

Language Proficiency Tests in the Iranian Context: Do They Represent Communicative Language Testing Model?

Seyyed Ayatollah Razmjoo
Shiraz University

Razmjoo, S. A. (2011). Language proficiency tests in the Iranian context: Do they represent communicative language testing model? *Journal of Pan-Pacific Association of Applied Linguistics*, 15(2), 85-96.

The Communicative Ability in language testing originates from a theory of language as communication proposed by Hymes (1972) and known as "communicative competence". The literature on language testing suggests that the practicality of communicative language testing (CLT) varies depending on how the instructors and teachers conceptualize CLT. As far as the researcher knows, no study has been done on the type of proficiency model behind language tests in the Iranian context. Therefore, the objective of this study is, firstly, to explore the language proficiency model based on which the tests are designed and constructed in Iranian high school and institutes. Secondly, the study aims at developing language proficiency model for the context of Iran as a basis so that teachers construct and design tests. To fulfill the objectives, two final tests of Iranian high school and institutes were analyzed. The analysis of the data indicated that high school and institute tests represent a short version of Lado's model (1961) which is not in line with new trends in language testing and aims at discovering whether the correct habits have been formed.

Key Words: proficiency model, communicative testing, CLA, CLT, EFL institutes

1 Introduction and Background of the Study

Testing and teaching are like two sides of a coin; that is, testing independent of teaching is unimaginable. (Hughes, 2003). "Both testing and teaching are so closely interrelated that it is virtually impossible to work in either field without being constantly concerned with the other" (Heaton, 1990, p. 5). Testing must be seen as a method of providing information that may be used for teaching and other purposes. However, the reality, as Davies (1990) puts it, is that "testing is always used in teaching, in the sense that much teaching is related to the testing which is demanded of the students" (p. 24).

The fact that testing affects the four most important factors of the curriculum -- teacher, learner, content and context -- is commonplace in the

Seyyed Ayatollah Razmjoo

educational and applied linguistics literature (Farhady *et al.*, 1998; Hughes, 2003; Alderson & Wall 1993). A great deal of research has been conducted about the effect of examinations on what takes place in the classroom. Pearson (1988 as cited in Alderson & Wall, 1993) states that “it is generally accepted that public examinations influence the attitudes, behavior and motivation of teachers, learners and parents” (p. 115).

Communicative language testing is intended to provide the tester with information about the testee’s ability to perform in the target language in certain context-specific tasks. It has to be recognized that given the constraints of time and practicality, only a small sample of the testee’s language can be collected, and that however realistic the tasks may be intended to be, the testee’s performance will inevitably reflect the fact that s/he was performing under test conditions. The following section deals with the presenting language testing models in general and the CLT in particular.

1.1 Differences between communicative language testing and other forms of testing

We will address this by first briefly identifying other testing methods in the ‘eras’ preceding the emergence of communicative language testing, looking at what they were intended to measure and their theoretical basis. We will then turn to communicative testing and examine two of the communicative models on which it is based, and the characteristics which set it apart from other testing techniques.

Spolsky (1975) identified three periods of language testing: the pre-scientific, the psychometric-structuralist and the psycholinguistic-sociolinguistic. Although he has since (Spolsky 1995) offered an alternative view, we will use his original phases in this section.

Spolsky first identifies the pre-scientific era. He recognizes it as dating back to the Chinese civil service exams two thousand years ago, but believes it took its present form from the 18th century Cambridge Tripos (Spolsky 1995). It was characterized by “the use of essays, open-ended examinations, or oral examining, with the results determined intuitively by an authorized and authoritarian examiner (Spolsky, 1995, p.353). As the name suggests, “testing in the pre-scientific era did not rely on linguistic theory and reliability was considered less important than the production of a test” (Spolsky, 1995, p.356).

After the pre-scientific era came the psychometric-structuralist era. The name was intended to reflect the joint contribution of the structural linguist, who identified elements of language s/he wanted to test and the psychometrist, who produced objective and reliable methods of testing the candidates’ control of those elements.

One of the first people to claim the need for input from these two sources was Lado, who was also responsible for the discrete point approach.

Language Proficiency Tests in the Iranian Context

The discrete point approach broke language down, using structural contrastive analysis, into small testable segments. Each test item was intended to give information about the candidate's ability to handle that particular point of language including the four language skills and the three components.

The main advantage of this was that it provided easily-quantifiable data. However, it also had numerous drawbacks, perhaps the greatest of which was pointed out by Morrow (1981, p.11), "An atomistic approach to test design depends utterly on the assumption that knowledge of the elements of a language is equivalent to knowledge of the language." As he says, knowledge of discrete elements is worthless unless the user can synthesize those elements according to the linguistic demands of the situation, or, in the words of Oller (1979, p.212, as cited in Weir 1990), "the whole is greater than the sum of its parts..."

By the 1970s discrete point testing was no longer felt to provide a sufficient measure of language ability and testing moved into the psycholinguistic-sociolinguistic era, with the advent of global integrative testing. Oller (1979, as cited in Weir 1990) argued that global integrative testing, such as cloze tests, which required candidates to insert suitable words into gaps in a text and dictation, provided a closer measure of the ability to combine language skills in the way they are used for actual language use than discrete point testing.

However, Oller's unitary trait hypothesis, which supposed that language proficiency consists of a single unitary ability (Bachman, 1990) and upon which cloze and dictation were based, has since been disconfirmed (Bachman, 1990) and the techniques have been heavily criticized. Alderson (1978, as cited in Weir 1990) pointed out that results of cloze tests were affected according to the number of deleted items and where the deletions began.

Morrow (1979, as cited in Weir 1990) states that neither technique allows for spontaneous production by the candidate, relying instead on the examiner for the language input. He also criticized the techniques on the grounds that they tested competence rather than performance, in other words, they tested knowledge of how the language worked rather than an ability to use it.

The fact that discrete point and integrative testing only provided a measure of the candidate's competence rather than measuring the candidate's performance brought about the need for communicative language testing (Weir 1990). Before we look at the features which distinguish this form of testing, we will outline the models of communicative competence on which it is based.

According to Spolsky (1989), Language tests involve measuring a subject's knowledge of and proficiency in the use of a language. A theory of communicative competence is a theory of the nature of such knowledge and

Seyyed Ayatollah Razmjoo

proficiency. One cannot develop sound language tests without a method of defining what it means to know a language, for until you have decided what you are measuring, you cannot claim to have measured it.

Several attempts have been made to define what it means to know a language, but we only propose to discuss two of the more influential models. The work of Canale and Swain began in an attempt to “determine the feasibility and practicality of measuring what we shall call the ‘communicative competence’ of students” (Canale and Swain, 1980, p.1). Canale and Swain proposed a set of three competences which combine to produce communicative competence. The first, grammatical competence, included “knowledge of lexical items and rules of morphology, syntax, sentence grammar semantics and phonology” (Canale and Swain 1980, p.29). The second was sociolinguistic competence. This was made up of “sociocultural rules of use and rules of discourse” (Canale and Swain, 1980, p. 29). The third competence they proposed was strategic competence, which related to “verbal and non-verbal communicative strategies that may be called into action to compensate for breakdowns in communication due to performance variables or to insufficient competence” (Canale and Swain 1980, p. 29). In 1983, Canale updated this model by subdividing sociolinguistic competence, which still relates to sociocultural rules, but he introduced a further competence, that of discourse. Discourse competence concerns mastery of cohesion and coherence in different genres.

The main implication this model had for communicative language testing was that since there was a theoretical distinction between competence and performance, the learner had to be tested not only on his/her knowledge of language, but also on his/her ability to put it to use in a communicative situation (Canale and Swain, 1980).

Bachman’s framework (1990) was an extension of earlier models “in that it attempts to characterize the processes by which the various components interact with each other and with the context in which language use occurs” (Bachman, 1990, p.81). The framework included three components: language competence, strategic competence and psychophysiological mechanisms (Bachman, 1990). Language competence comprises two further competences, organizational competence and pragmatic competence, each of which he further breaks down, with organizational competence covering grammatical and textual competence, and pragmatic competence covering illocutionary and sociolinguistic competence. Bachman defined language competence as “a set of components that are utilized in communication via language” (Bachman 1990, p.84).

Strategic competence consists of three components: assessment, planning and execution. It is the mental capacity to implement language competence appropriately in the situation which communication takes place, and involves sociocultural and real world knowledge. Psychophysiological

Language Proficiency Tests in the Iranian Context

mechanisms refer to the neurological and psychological processes involved in producing and comprehending language.

One notable advance on the Canale and Swain model is that Bachman acknowledges that test design and scoring might have a significant effect on the testee's performance as a result of strategic competence. Certain tasks lend themselves to use of strategic competence to compensate for a lack of competence in other areas, while tests which are assessed according to the "practical effect of the language performance" (Bachman, 1990, p.105) may be affected by the strategic competence factor.

Canale and Swain and Bachman's are two of the more influential models of language competence, and, along with several others, they provide a useful framework for designing communicative language tests (Weir, 1990). We will now go on to describe some of the features of communicative language tests which set them apart from other forms of testing.

1.2 CLT features

Communicative language tests should have high content validity. If they are to be used to make judgments about how an individual can function in a normal situation outside the test, the test has to be as accurate a reflection of that situation as possible. This means that the sample of language collected and the tasks the candidate is called upon to perform should be as representative as possible of the language and skills needed to function in the real life context. Tests, therefore, need to be context-specific. If, for example, the objective is to test candidates to determine whether their second language ability is adequate to undertake a course at a higher education establishment, conducted in that second language, the tasks included in the test should be a fair reflection of the type of tasks the candidate will be required to perform as part of the course itself. As Weir (1990) points out, inauthentic tasks may interfere with the measurement of the construct which we seek. "Tests of communicative language ability should be as direct as possible (attempt to reflect the 'real life' situation) and the tasks candidates have to perform should involve realistic discourse processing" (Weir 1990, p. 12). He advocates the use of genuine texts and that care be taken with regard to task length and processing in real time.

Face validity is also related to authenticity of tasks. Although not universally agreed upon, many testers believe it is easier to gain acceptance for a test which appears to test real life skills than those which use formats such as cloze, which are not seen outside the test itself. Employing tasks which the testees might recognize also makes it easier to explain and justify the test to them. According to Morrow (1981, p.18), "Reliability, while clearly important, will be subordinate to face validity."

Tests of communicative spoken ability should have certain characteristics. They should reflect normal spoken discourse and give the candidate chances

Seyyed Ayatollah Razmjoo

to initiate. There should also be an element of unpredictability. As Morrow (1981) points out, the processing of unpredictable data in real time is a vital aspect of using language. The final aspect of communicative language testing we would like to address is that of assessment. Communicative tests should be assessed qualitatively rather than quantitatively (Morrow, 1981). The behaviorist view was that learning took place through habit formation. Following from this, tests such as Lado's aimed to discover whether the correct habits had been formed. If they had, they were rewarded, but if they hadn't, they weren't. Passing the test meant obtaining a certain number of correct responses. However, Morrow (1981) argues that answers to tests are more than simply right or wrong and that they should be assessed on the basis of how far toward an approximation of the native speaker's system they have moved. Tests should reveal the quality of the testee's language performance. Assessment which relates test performance to external criteria is called criterion referencing. It is an area of some contention and it is the starting point for the next part of this paper.

1.3 Studies done on language tests in Iran

During the past twenty years or so, different studies have been conducted as well as reported with respect to the shortcomings of the University Entrance Exams in Iran. For example, investigating the exams administered from 1983 to 1985, Farhady (1985) found little correspondence between the manner materials are taught to the students and the manner the students are tested on them. In a similar attempt, Yarmohammadi (1986) came to this conclusion that in such exams, the role of communication is, to a great extent, ignored. Moreover, among language skills only one skill (reading, namely) and among language components only grammar and vocabulary are tested. Jafarpur (1996; see also Jarfarpu, 1986), ifor instance, in a study critically reviewed the TEFL MAEE of 1996 and concluded that "the validity of that exam is under question due to the problems related to the items in terms of testing principles" (p.20). Razmjoo (2006) stated that the University Entrance Examination held in Iran during the past two decades are frequently criticized as being invalid, and non-standardized with lots of problems in terms of principles of testing in general and test construction in particular. Yet some of the basic principles of language testing are not observed in the process of constructing the exam.

1.4 Objectives of the study

This study investigates the extent to which high school and EFL institute tests represent the CLT model.

The first objective of the study deals with the analysis of high school tests from the perspective of repressing the CLT model.

Language Proficiency Tests in the Iranian Context

The second objective is to find out to what extent EFL institute tests represent the CLT model.

Third, the current study aims at comparing the amount of representing the CLT model in the tests of each domain.

The fourth and in effect the chief objective of the study is to develop a testing model suitable for the Iranian context based on the results of the study.

1.5 Research questions

Regarding the purpose of the study, the following research questions are posed:

1. To what extent, do the high school tests represent the CLT model?
2. To what extent, do the EFL institute tests represent the CLT model?
3. Are there any significant differences among high school and institute tests in terms of representing the CLT model?

2 Method

In this section, the materials of the study, the instrument and the data collection procedure are presented.

2.1 Materials

The materials include the final test of grade three of high school which is administered nationwide as well as a final test prepared by EFL institutes. Each material is discussed below in detail.

2.1.1 The content and organization of EFL institute test

The test includes 50 multiple-choice items and the language learners are supposed to answer them in at most one hour. It consists of the following sections:

- I. Listening Comprehension
- II. Vocabulary
- III. Grammar
- IV. Reading Comprehension
- V. Cloze Passage

2.1.2 The content and organization of the high school test

The test includes a variety of item forms such as fill in the blanks and missing letters, multiple-choice items, short answers and matching. The

Seyyed Ayatollah Razmjoo

students are supposed to answer the items at most two hours. The test is made of the following sections:

- I. Dictation
- II. Vocabulary
- III. Word Formation
- IV. Grammar
- V. Language Functions
- VI. Pronunciation
- VII. Sentence Comprehension
- VIII. Cloze Task
- IX. Reading Comprehension

2.2 Instrument

2.2.1 A test analysis checklist

The final test of grade three of high school which is administered all over the country as well as a final test prepared by EFL institutes were evaluated based on the language proficiency model proposed by Bachman (1990) because he believes that the underlying theory of the model is the communicative competence proposed by Hymes (1972). (Appendix C).

To get the reliability of the test analysis checklist, it was tried to gain both intra-coder and inter-coder reliability for the scheme. For intra-coder reliability, the researcher analyzed and recorded the same test with a two-week time interval and the correlation between the analyses and the recordings was computed. To gain inter-coder reliability, an independent encoder who was sufficiently trained, skilled and familiar with the task accompanied the researcher. The test was analyzed by the coders independently at the same time and the correlation of marking the items done in the test represented the inter-coder reliability and were shown to be .98 and .92 respectively.

2.3 Data collection and analysis

The two tests of public and private institutes were analyzed based on a language proficiency model which is in line with the CLT principles and theoretical framework. So, the tests were surveyed based on the model to see how much they include the components of communicative competence of language.

Language Proficiency Tests in the Iranian Context

3 Results and Discussion

In the following sections, the results of the test analysis in the two domains is presented and discussed.

Table 1. Analysis of the Public School English Test

Category	Components	Type of Items	Sub-category Items	Weight
Grammatical Comp.	Graphological	Dictation	16 incomplete words	4 points
Grammatical Comp.	Vocabulary	Vocabulary	Fill in the blank with the appropriate words	4 points
		Vocabulary	6 incomplete sentences	3 points
Grammatical Comp.	Morphology	Word Formation	6 incomplete sentences	3 points
		Structure	6 MC items	3 points
Grammatical Comp.	Syntax	Structure	2 Word Order	2 pints
		Structure	2 Incomplete sentences	2 points
		Structure	2 picture completion	2 points
Grammatical Comp.	Vocabulary + Syntax	Language Function	8 matching items	4 points
		Pronunciation: Sound	2 MC items	1 point
Grammatical Comp.	Phonology	Pronunciation: Stress	2 MC items	1 point
		Sentence Comprehension	4 MC items	4 points
Textual Comp.	Cohesion	Reading Comprehension	3 open-ended items	3 points
		Reading Comprehension	2 T/F items	1 point
Textual Comp.	Cohesion	Cloze task	6 MC items	3 points

The content analysis of the English test of the public school shows that that the different language components are tested in discrete items, an evidence for partial representation of the skills/component model (Lado, 1961) as the underlying theory of language proficiency in such tests. Except for the cloze passage, the grammar, vocabulary, language function, pronunciation and reading comprehension items are tested within very limited contexts of use.

Table 2. Analysis of the Private Institute English Test

Category	Components	Type of Items	Sub-category Items	Weight
Grammatical Comp.	Graphological	Listening	6 MC items	6 points

Seyyed Ayatollah Razmjoo

Grammatical Comp.	Vocabulary	Vocabulary	17 MC items	17 points
Grammatical Comp.	Syntax	Structure	18 MC items	19 points
Textual Comp.	Cohesion	Reading Com	4 MC items	4 points
Textual Comp.	Cohesion	Cloze task	5 MC items	5 points

Although the private institute English textbooks aim to develop “all language skills” in the learners, speaking and writing skills are not taken into account in the tests, probably due to the difficulties in administration and scoring. Thus, the private institute English test seems to concentrate mostly on the learners’ reading comprehension skills and their knowledge of grammar and vocabulary as the requirement for obtaining passing scores on the tests. So, it can be concluded that the exams of the two domains do not represent the communicative language ability model proposed by Bachman (1990). This is in sharp contrast with the new trends in language testing (Bachman 1990; McNamara, 2000).

Bachman (1990) stated that from both theoretical and practical points of view, language tests should be developed in a way that they:

... reflect current views of language and language use, in that they are capable of measuring a wide range of abilities generally associated with ‘communicative competence’, or ‘communicative language ability’, and include tasks that themselves embody the essential features of communicative language use. (p.297).

All in all, the public and private domain English tests analyzed in this study are far from the present approaches to language tests design (e.g., performance based assessment, communicative language testing). Neither do the tests include all the points (the oral language skills and the pronunciation components) of the Lado's Model (1961). Even if they aim to measure the applicants’ “knowledge of grammar and vocabulary and their reading ability”, they can be developed in context-based items so as to meet the basic requirements for moving beyond discrete-point testing. As such, following model can be proposed so that the Iranian test constructors and designers stick to it while preparing items for the two domains:

4 Conclusions

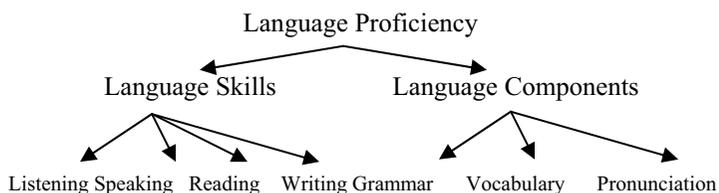
The analysis of the two tests reveals the fact that the elements of communicative language testing are not represented and practiced in the tests constructed by the Iranian testers. Moreover, tests of the two domains do not differ to a great extent; that is, neither of the tests taps all the CLT components. As such, they have the following shortcomings:

Language Proficiency Tests in the Iranian Context

1. The two tests are not as accurate a reflection of the real situation as possible. They, therefore, are not context-specific.
2. The two tests are quantitative rather than qualitative. They follow the behaviorist view that learning took place through habit formation. As such they do not reveal the quality of the testee's language performance.

Obviously, more definite conclusions must await further studies with other tests of other levels and contexts. Till then, these results must be considered as tentative. As such, and for the time being, the following model can be proposed so that the Iranian test constructors and designers stick to it while preparing items for the two domains:

Figure 1. A proposed language proficiency model for the Iranian context



References

- Alderson, J. C. (1981) Report of the discussion on communicative language testing. In J. C. Alderson, & A. Hughes (Eds.). *Issues in Language Testing. ELT Documents 111*. London: The British Council.
- Alderson, J. C., & Wall, D. (1993). Does washback exist? *Applied Linguistics*, 14(2), 115-129.
- Bachman, L. F. (1990) *Fundamental considerations in language testing*. Oxford: Oxford University Press.
- _____. (1991) What does language testing have to offer? *TESOL Quarterly*, 25(4).
- Brindley, G. (1991) Defining language ability: The criteria for criteria. In Anivan, S (Ed.), *Current development in language testing*. Singapore: SEAMEO Regional Language Centre.
- Canale, M. (1983). From communicative competence to communicative language pedagogy. In Richard, J. C. & Schmidt, R. W. (Eds.), *Language and communication*. London: Longman. 2-27.
- Davies, A. (1990). *Principles of language testing*. Oxford: Blackwell.
- Farhady, H. (1985). A survey of English tests of University Entrance Examination. *Roshad Language Teaching Journal*, 31, 15-17.
- Farhady, H., Jafarpur, A., & Birjandi, P. (1998). *Testing language skills: From theory to practice*. Tehran: SAMT.

Seyyed Ayatollah Razmjoo

- Heaton, J. B. (1990). *Writing English language tests*. London: Longman.
- Hughes, A. (2003). *Testing for language teachers*. (2nd ed.) Cambridge: Cambridge University Press.
- Hymes, D. (1972). On communicative competence. In J. B. Pride, & J. Holmes (Eds.), *Sociolinguistics*, Harmondsworth: Penguin. 269-293.
- Jafarpur, A. (1997). A survey of English Entrance Exam of Graduate School. *Journal of Isfahan University*, 8(1), 15-22.
- _____. (1986). A survey of one of the final tests of English of the Ministry of Education. *Journal of Social Sciences and Humanities, Shiraz University*, 2(1), 92-105.
- Morrow, K. (1981). Communicative language testing: Evolution or revolution? In J. C. Alderson and A. Hughes (Eds.), *Issues in Language Testing*. ELT Documents 111. London: The British Council.
- Razmjoo, S. A. (2006). A Content Analysis of University Entrance Examination for English Majors in 1382. *Journal of Social Sciences and Humanities, Shiraz University*, 46(1), 67-75.
- Skehan, P. (1991) Progress in Language Testing: The 1990s. In J. C. Alderson and B. North (Eds.). *Language Testing in the 1990s*. London: Macmillan.
- Spolsky, B. (1989) Communicative Competence, Language Proficiency and Beyond. *Applied Linguistics*, 10(2), 138-155.
- _____. (1990) *Measured Words: The Development of Objective Language Teaching*. Oxford: Oxford University Press.
- Weir, C. J. (1990). *Communicative language testing*. London: Prentice Hall.
- Yarmohammadi, L. (1986). A survey of English tests of University Entrance Examination in 1365. *Journal of Social Sciences and Humanities, Shiraz University*, 2(1), 80-91.

Seyyed Ayatollah Razmjoo
Department of Foreign Languages and Linguistics
College of Literature and Humanities
Shiraz University, Shiraz, Iran, 7194685115
Tel: 00987116288721
E-mail: arazmjoo@rose.shirazu.ac.ir

Received: August 17, 2011
Revised: November 23, 2011
Accepted: December 7, 2011