Language Proficiency and Assessing Classroom Achievement: A Literature Review

Parul Sood

Introduction

The goal of placement tests in language programs is to divide students into homogenous groups for the purpose of class assignment. In most intensive English language programs such as the AEC, the most typical criterion for dividing students into classes is proficiency in the target language (English). Accurate and reliable placement tests are crucial to the successful functioning of any English language program. This literature review discusses the concept of language proficiency and ways to measure it.

Language Proficiency

In order to place students into groups by proficiency it is first necessary to determine the nature of proficiency itself. Unfortunately, this task has proven to be surprisingly difficult, and researchers are still divided. Some have claimed that proficiency is essentially undefinable; for instance, Vollmer writes, "language proficiency is what language proficiency tests measure" (Vollmer, 1981, p.152). This circular statement was all that could be firmly said when asked for a definition of language proficiency in 1981. Recent research has shed a little more light on the concept of proficiency. Let us therefore seek a clearer account of proficiency that can serve as a theoretical guide in the construction of our assessment measures.

Language proficiency is defined as the level of competence, at which an individual is able to use language for both basic communicative tasks and academic purposes (TESOL, 1997). Language proficiency tests capture a student's holistic ability in a language generally showing the extent to which a student has acquired a second language. It is used to ascertain a learner's global proficiency in the areas of listening, speaking, reading and writing. Teachers of English to Speakers of Other Languages, Inc. (TESOL) developed a number of general principles derived from current research and theory on language acquisition. One of them is that language proficiency, which is a measure of implicit language acquisition as well as explicit language learning, is a gradual process moving through developmental stages.

Language Proficiency as Communicative Competence

In the sixties, understandings of language proficiency were primarily defined in terms of grammatical knowledge. This conception was largely based on Chomsky's view of language. Chomsky (1965) defined language in terms of a dichotomy: 'performance' and 'competence'. Competence refers to the intuitive knowledge that an 'ideal speaker-listener' has of the linguistic system of his language including rules of syntax and vocabulary; and performance refers to the individual's ability to use the language. He believed that the actual language produced by an individual was so degenerate or imperfect due to a number of variables (e.g., pauses, false starts, omissions, repetitions) that the only way to study linguistic competence was to remove language from its real-life production.

In the 1980s, the concept of language proficiency was defined as communicative competence and researchers included the communicative dimension of language in models of language proficiency. Like Chomsky, Canale & Swain (1980) made a distinction between knowledge of use and a demonstration of this knowledge. Knowledge of use is often referred to as communicative competence, and the demonstration of this knowledge as performance. They came up with a model which included four areas that are encompassed in proficiency: grammatical competence, sociolinguistic competence, discourse competence, and strategic competence. Grammatical competence is knowledge and mastery of the linguistic code, including vocabulary use, pronunciation, spelling, morphological rules and syntactic rules. The second aspect of proficiency, sociolinguistic competence, is the use of appropriate language in a variety of social contexts, i.e., the ability to

behave in a linguistically appropriate manner so as to achieve the desired results with the person or people being addressed. For instance, if a learner wants to go for a coffee break, s/he needs to know not only how to formulate a grammatical utterance, but also how to phrase the request so as to avoid offending the listener. The third aspect of proficiency, discourse competence, is the ability to choose meanings and forms that together create a unified and organized text or message. For instance, in order for a speaker to be able to express him/herself meaningfully in a phone conversation, s/he needs to master the use of greetings, introduction to a topic, and closure. The last aspect of proficiency is strategic competence, the ability to compensate for lack of second language knowledge, such as with the use of body language and gestures.

Language proficiency as it relates to communicative competence is also defined as the ability to get a message across. This ability includes social rules of language and use as well as knowledge of linguistics, socio-linguistics and discourse rules (Rivera, 1984). This means that the learner is able to use the language to receive and to respond to communication.

Social and Cognitive Academic Language Proficiency

A more holistic definition of proficiency, which increasingly guides language policies, research, and practice and, most recently, proficiency test development, is that proposed by Cummins (1981). Cummins (1981) distinguished between social language proficiency and academic language proficiency. Such skills may be observed through "visible language proficiencies of pronunciation, basic vocabulary, and grammar" (Cummins, 1981, p. 21). Academic language proficiency, or Cognitive Academic Language Proficiency (CALP), refers primarily to literacy related skills reflecting "students' ability to understand and express, in both oral and written modes, concepts and ideas that are relevant to success in school" (Cummins, 2008, p. 71). Interactive language corresponds to Cummins' Basic Interpersonal Communication BICS (Cummins, 1981). Basic interpersonal skills involve the ability to function in everyday communication between individuals; for example, being able to greet, apologize, go grocery shopping or clarify information.

Contextually, language proficiency can be expressed as "context-embedded" and "context-reduced" communication (Ovando & Collier, 1985). Context-embedded communication (e.g., greetings, making plans) is supported by meaningful paralinguistic and situational cues (e.g., gestures, intonation) in which the student can actively negotiate meaning. Context-reduced communication (e.g., listening to a lecture) is dependent upon one's knowledge of the language and requires a more active cognitive involvement. The degree towards which the student has automated context-embedded communication and is able to negotiate context-reduced communication affects the student's language proficiency level. Research findings suggest that ELLs may achieve social language proficiency within a period of up to three years of schooling in the United States, and academic language proficiency within a period of up to seven years (Cummins, 1981 as cited in Hakuta et al., 2000).

Today, most language tests are based at least on a theory of language proficiency. However, a lack of consensus as to the exact nature of language proficiency has led to numerous theories, and definitions of language proficiency. Therefore, the task of operationalizing these theories of language proficiencies into test questions is indeed daunting.

Assessing Classroom Achievement

Discrete-point Vs. Integrative Tests
Discrete-point tests are the traditional paper-pencil tests that are constructed on the assumption that language can be broken down into various components: the skills of listening, reading, speaking, and writing. Each item on a discrete-point test is intended to measure one linguistic element. Discrete-point tests have been met with

some criticism, particularly in light of more recent trends toward viewing language and its communicative nature and purpose and viewing language not merely as the arithmetic sum of all its parts. Oller (1979) argued against the notions of a language being segmented and tested in discrete points. Language testing needs to acknowledge second-language learning as comprised of many components and as communicative in nature. According to him, "language competence is a unified set of interacting abilities which cannot be separated apart and tested adequately" (Oller, 1979, p.37).

In response to discrete-point tests, Carroll (1961) introduced integrative tests (as cited in Oller,1983). Examples of integrative tests are dictation, oral interviews, conversation, essays and other coherent writing tasks. As the term suggests, an integrative test requires a test-taker to utilize one or more language skills (e.g. listening, reading, speaking, and writing) and one or more levels of language (e.g. syntax, semantics, lexicon, phonetics, morphology, phonology, discourse) at the same time. Additionally, integrative tests are much more direct and a more authentic way to examine whether the test-takers respond within the target context (Smith, 1994). In an effort to compromise between the discrete-point and integrative procedures, Smith (1994) further proposes that a test might be integrative in task but discrete point in evaluation. More specifically, a student may write a letter of complaint-an integrative test task, but it can be evaluated on particular errors in grammar and vocabulary-a discrete point evaluation. He believes that a test like this is best used for diagnostic purposes. For the purpose of proficiency levels, however, integrative tests were recommended both in task and in evaluation.

Problems with Traditional Proficiency Tests

Peregoy and Boyle (2005) argue that traditional language proficiency tests have limitations that can lead to inappropriate program placement. The limitations on standardized language proficiency tests are (1) the score is based on a single performance; (2) non-linguistic barriers such as unfamiliarity with the test format or procedures may interfere with the testing process; (3) test anxiety; and (4) different standardized language proficiency tests may give different levels for the same student.

Stoynoff & Chapelle (2005) point out that the kind of placement tests used in English language programs at various universities do not have the qualities of authenticity and interactiveness that are thought to be the hallmark of communicative language testing. For instance, there is no effort to keep topics consistent throughout the test taking experience; instead, the questions move from topic to topic in random fashion. In addition, Huerta-Macias (1995) remarks that the testing situation itself often produces anxiety within the student such that s/he is unable to think clearly. Huerta-Macias contends further that problems associated with traditional testing often mask what the student really knows, or in the case of ESL, what the student can do in his second language.

Although proficiency tests in general have shortcomings in terms of their authenticity and communicativeness, they do have the advantage of being relatively inexpensive, quick to take and easy to administer. Regardless of the difficulty for test designers to construct such tests, scoring them is relatively easy, fast, and objective (Brown & Hudson, 1998; Bailey, 1996). For this reason, they are still used as placement tests in many intensive English programs.

Farhady (1982) mentioned another alternative to testing in his works. According to him, one of the major problems with language programs is that none of them includes test taker characteristics as a potential dimension in language testing. He further argues that students with different educational backgrounds tend to perform differently on various language tasks. Therefore, Farhady suggested all language programs should do an initial detailed analysis of learner needs. Testing should be done in two phases; the first phase should assess the learners' general English. After a satisfactory score on this test, learners take a test that is developed on the basis of careful analyses of learner needs. This test would assess a selective functional proficiency of the

learners in various academic areas. This way we can avoid learner variables interfering with the decisions made on the basis of the test scores.

Finally, the validity of the commonly used proficiency tests is questionable (Oller, 1979). It is common that an individual might have different levels of proficiency in different areas. For example, someone might have better command of grammar in reading than in speaking; however, vocabulary command in speaking might be greater than in writing. No studies or tests to date have developed a set of criteria for determination of proficiency covering several aspects of linguistic elements (Oller, 1979).

What Are the Alternatives?

There is a lack of literature available on whether we should use classroom achievement as a criterion for placement over a proficiency test. If we do not use the proficiency tests, what are then the alternatives? How can we assess a student's achievement in a second language in a valid and reliable way? Current trends in the assessment of student learning recommend that alternative forms of assessment rather than standardized tests alone should be used to evaluate students' progress. In recent years, as pointed out by Brown (2001), language educators have made efforts to develop these non-test assessment tasks to distinguish them from traditional formal assessments. Most often, such innovations are referred to as alternative assessment (Huerta-Macias, 1995). Though there is yet no single definition for it, a variety of labels have been employed to distinguish alternative assessment from traditional standardized testing. Garcia and Pearson (1994), Hamayan (1995) and Smith (1994) reviewed those labels and created a list: performance assessment, authentic assessment, portfolio assessment, informal assessment, situated (or contextualized) assessment, and assessment by exhibition. For these types of assessment, students are required to "accomplish approximations of real-life, authentic tasks, usually using the productive skills of speaking or writing but also using reading or writing or combining skills" (Brown & Hudson, 1998, p. 662). The characteristics of performance assessments, by definition, lie in the fact that they are "direct, authentic, and highly contextualized" (Bailey, 1996, p. 215).

Alternative assessment can take many forms that can be adapted to varying situations. Examples include portfolios, journals, reading logs, conferences, videos of role plays, storytelling, oral interviews, self and peer evaluation questionnaires, work samples, and teacher observations or anecdotal records (Hamayan, 1995). Proponents of alternative assessment insist that indeed the method is valid and reliable, in as much as it has credibility (i.e. truth value) and auditability (i.e. Consistency) (Huerta-Macias, 1995). They also claim that alternative assessment is different from traditional testing in that it actually asks students to show what they know and can do. Students are evaluated on what they integrate and produce rather than on what they are able to recall and reproduce. The main goal of alternative assessment is to "gather evidence about how students are approaching, processing, and completing real-life tasks in a particular domain" (Garcia and Pearson, 1994). Hancock (1994) states that through the use of alternative assessments language programs will become more responsive to the differing learning styles of students and value diversity.

Alternative assessment procedures are nonintrusive to the classroom because they do not require a separate block of time to implement them, as do traditional tests. Moreover, the same day-to-day activities that a student is engaged in (e.g. writing, role playing, group discussion) are the basis for alternative assessment. Thus, there is little or no change required in classroom routines to implement them. Because alternative assessment is based on the daily classroom activities, it also reflects the curriculum. Because the data collected are based on real life tasks, furthermore, alternative assessment provides information on the strengths as well as the weaknesses of a student (Huerta-Macias, 1995).

Lewis-White (1998) conducted a study that investigated the use of alternative assessments and traditional tests for the purpose of placement. The results indicated that alternative assessments provided fuller more detailed

information while traditional testing was considered vague and yielded insufficient information upon which to base an instructional decision. In addition, The Arlington Education and Employment Program (REEP) conducted a two-year project to research, develop, field test and establish recommendations for alternative assessment for adult students. The research found that the alternative assessment process was helpful for teacher development and it enabled teachers to substantiate promotion and retention decisions. REEP used alternative assessment methods to determine language proficiency and to make decisions concerning placement.

In recent years, some ESL experts have emphasized the importance of a different type of assessment: self-assessment, where students are provided opportunities before and after units of instruction to assess their own performance. Pierce, Swain & Hart (1993) reported a study of five hundred 8th graders that suggested that self-assessment was a valid and reliable measure of language proficiency. To facilitate autonomy among university ESL students in this learning process, McNamara & Deane (1995) have designed a variety of activities that foster self- assessment. These activities engage the students in routine and ongoing self-assessment so that they develop a critical awareness of their language learning process (McNamara & Deane, 1995).

As alternative assessment is implemented in the classroom settings, it is worth discussing how teachers can best make rational choices among the available options and what these options mean in terms of mainstream communicative language testing at the levels of both theory and practice (Brown & Hudson, 1998; Brown, 2001). Determining how effective such classroom-based teacher assessment practices are may involve the incorporation of positive washback, feedback, and learner autonomy, each of which is addressed briefly.

Washback is generally defined as the influence of testing on teaching and learning (Bailey, 1996). According to Brown and Hudson (1998), the impact of washback on student learning can be either negative or positive. Emphasis on promoting positive washback is one main difference between traditional language proficiency tests, with a primary focus on linguistic competence, and communicative language tests, which attempt to measure the broader construct of communicative competence. Communicative language tests, in contrast to traditional proficiency tests promote positive washback. Altogether, such testing tends to have dramatic washback both in terms of curriculum innovations and student intrinsic motivation to prepare for the test (Bailey, 1996).

Oller (1979) has pointed out that the value of instruction is dependent on teachers making sacrifices of time and effort to offer their students meaningful and useful feedback. Consequently, assessment information must be more detailed and appropriate than a general score so that such feedback can become an integral part of the learning process (Bailey, 1996). Unlike traditional proficiency tests, Brown and Hudson (1998) noted that "conference, portfolios, and self-assessments all provide rich forms of feedback to the students that can be integrated into their learning" (p. 669).

Learner autonomy refers to "the philosophy that learners should have a large amount to say about what, how and how fast they learn" (Bailey, 1996, p. 270). In other words, autonomous learners can develop their own internal values in order to enable them to diagnose their weak and strong points, and to get a more realistic view of their proficiency and progress in general. Such incorporation of informal assessment and constant feedback is dynamic and formative, providing information about whether some students need extra work, and what steps to take next.

Implications

Vollmer (1981) argues that it is important to use different types of tests in order to get a broader view of a person's foreign language proficiency or communicative competence, and make sound decisions based on it. In addition, " any proficiency measurement should not be done at a single point in time alone but should definitely be repeated under varying circumstances because in all probability each single measurement will add

information to our understanding of a person's language ability and language use" (Vollmer, 1981, p.169). According to him, foreign language acquisition is a dynamic and a complex process. Therefore, he advised the use of different methods or different versions of the same test for language proficiency assessment, so that it can reflect complete assessment of all the aspects of language command.

Farhady (1979) advised that learner variables like previous educational and linguistic background of students would cause students to perform differently on different forms of tests (discrete and integrative). For example, a foreign student who has not experienced a cloze type test probably should not be expected to do well on it regardless of his/her language proficiency. Therefore, language proficiency tests need to consist of both integrative and discrete point subtests. He further argues "integrative tests enjoy a theoretical but not statistical or practical superiority over discrete point tests. However, it is also true that ignoring one type over the other, given the present state of affairs, may create some unwanted and unreliable biases for or against students from different countries (Farhady, 1979, p.356). The underlying processes of the language are too complex to be easily tested. A single type of test will not give a complete assessment of all the aspects of a language (Ingram, 1978, as cited in Farhady, 1979).

Drawing on the Canale and Swain model of communicative competence, Bachman and Palmer (1996) developed a theoretical model of language ability for language testing purposes. This language ability is specified within an interactional framework which considers two sets of characteristics that affect language use and language test performance: 1) the characteristics of the task or test situation (setting, test structure and instruction), and 2) the characteristics of the individuals (sex, age, native language, etc.). According to this model, tests should emphasize the communicative ability of test takers in meaningful contexts in which they are likely to need to use the language. Additionally, tests should ensure both authenticity, which is the degree to which the test tasks match the language use situation and interactiveness, which is the extent to which the test taker's language ability is engaged in accomplishing the task.

Conclusion

Language testing is at a critical stage of evolution. The current approach to language testing is a pragmatic one; incorporating ideas from both the integrative and discrete point approaches. It maintains some of the linguistic analysis of the discrete point approach while using context. The trend has shifted from testing linguistic elements to testing communicative functions. In addition, an effort is made to link the language-testing situation with the test-takers' experience. Consequently, pragmatic language tasks are intended to be as authentic as possible. For example, if a test-taker is going to listen to a story and retell it, a pragmatic test might have an adult actually read it (rather than an audiotape) with normal visual input (e.g., reader's gestures, number of story-linked pictures, etc.). We need to use meaningful authentic assessments that involve students in selecting and reflecting on their learning along with traditional proficiency tests. This will provide language teachers with a wider range of evidence on which to judge whether students are becoming competent, purposeful language users. Also, this will make our language program more responsive to the differing learning styles of students and value diversity therein. Finally, language programs that focus on alternative assessment along with traditional proficiency tests are likely to instill in students lifelong skills related to critical thinking that build a basis for future learning in any academic setting, and enable them to evaluate what they learn both in and outside of the language class.

References

Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice*. Oxford: Oxford
    University Press.

Bailey, K. M. (1996). Working for washback: A review of the washback concept in
    language testing. *Language Testing*, 13(3), 257-279.

Brown, H. D. (2001). *Teaching by principles: An interactive approach to language
    pedagogy* (2nd ed.). New York: Addison Wesley Longman, Inc.

Brown, J.D., & Hudson, T. (1998). The alternatives in language assessment. *TESOL
    Quarterly*, 32(4), 653-675.

Canale, M. & Swain, M. (1980). Theoretical bases of communicative approaches to
    second language teaching and testing. *Applied Linguistics* 1, 1-47.

Cummins, J. (1980). The entry and exit fallacy in bilingual education. *NABE Journal*, 4
    (3), 25-59

Cummins, J. (1981). The role of primary language development in promoting educational
    success for language minority students. In *Schooling and language minority students: A theoretical
    framework* (pp. 3-49). Los Angeles; National Evaluation, Dissemination, and Assessment Center.

Cummins, J. (2008) BICS and CALP: Empirical and Theoretical Status of the Distinction
    in Street, B. & Hornberger, N. H. (Eds.). (2008). *Encyclopedia of Language and Education, 2nd Edition,
    Volume 2: Literacy*. (pp. 71-83). New York: Springer Science Business Media LLC.

Farhady, H. (1979). The disjunctive fallacy between discrete-point and integrative tests.
    TESOL Quarterly, 13(3), 347-357.

Farhady, H. (1982). Measures of language proficiency from the learner's perspective.
    TESOL Quarterly, 16(1), 43-59.

Garcia, G. E., & Pearson, D. (1994). Assessment and diversity. *Review of Research in
    Education*, 20, 337-391.

Hakuta, Kenji. (2000). *How Long Does It Take English Learners to Attain Proficiency*.
    UC Berkeley: University of California Linguistic Minority Research Institute.

Hamayan, E. V. (1995). Approaches to alternative assessment. *Annual Review of Applied
    Linguistics*, 15, 212-226.

Hancock, C. (1994). Alternative assessment and second language study: what and why?
    Retrieved April 2, 2013 from the CAL Website: http://www.cal.org/resources/digest/hancoc01.html

Huerta-Macias, A. (1995). Alternative assessment: Responses to commonly asked
    questions. *TESOL Journal*, 5, 8-11.

Lewis-White, L. (1998). Assessing oracy and literacy in bilingual students: Getting the whole picture. In E. Sturtevant, J. Dugan, P. Linder & W. Linek (Eds.), *Literacy and community. the twentieth yearbook: A peer reviewed publication of the college reading association* (pp. 147-175). 851 N. Lassiter Rd., Readyville, TN 37149: College Reading Association.

McNamara, M. J., & Deane, D. (1995). Self-assessment activities: toward language autonomy in language learning, *TESOL Journal*, 5, 17-21.

Oller, J. W., Jr. (1979). *Language tests at school*. London: Longman.

Oller, J. W. (1983). *Issues in language testing research* Newbury House Publishers, , Rowley, MA 01969

Ovando, C. & Collier, V. (1985). *Bilingual and esl classrooms: Teaching in multicultural contexts*. New York: McGraw-Hill Book Company.

Peirce, N. B., Swain, M., &Hart, D. (1993). Self-assessment in two French immersion programs. *Applied Linguistics*, 14, p25-42.

Peregoy, S. F. and Boyle, O. F. (2005). *Reading, Writing, & Learning in ESL: A Resource Book for K-12 Teachers*. Longman Publishers USA, 88-97.

REEP alternative assessment project (353 special project) final report, July 1, 1994-June 30, 1995.

Rivera, C. (Ed.). (1984). *Communicative competence approaches to language proficiency assessment: research and application.* Clevedon, A von, England: Multilingual Matters,Ltd.

Smith, M.S. (1994). *Second language learning: Theoretical foundations*. New York: Addison Wesley Publishing Co.

Spence-Brown, R. (2001). The eye of the beholder: Authenticity in an embedded assessment task. *Language Testing*, 18(4), 463-481.

Stoynoff, S. & Chapelle, C. (2005). *ESOL tests and testing*. Alexandria: Teachers of English to Speakers of Other Languages

Teachers of English to Speakers of Other Languages, Inc. (1997). *The ESL standards for pre-k-12 students*. Alexandria, Teachers of English to Speakers of Other Languages, Inc.

Vollmer, H. (1981). Why are we interested in general language proficiency? In J. Alderson & A. Hughes (Eds.) *Issues in Language Testing*. London: The British Council, 152-171.