

Types of Language Tests

The needs of assessing the outcome of learning have led to the development and elaboration of different test formats. Testing language has traditionally taken the form of testing knowledge about language, usually the testing of knowledge of vocabulary and grammar. Stern (1983, p. 340) notes that '*if the ultimate objective of language teaching is effective language learning, then our main concern must be the learning outcome*'. In the same line of thought, Wigglesworth (2008, p. 111) further adds that "*In the assessment of languages, tasks are designed to measure learners' productive language skills through performances which allow candidates to demonstrate the kinds of language skills that may be required in a real world context.*" This is because a "*specific purpose language test is one in which test content and methods are derived from an analysis of a specific purposes target language use situation, so that test tasks and content are authentically representative of tasks in the target situation*" (Douglas, 2000, p. 19).

Thus, the issue of authenticity is central to the assessment of language for specific functions. This is another way of saying that testing is a socially situated activity although the social aspects have been relatively under-explored (Wigglesworth, 2008). Yet, language tests differ with respect to how they are designed, and what they are for, in other words, in respect to test method and test purpose. In terms of method, we can broadly distinguish traditional *paper-and-pencil language tests* from *performance tests*.

Paper-and-pencil language tests are typically used for the assessment either of separate components of language knowledge (grammar, vocabulary etc.), or of a receptive understanding (listening and reading comprehension). In performance-based tests, the language skills are assessed in an act of communication. Performance tests¹ are most commonly tests of speaking and writing, for instance, to ask a language learner to introduce himself or herself formally or informally and to write a composition, a paragraph or an essay,

¹ A performance test is "*a test in which the ability of candidates to perform particular tasks, usually associated with job or study requirements, is assessed*" (Davies et al., 1999, p. 144).

on the way he or she spent her summer holidays. These examples are elicited in the context of simulations of real-world tasks in realistic contexts. In terms of purpose, several types of language tests have been devised to measure the learning outcomes accordingly. However, each test has its specific purpose, properties and criterion to be measured². The test types that will be dealt with in this part have been laid-out not in terms of importance, they are all of equal importance, but on the basis of alphabetical order. Yet, dictation, the traditional testing device which focuses much more on discrete language items, will have its fair of attention in terms of its pro's and con's.

1. Achievement Test

An achievement test, also referred to as attainment or summative test, are devised to measure how much of a language someone has learned with reference to a particular course of study or programme of instruction, e.g. end-of-year tests designed to show mastery of a language. An achievement test might be a listening comprehension test based on a particular set of situational dialogues in a textbook. The test has a two-fold objective:

- 1) To help the teachers judge the success of their teaching.
- 2) To identify the weaknesses of their learners.

In more practical and pedagogical terms, Brown (1994, p. 259) defines an achievement test as '**tests that are limited to particular material covered in a curriculum within a particular time frame**'. In other words, they are designed primarily to measure individual progress rather than as a means of motivating or reinforcing language. Ideally, achievement tests are rarely constructed by classroom teacher for a particular class.

² Richards et al. (1985) define a criterion-referenced test (CRT) as: *a test which measures a student's performance according to a particular standard or criterion which has been agreed upon. The student must reach this level of performance to pass the test, and a student's score is therefore interpreted with reference to the criterion score, rather than to the scores of the students.* That definition is very different from their definition for a norm-referenced test (NRT) which they say is: *a test which is designed to measure how the performance of a particular student or group of students compares with the performance of another student or group of students whose scores are given as the norm. A student's score is therefore interpreted with reference to the scores of other students or group of students, rather than to an agreed criterion score.*

2. Cloze Test

A cloze test, also alternately referred to as cloze procedure, consists of a set of techniques for measuring, for example, reading comprehension. In a cloze test words are removed from a reading passage at regular intervals, leaving blanks. For example every fifth word may be removed. The reader must then read the passage and try to guess the missing words. For example, a cloze passage looks like this:

A passage used in cloze test is a of written material in words have been regularly..... . The learners must then to reconstruct the passage filling the missing

(Adapted from Richards et al., 1989, p. 41)

Here, the test-taker or the reader has to guess the following missing words: *a, passage, which, removed, try, by and words.*

The cloze test can also be used to judge the difficulty of reading materials. If the cloze procedure is being used for language testing, the test-taker is given a score according to how well the words guessed match the original words, or whether or not they make sense. Two types of scoring procedure are used:

- 1) The reader must guess the exact word which was used in the original (as in the example) above. This is called *exact word method*.
- 2) The reader can guess any word that is appropriate or acceptable in the context. This is called the *acceptable word method*.

Another illustrative example of close test looks something like the following: '*A week has seven*'. The only word which will fit in this blank is 'days'. But sometimes one can choose between two or more words, as in: '*We write with a.....*'. In this blank one can write 'pen' or 'pencil' or even 'chalk', 'computer' or 'typewriter' .

However, two substantial criticisms have been made to the cloze-test types (Broughton et al., 1980). The first of these criticisms is that such tests rarely afford the person being tested any opportunity to produce language spontaneously. The second is that they are fundamentally trying to test that knowledge of the language system that underlies any actual instance of its use –linguistic competence in the Chomskyan sense- they are not concerned with the ability to

master the language system for particular purposes with particular people in particular situations.

3. Diagnostic Test

As its name denotes, a diagnostic test is primarily designed to diagnose some particular linguistic aspects. Diagnostic tests in pronunciation, for example, might have the purpose of determining which particular phonological features of the English language are more likely to pose problems and difficulties for a group of learners. One of the well-known diagnostic tests in English is Prator's (1972) *Diagnostic Passage*. It consists of a short written passage that the learner reads orally; the teacher then examines a tape recording of that reading against a very detailed checklist of pronunciation errors. Basically, diagnostic language tests have a three-fold objective:

1. To provide learners with a way to start learning with their own personal learning programme or what would be called in the literature of testing *learning paths*.
2. To provide learners with a way to test their knowledge of a language.
3. To provide learners with better information about their strengths and weaknesses.

Ideally, diagnostic tests are designed to assess students' linguistic knowledge (knowledge of and about the language) and language skills (listening, speaking, reading and writing) before a course is begun. However, the term *formative* is sometimes used to designate a diagnostic test. One of the main advantages of a diagnostic test is that it offers useful pedagogical solutions for mixed-ability classes. In this very specific context, Broughton et al. (1980) contend that:

There will certainly be a large block in the middle of the ability range who can be separated off as a group for some parts of the lesson, or for some lessons, and will form a more homogenous teaching group. If this strategy is adopted, the poor ones and the better ones must receive their due time and attention.

(Broughton et al. 1980, p. 189)

4. Discrete-Point Test

The discrete-point test, also called discrete-item test, is a language test which measures knowledge of individual language items, such as a grammar test which has different sections on tenses, adverbs and prepositions. Discrete-point tests are based on the theory that language consists of different parts such as speech sounds, grammar and vocabulary, and different skills such as listening, speaking, reading and writing, and these are made up of elements that can be

tested separately. Test consisting of multiple-choice questions are usually regarded as discrete-point tests. Discrete-point tests are all too often contrasted with what are called integrative tests. An integrative test is one which requires a learner to use several skills at the same time. An essay-writing is an integrative test because it leans heavily on the knowledge of grammar, vocabulary, and rules of discourse; a dictation is also an integrative test as it requires knowledge of grammar, vocabulary and listening comprehension skills.

In this vein, Harmer notes the following distinction between discrete-point testing and integrative testing, “*Whereas discrete point-testing only tests one thing at a time such as asking students to choose the correct tense of a verb, integrative test items expect students to use a variety of language at any one given time – as they will have to do when writing a composition or doing a conversational oral test*” (Harmer, 2001, p. 323). In the same line of thought and Broughton *et al.*, more than some thirty years ago, noted that “*Since language is seen as a number of systems, there will be items to test knowledge of both the production and reception of the sound segment system, of the stress system, the intonation system, and morphemic system, the grammatical system, the lexical system and so on*” (Broughton et al., 1980, pp. 149-150).

5. Language Aptitude Test

Before one ventures into defining what a language aptitude test is, it would be wiser to start first by defining what a language aptitude is. Language aptitude, as a hybrid concept part linguistic and part psychological, refers to the genuine ability one is endowed with to learn a language. It is thought to be a combination of several abilities:

- *Phonological ability*, i.e. the ability to detect phonetic differences (e.g. of stress, intonation, vowel quality) in a new language.
- *Syntactic ability*, i.e., the ability to recognize the different grammatical functions of words in sentences.
- *Psychological ability*, i.e. rote-learning abilities and the ability to make inferences and inductive learning.

Additionally, Crystal (1989, p. 371) suggests other variables conducive to successful language learning such as ‘***empathy and adaptability, assertiveness and independence with good drive and powers of application***’. A high language-aptitude person can learn more quickly and

easily than a low language-aptitude individual. The evidence in such assertion is axiomatic in a language aptitude test.

A language aptitude test tends to measure a learner aptitude for language learning, be it second or foreign, i.e. students performance in a language. Thus, it is used to identify those learners who are most likely to succeed. Language aptitude tests usually consist of several different test items which measures such abilities as:

- *Sound-coding ability*, i.e. the ability to identify and remember new sounds in a new language.
- *Grammar-coding ability*, i.e. the ability to identify the grammatical functions of different parts of sentences.
- *Inductive-learning ability*, i.e. the ability to work out meanings without explanation in the new language.
- *Memorization*, i.e. the ability to remember and to recall words, patterns, rules in the new language.

Two well-known standardized language aptitude tests have been used in the United States, the Modern Language Aptitude Test (Carroll and Sapon, 1958) and the Primsleur Language Aptitude Battery (Primsleur, 1966). Both of these are English tests and require students to perform such tasks as learning numbers, listening, detecting spelling clues and grammatical patterns and memorizing (Brown, 1994).

6. Placement Test

A placement test, as its name implies, is originally designed to place learners at an appropriate level in a programme or course. The term “placement test” as Richards et al. (1989) note does not refer to what a test contains or how it is constructed, but to the purpose for which it used. Various types of testing procedures such as dictation, interview or a grammar test (discrete or integrative) can be used for placement purposes. The English Placement test (EPT), which is a well-known test in America, is an illustrative example of this test-type. The EPT is designed to assess the level of reading and writing skills of entering undergraduate students so that they can be placed in appropriate courses. Those undergraduate students who do not demonstrate college or university-level skills will be directed to remedial courses or programmes to help them attain these skills.

7. Proficiency Test

A proficiency test is devised to measure how much of a language someone has learned. It is not linked to any particular course of instruction, but measures the learner's general level of language mastery. Most English language proficiency tests base their testing items on high frequency-count vocabulary and general basic grammar. Some proficiency tests have been standardized for worldwide use, such as the well-known American tests, the TOEFL, and the English Language Proficiency Test (ELPT)³ which are used to measure the English language proficiency of foreign students intending further study at English-speaking institutions, namely the USA.

However, the Cambridge Certificate of Proficiency in English or CPE, as it is generally referred to, is the most advanced remains the only British top-value and high-prestige standardized⁴ language test. It is the most advanced general English exam provided by the University of Cambridge. The Certificate is recognized by universities and employees throughout the world. The English level of those who pass the CPE is supposed to similar to that of a fairly educated native speaker of English. Clearly, as Valette posits, '*the aim of a proficiency test is to determine whether this language ability corresponds to specific language requirements*' (Valette, 1977, p. 6)

Actually, there are four other types of Cambridge proficiency tests, the Cambridge Key English Test (KET), the Cambridge Preliminary English Test (PET), The Cambridge First Certificate of English (FCE) and the Cambridge Certificate in Advanced English (CAE). The material contained in proficiency tests can be used for teaching as well as for testing. In essence, a proficiency test measures what the student has learned in relation to a specific purpose, e.g. does the student know enough English to follow a course offered in English?

³ The English Language Proficiency Test (ELPT) was the name of a test last administered in January 2005. It was a one-hour multiple choice question given on English language proficiency. A student whose native language was not English could have chosen to take this test instead of or in addition to the TOEFL for college or university entrance depending upon the requirements of the schools in which the student was planning to apply. Until 1994, the tests were known as Achievement Tests. The ELPT assessed both the understanding of spoken and written standard American English and the ability to function in a classroom where English is spoken.

⁴ A *standardized test* is an exam which has been developed from tryouts and experimentation to ensure that it is reliable and valid. It is also a test for which norms have been established and it provides uniform procedures for administering (time limits, response format, and number of questions) and for scoring the test. "*Standardized tests are often used by school systems for high-stakes decision making*" (Menken, 2008, p. 402).

8. Progress Test

A progress test is an achievement-like test. It is closely related to a particular set of teaching materials or a particular course of instruction. Progress tests are usually administered at the end of a unit, a course, or term. A progress test may be viewed as similar to an achievement test but much narrower and much more specific in scope (Richards et al., 1989). They help examiners in general and language teachers in particular to assess the degree of success of their programmes and teaching and therefore to identify their shortcomings and weaknesses respectively. Progress tests can also be diagnostic to some degree, in the sense that they help identify areas of difficulties encountered by learners in general.

9. TOEFL

The Test of English as a Foreign Language, or TOEFL for short, is a large-scale language assessment. It is, “***arguably the most well-known and widely used large-scale language assessment in the world***” (Kunnan, 2008, p. 140). It was first developed in 1963 in the United States to help in the assessment of the language competence of non-native speakers. As a test type, it is a standardized test of English proficiency administered by the Educational Testing Service, Princeton. It is widely used to measure the English-language proficiency of foreign students wishing to enter American colleges and universities. According to Taylor and Angelis (cited in Kunnan, 2008) the first TOEFL was administered in 1964 at 57 test centres to 920 test candidates. Recently, the TOEFL has widely been recognized as a model test and have-take-test for our students, graduate and postgraduate, as well as our teachers and researchers in universities and higher education institutions wishing to read for higher degrees and develop further their research potential in North American universities⁵. Kunnan (2008, p. 141) notes that, “***Over the years, the TOEFL became mandatory for non-American and non-Canadian native speakers of English applicants to undergraduate and graduate programs in U.S. and Canadian English-medium universities***”.

One of the most important realizations in the TOEFL enterprise was the launching of a more innovative test, the iBTOEFL, internet-based TOEFL, in 2005. This iBTOEFL is

⁵ The International English Language Testing System, IELTS, is designed to assess the language ability of candidates who wish to study or work in countries where English is the language of communication. IELTS is required for admission to British universities and colleges. It is also recognized by universities and employers in Australia, Canada, and the USA. IELTS is jointly managed by the University of Cambridge, British Council and IDP Education.

regarded as a significant development over the previous TOEFL forms and the TOEFL CBT, Computer-Based Test, launched in 1996. The novel features of the iBT[®]TOEFL are a speaking section consisting of independent and integrated skills tasks, a listening section with longer lectures and conversations with note-taking, a reading section made up of questions that ask test-takers to categorize information and fill in a chart or complete a summary and a writing section that has both an independent and integrated task.