

# APPROACHING COMMUNICATIVE LANGUAGE TESTING

Stephen Burrow

## Abstract

This paper is concerned with communicative language testing (CLT), which has long been a goal of the language testing world. CLT began in the 1980s as a complement to communicative language teaching, and now it is fair to say that most language tests are made with the concept of CLT in mind. It is difficult to ascertain exactly what CLT is because there is disagreement as to what is good communication in a foreign language is, and what it means to know a language. Several features of CLT are defined, and the paper contains a discussion of the problems of making communicative language tests.

## What is Communicative Language Testing?

Language testing, to a large extent, follows language teaching in both theory and practice. The communicative approach to language teaching, which first appeared in the 1980s, and now task-based language teaching, which has become more widespread in recent years in Japan (Sato, 2010), both place emphasis on the end message and whether real communication has taken place. Communicative language testing (CLT) also has its roots in the 1980s, and although related to the communicative Approach to

language teaching, it is possible to identify several aspects of what it involves.

### **Features of Communicative Language Testing :**

#### **Authenticity :**

The first of these is authenticity, and for a test to be considered authentic it should involve language actually used in the real world. Therefore, a communicative language test should display a high level of what is known as content validity (Hughes, 2000). Content validity refers to whether we are really testing the real language we say we are testing. The ultimate aim of most language learners is to be able to communicate in a foreign language in any context, be it on holiday or in the workplace, and so a test should reflect those aims. In addition to authentic language, CLT should involve authentic situations or contexts, in which hopefully test takers will actually be able to communicate in outside of a test situation. A communicative language test, therefore, should allow us to make inferences as to whether test takers can use language in situations we think they will encounter outside of the test (Bachman and Palmer, 1996).

Bachman and Palmer also identified a concept of target language use (TLU) domain, a set of 'specific language tasks that a test taker is likely to encounter outside of the test itself'. Therefore, tests such as the TOEFL (Test of English as a Foreign Language), or IELTS (International English Language Testing System), both of which are used to assess whether a test taker is able to study in an English-speaking university, might have a TLU domain identified as 'the ability to use and understand

English the university level'. A set of language functions can be identified within this domain, such as note-taking, stating an opinion in a seminar class or participating in a discussion in a university seminar class. Therefore, if a test is authentic, it must address the situations or functions described in the TLU domain. The focus should be on the end product, rather than on how a test taker reached there. In addition, the language used in CLT should be authentic and within the TLU domain required, as opposed to simplified for the purposes of teaching. Although this might be difficult to achieve, ultimately test takers will be better prepared for communication in the real world.

### **Criterion-Referencing :**

CLT should also be criterion-referenced, and so relate the performance of students to various benchmarks or criteria according to what they can or cannot do. This is mainly true of speaking and writing tests, in which students are asked to produce language, as opposed to listening and reading tests which concentrate on receptive skills. An example of this is in the TOEFL test writing section, in which essays are graded according to various criteria on a scale of one to six. The advantage of this and other criterion-referenced testing is that you are actually testing what a test taker can do absolutely, not relating a test taker to another and so can clearly ascertain a test taker's performance and ability to communicate. It also can help teachers focus on what to teach their students, and if clearly defined, helps a test taker by providing specific learning goals.

## **Integration of Skills**

A third feature of CLT is that it involves more than one language skill at a time; in other words it is integrative. This reflects the real world because we usually use more than one language skill at a time. In a spoken test, in which a test taker is one-to-one with a tester, listening is important as well as spoken communication. A reading test can involve writing answers to questions, or a summary, and hence the ability to communicate what a test taker can understand becomes important. Integrative testing distinguishes CLT from discrete-point testing, whereby students are tested on one skill at a time (Hughes, 2000). Although discrete-point testing has some use, it is limited and does not really test communicative ability. A good example of the integration of skills in a test can be found in some sections of the internet-based TOEFL test (iBT).

## **Language Performance**

Another feature of CLT is that it tests knowledge of a language and the ability to use that knowledge. Chomsky's (1965) concept of competence versus performance is valid here, where competence is knowledge of a language and its grammar, and performance is how well a language is used. Spolsky (1985), in an influential book on language testing, stated that there were three main approaches to language knowledge and use. Firstly, the structural approach is how much can a student control the structures of a language, for example, its verb tenses. The functional approach, however, refers to how well a student can use

language in a given situation, such as communicating a past experience or buying a ticket at a travel agent. Thirdly, he identified the concept of general proficiency, which is basically what *knowing* a language means. CLT is not really concerned with the structural approach, but it is with the other two because two important constructs it tests are whether a test taker can function in different situations and general proficiency. Therefore construct validity, whether a test measure what it is claims it does (Hughes, 2000), becomes a vital issue.

### Language Production

Another issue concerned with reliability is that of actual language production. Fulcher (2000) argues that students should actually have to produce language for a test to be communicative. Communication involves production of language, and so for a test to communicative test takers must produce, either in writing or orally, language within the TLU domain.

### Problems of Communicative Language Testing

CLT presents a strong case, but in practice attempting to implement it can produce many problems. Weir (1990) and Hughes (2000) imply that there are three major criteria within which language tests have limitations . validity, reliability and efficiency :

## Validity

"Validity is concerned with whether a test measures what it is intended to measure. Reliability is concerned with the extent to which we can depend on the test results. Efficiency is concerned with matters of practicality and cost in test design and administration." (Weir 1990:1)

The first of these is validity, which is basically how much a test really tells us about a test taker, and their capabilities in a foreign language. Content validity, which relates to the authenticity of a test in terms of language and context or situation, becomes important here. It is virtually impossible to authentically replicate a real world situation in a language-testing situation. As Bachman (1990:356) says :

"The most complex and persistent problems in language testing are those presented by the consideration of the relationship between the language use required by tasks on language tests and that which is part of our everyday communicative use of language. That there must be a relationship between the language used on tests and that used in 'real life' cannot be denied, since if there is no such relationship, our language tests become mere shadows, sterile procedures that may tell us nothing about the very ability we wish to measure."

CLT must involve authentic and relevant language and contexts, but there is the problem of defining what the real world actually is. According to Fulcher (2000:489),

## APPROACHING COMMUNICATIVE LANGUAGE TESTING(Stephen Burrow)

"This 'real world' involves interaction, unpredictability, and integration of skills."

It is important to understand that in real life language is varied, and it is impossible to understand the simple fact that a test simple cannot cover all of it. The variety of language used in the real world is too great to be tested in one test, and there are countless variations of age, gender, region, country and others to consider Test takers may do well on a test designed to prepare them for work in the hotel industry and be perplexed when they encounter a guest who uses different language to what the have learnt and passed a test with. In this way, it may not be possible to depend on the results of a test.

### Reliability

Different tests should give the same student the same score for a test to be completely reliable. For a test to be reliable, it must produce scores which can be 'trusted' (Hughes, 2000). In practice, this is difficult to achieve. One major reason is that CLT for speaking and writing is mainly criterion-referenced, and so there is always some subjectivity on the part of the tester. There is no guarantee that different testers will agree on a score for the same student. This is because there is always a certain amount of subjectivity involved as to whether a test taker has achieved the benchmarks set in the test or not. The criteria could be written quite vaguely and leave quite a large area open to discretion.

Another problem of reliability is that some students are better at taking tests than others. Some are more confident than others in a spoken

test *situation*, and hence are more likely to perform better in that test but not necessarily outside of the test situation. One way in which more confident test takers perform better is that they simply respond faster than less confident ones. This is not necessarily a factor of language ability but more one of confidence. The latent or underlying ability is there, but it so often does not come out.

Another are in which reliability of a test can be questioned is in multiple-choice tests like TOEIC or TOEFL tests that only require test-takers to recognise a correct answer, not produce one (Brindley, 1998). The ability to recognise a correct answer can be improved by practising a certain test. This can, of course, improve a test taker's score but does not necessarily mean that he or she 'knows' a language, or can communicate in it better. Test-taking technique therefore becomes vitally important. An example of test technique being important can be seen in the TOEFL exam and other similar tests. A question from the short conversation section of the listening comprehension reads like this :

12. (A) His desk is made of metal  
 (B) He never meddles with his office workers  
 (C) His desk is on the right side of the office  
 (D) He works in the center of the office

Using two techniques which help improve scores in TOEFL, and are often contained in TOEFL textbooks, it is possible to surmise that the answer is (D) before listening. Firstly, do not choose sound-alike answers, and the answer is often a restatement of one of the speaker's utterances (usually the second). Obviously, metal and meddle sound like middle,

which is a synonym of centre. Most TOEFL course books contain the same advice, and an experienced TOEFL test taker could also get the answer without listening to the actual conversation. Although, this is perhaps an extreme example, it does raise serious questions of reliability for tests such as TOEFL and multiple-choice tests in general.

### **Efficiency**

Weir's (1990) and Hughes's (2000) third limitation of language tests is efficiency. This refers to practical considerations, such as time, space and money. These often restrict our ability to give a communicative test because constraints mean that it is sometimes difficult to include language production and important consideration in CLT (Fulcher, 2000). This lack of test efficiency means it is difficult to approach real CLT, although new technology does help with these considerations. The internet-based TOEFL introduced in 2007 is a very good example of the use of technology to improve a test, and to bring it closer to what can be considered a communicative language test. Sections of it have skill integration, it requires test takers to produce oral and written language and the productive sections of the test are criterion-referenced.

In the future, advances in computers will surely further improve our ability to provide communicative language tests.

### **Approaching Communicative Language Testing**

In conclusion, although certain features of CLT can be identified, the lack of agreement as to what knowing a language is still hindering its understanding. It is of course desirable to have tests which can measure

communicative competence (Barker and Zeff, 2000) and also which make reliable inferences as to whether a test taker can function in the real world. As discussed, the lack of agreement amongst theoretical models concerning these two major constructs, and hence the lack of construct validity of tests hinders progress towards approaching what can be called real communicative language tests.

## References

- Bachman, L and Palmer, A (1996) *Language Testing in Practice* Part 1, Ch 3 Oxford: Oxford University Press
- Bachman, L (1996) *Fundamental Considerations in Language Testing* Ch 8 Oxford: Oxford University Press
- Barker, D and Zeff, B.B (2000) *Communicative Language Testing* Sapporo: Hokkai Gakuen Journal of Humanities
- Brindley, G. (1998) *Assessing Listening Abilities* W. Grabe (ed.), Annual Review of Applied Linguistics, 18.
- Educational Testing Service. (2009). *Test of English as a Foreign Language*. Princeton, NJ: Author. Retrieved from <http://www.ets.org/toefl/ibt/about/>
- Fulcher, G (2000) *The 'communicative' legacy in language testing'*
- Hughes A. 2000 *Testing for Language Teachers*. Glasgow: Cambridge University Press.
- Sato, R (2010) *Reconsidering the Effectiveness and Suitability of PPP and TLBT in the Japanese EFL Classroom* JALT Journal Volume 32 Tokyo: JALT
- Spolsky, B (1985) *What does it mean to learn a language?* Language Testing 2.2: Bar-Ilan University
- Weir, C (1990) *Communicative Language Testing* Ch1 London: Prentice Hall