't sleeping!*

*His exams \_\_\_\_\_ in two weeks, so he is \_\_\_\_\_ about missing school, but has managed to \_\_\_\_\_ quite a lot in spite \_\_\_\_\_ feeling ill.*

However, such an item cannot cover all what is needed to be a good communicator. It cannot test the level of testees' fluency and all their skills' deficiencies

**I.4.3.2.3. Transformation and Paraphrasing test item:**

In this item, the testees can be asked to rewrite a sentence or a passage with their own words without changing their main ideas. Therefore, they have to comprehend their meanings carefully, then to form their equivalents that are to be grammatically correct. For an example: tester can give a short story of one or two paragraphs which they will understand, then they will re-narrate individually without referring to the main passage. It is good in terms of its measuring the testees' ability to comprehend a spoken or written passage in one hand. In another hand, it can check their ability to manipulate the language grammar, vocabulary and organisation.

**Examples of Transformation and Paraphrasing:**

Example one: (*Harmer2001*: 324)



*-I'm sorry that i didn't get her an anniversary present.*

*-I wish \_\_\_\_\_*

Example two:

*Transform the charts to a written composition*

#### **I.4.3.2.4.Sentence /Sentences Re-ordering test item**

In this item, the testers can provide a group of mixed words, and ask them to re-order them appropriately. They can also be asked to re-order a group of sentences that form a passage appropriately respecting its order and its sentence links.

It is, in part, good for it measures the testees' ability to form a grammatically correct sentence, and it can also measure the testees' ability to organise their language production. However, it is not sufficient to be a test for the communicative ability of the testees.

#### **Examples of Sentence /Sentences Re-ordering**

Example one: (*ibid*: .325)

*Called/ I/ I'm/sorry/wasn't/when/you*

---

Example two: (*see the appendix04*)

*...2) Re-order the following sentences and write a coherent paragraph*

- Because there is a guide who gives them information.*
- " le bardo" is a museum in Algiers*
- They can take photos,*
- where visitors can see traditional clothes*
- but they must not touch anything*
- And other objects of art.*

#### **I.4.3.2.5. Essay test item**

It is an important indirect test items for its wide use. It shows the testees' abilities to get the appropriate answer using their appropriate competencies: linguistic, discourse, strategic ...etc. And on these bases, the testees are to be evaluated. In addition, it is mostly directed to learners of advanced levels for they use the target language well.

In this item, testees are to write a composition of a usually limited number of lines, about certain topic, that are divided into three parts: introduction [ the framework of the topic and the problematic], body [it contains the details about the topic], and finally, the conclusion[it is the sum-up of the details and an answer of the problematic]. The problem that faces the testees in such an item is the terminology used for the way to manipulate the ideas through an essay. Thus, they need to know about them in order to avoid misunderstanding of the test. Alderson (1999: 59) notes its main drawback, that is, terminology; especially 'for conventions behind the technical use of words' that are used for essay developing like analyse that must be clear for the testees before holding the test.

There are many *key words* that are used by test designers that determine the way of manipulating the ideas in the essay they are asked to do. These are some of them:

***Analyse, assess, compare, contrast, criticise, define, describe, discuss, enumerate, explain, justify, outline, list, review, summarise...etc.*** (*Key words in essay topics and exam questions*)

**Examples of essay items:**

Example one: *(See the appendix02)*

*Read the following essay questions then choose only one.*

*.....2- Discuss the British electoral system in the light of its organisation and legislation.*

Example two: *(See the appendix02)*

*Write an essay of 350 words on one of the following topics:*

*....II. Discuss the symbolic meaning of the “two roads” in the forest and “the wall” in Robert Frost’s two poems and speak about his genius in writing about simple things to give them wider applications.*

Essay item is regarded as the most difficult to be scored and evaluated, so that it needs much time and more concentration for a fair evaluation, and in order to attain reliable scores as a result; as described by Cohen and Wallack:

*“This type of item, however, is difficult to score reliably and can require a significant amount of time to be graded.”*

Normally, the learners have to have an idea about the purpose of using each of these key words.

These items are directed to written test more than oral ones, thus they need direct test items to complete the purpose of foreign language testing.

#### **I.4.4. Evaluation of English language tests**

Test procedure is not normally directed only to evaluate the *learners' language ability* or *knowledge about this language*, but also to evaluate *language program* as well. Both of the learners and the program are to be involved in evaluating the gathered data through the test procedures. Hutchinson and Waters (1987:145) state that there are

two forms of evaluation that cannot be distinct, the first one is of the learner, and the second is of the course itself. Thus; test outcomes are a mirror of the language program level through its takers outcomes (scores). Evaluation does not only consider the learner and the program, but also the test itself as well. The evaluation of the test determines its validity and reliability.

#### **I.4.4.1. Types of evaluation**

The evaluation of the testees can be objective or subjective, and this depends on the test item types the tester holds. As a result, we have two types of evaluation:

##### ***I.4.4.1.1. Objective Evaluation***

This can be reached with some indirect test items types (see 1.3.4.2.) Evaluator can find it easy to evaluate one's testees. Alderson (1999: 106) denotes that objective evaluation can be reached through using certain direct test items such as: multiple-choice questions, true-false, error -recognition..etc that means the questions that have precise answers. To get this evaluation is something approximate since these test items types cannot really cover the main language test purpose [communication]. Tester, then, needs more complicated outputs that clearly display the testees' language abilities.

##### ***I.4.4.1.2. Subjective Evaluation***

Alderson (1999: 107) notes that subjective evaluation is usually held for evaluating the writing or speaking skills [productive skills]. Testers are required to judge more complicated items. They are to assess the extent to which they can make a

given task. Tester, here, needs a rating scale that contains numbers, letters or other labels, such as: excellent, good, bad...etc.

#### **I.4.4.2. Interpreting the Evaluation of the test**

Tester can also add statements to each point in scale. As said the evaluation is “*the process of making value judgment*” (unknown author). This value judgment is interpreted in what is called **MARKING CRITERIA**, however; some call it scoring and some call it grading. It is reflected in a scale that distributes the marks for each item according to its importance in the whole test.

#### **I.4.4.3. Ways of marking the test outcomes**

There are various scales for evaluating EFL testees. They differ in their concrete material. Generally, there are three main scales that are as follows:

Grading can be through **total scale** like giving 20/20. It can also be based on **percentage scale** like 100%. Furthermore, it can be on the form of **category weights** like < A, B, C>.

Evaluation is regarded as the difficult task in the test procedures. Its objectivity or subjectivity can also be due to other factors such as the examiners evaluation training that is needed for all EFL testers to reach the goal of the foreign language teaching.

## **I.5. English Language Learning Levels**

Teaching foreign language must consider learners' levels in terms of the target language, and so does their testing. As a result, the categorisation of the learners' levels has a great benefit to characterise the level of foreign language programs as well as its testing which will describe the degree of the learners' level. Moreover; there are different models of categorisation; however, the majority of scholars agree on three main levels: *beginner, intermediate, advanced*.

### **I.5.1. Beginner Level**

In this level, the learners may not know English, or they are unable to communicate only by some language materials that were learned. It is the characteristics of the foreign language learner in the early years of learning (middle schools in Algeria). This level is also divided into two degrees; (see fig1). False beginner who cannot use any English. Furthermore, elementary who are able to use English by some basic language materials?

In this level, testers can check the learners' ability in terms of their knowledge of some basic materials in terms of grammar, word order and some vocabulary. Thus, the learners are to be tested easily, but this does not really confirm success.

### **I.5.2. Intermediate Level**

In this level, the learners are able to comprehend a simple heard speech or a read piece of writing. Not only this; but they are also able to hold a simple dialogue or conversation, and write simple piece of writing as well. Harmer (1998:12) indicates:

*“...intermediate suggests a basic competence in speaking and writing, and an ability to comprehend a fairly straightforward listening and reading.”*

Their testing is based on their fluency and general comprehension and elements concerning grammar and vocabulary use; however, the intermediate learners do not achieve complicated thought and topics that are achieved by advanced learners.

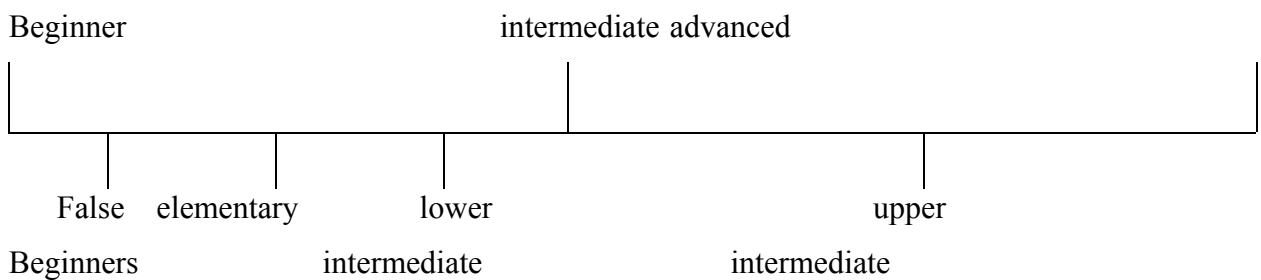
### **I.5.3. Advanced Level**

In this level, learners are able to communicate easily and fluently in the target language in real context. Harmer also considers that they can also use it outside with its native speakers:

*“...Advanced students are those whose level of English is competent, allowing them to read unsimplified fact and fiction and communicate fluently with native speakers.”*

Their testing is based on the learners’ communicative language ability involving its components like: linguistic competence, strategic competence...etc.

Figure 1



*Terms for student levels: Jeremy Harmer*

Knowing these levels is to determine the level of testing foreign language learning to be held at the end of any language program, so that, each level will be tested differently. Each one has its own tasks that differ in terms of level of complication for the use of grammar, vocabulary, pronunciation, fluency, organisation and content.



## **Conclusion**

Testing English language is a major part in teaching English language since it is a decision-making component to diagnose the learners' weaknesses, to place them in the appropriate position, to check their proficiency in the target language, or to check their achievement of the presented program. This test is held under certain conditions, and must have certain characteristics to gain reliable and valid results. And to guarantee the success of this test, the tester must consider the construction of this test and its items as well as the levels of the testees, and that will decide the test level itself. Not forgetting that this test must be suitable for the current communicative approach that is the essential purpose of learning any English language, and it must take the two forms (written and oral)