Considerations of Conducting Spoken English Tests for Advanced College Students

Dr. Byron Gong

English Department
Soochow University, Taipei, Taiwan
byronygong@yahoo.com

Abstract

The ability to speak the target language has become a major focus of language teaching and testing. How to assess college students' speaking ability within TEFL teaching programs at universities has become more important during the past years. However, it remains a difficult task for many TEFL teachers to conduct a reliable and valid speaking test for English-majored students. This phenomenon is widespread among universities in Taiwan. This paper provides a notional discussion of adopting task approach as a valuable means of assessing English majors' speaking proficiency. Although task-based approaches still need much investigation both theoretically and empirically, it is believed that a task-oriented approach is a valid assessment of the productive communicative ability of the test taker, especially for English-majored students at advanced level.

Key words: speaking tests, evaluation of speaking tests, task approach

Introduction

Along with the development of ELT theories and teaching practice, the ability to speak the target language has become a major focus of language teaching and testing. The status of speaking ability within ELT teaching programs at colleges and universities, especially in Taiwan, has risen considerably during the past years. Currently, an English-majored student's English proficiency is usually not assessed by direct testing of the student's oral ability. Instead, a college student's ability of spoken English is assumed to be acceptable according to his/her paper test scores of other courses. In reality, most universities in Taiwan do not conduct a graduation examination that contains test of spoken English as part of the assessment. However, it remains a difficult task for many college ELT teachers to conduct a reliable and valid speaking test in their classrooms, especially at the advanced level for English-majored college students.

In view of this situation, there are good reasons for language teachers to adopt a communicative approach and feasibly use an oral proficiency assessment to test students' speaking ability by asking students to complete tasks. The purpose of this paper is to present an analytical study of a task-oriented method, or task approach, that is used to test students' oral proficiency, which is an option for teachers to evaluate their students' speaking ability in terms of productive and interactive performance. In this paper, the writer will start with a brief discussion of theoretical consideration of conducting an oral proficiency test in view of productive and interactive performance. Backed up with a small-scale experimental study, the writer of this paper found task-oriented approach appears to be a kind of more valid assessment of the productive communicative ability of test takers, especially for English-majored students at advanced level. This paper presents a further discussion of the considerations of speaking tests and concepts underlying the test validity of task-oriented testing in college ELT programs. Considerations of task-oriented approach in terms of its important features, problems, and washback effect on teaching are discussed. The paper concludes with some suggestions on the practical implications of the task-oriented approach in classroom. The writer hopes that the notion presented in this paper will provide a new starting point for a possible exchange of experience in testing of spoken English at tertiary level.

Theoretical Discussion

In English teaching practice, many college ELT instructors in Taiwan often ask questions, "What type of speaking test should we use for this final exam: interview, role-play, or prepared monologue?" "How should I score the speaking test?", "What can we do to reduce interview bias?", and "How valid is this midterm speaking test?" (Li, 2001; Liu and Hu, 2000; Zhou, 1998;). It is not the intention of this paper to answer these questions, but these questions would lead us to a wider scope of theoretical framework and practical experience. Although there is no "best answer" to these questions as situations and purposes differ from one context to another, this paper offers a discussion of some key issues of speaking tests in order to provide practical suggestions for testing of oral English ability with the notion of "task approach" (Bachman, 1990; Bachman and Palmer, 1996; Fulcher and Davidson, 2007; Wigglesworth, 2008). So, let us now take a look at a wider range of theoretical considerations.

Speaking test in general has the positive washback effect of encouraging students to study the language skills required for oral communication. Many college teachers in Taiwan realize that speaking ability is an important part of the learner's general English proficiency. As Weir (2004:103) pointed out: "If we wish to make statements about capacity for spoken interaction we are no longer interested in multiple-choice, pencil-and-paper tests, that is, indirect tests of speaking where spoken language is conspicuously absent." In line with this perspective, the writer proposes that the task-oriented approach should be especially considered in the test of spoken English as a type of performance test to assess language proficiency by asking students to give productive and interactive performance. It is believed that, if well designed according to the needs of various real situations, the task approach, as an option of speaking test, not only increases test authenticity and validity but also can be used to considerably reduce most test deficiency such as halo effect, interview bias, intra-rater reliability, etc. This rationale is especially significant when we take a look at the current methods used in speaking tests at colleges and universities, especially in Taiwan where few universities have established a complete speaking testing system that ensures a productive and interactive performance even for the English majors.

Traditionally, due to various factors, the test of spoken English, if it is to be conducted at advanced level at universities in Taiwan, is conducted indirectly and invalidly. In many cases, such an oral test of English ability would contain mechanical greeting dialogues, sentence reading, story telling, passage recitation, situational conversations, or a prepared speech based on a given topic. However, such an assessment of students' speaking ability may be criticized for lack of authenticity, because much of the testing material in such a test would be artificial or it is a piece of prepared speech. Besides, as negative washback effect, students may be encouraged to produce well-memorized answers in a prepared speech topic. Therefore, it would be more valid if oral English tests can be authentic. In other words, an oral test should be a true assessment of spoken ability rather than an indication of the ability. As Fulcher and Davidson (2007:249) pointed out that the best way of assessing speaking performance is in "the real world", but the problem is how to define "the real world". In addition, how to score the limited samples collected in an oral test remains a problem for many college ELT instructors in terms of both methodology and empirical evidence. Regarding these questions, ELT instructors may consider adopting different assessment methods based on the conception the test of spoken English should be direct and productive rather than to assess how well a student can read a passage in English, or recite a well-prepared speech. As for the general trend of language testing, Wigglesworth commented, "In the context of language testing and assessment, performance assessment has become increasingly important over the last two decades, and has been the focus of substantial empirical investigation" (2008:111).

Conceptually, the writer holds that a valid speaking test of English proficiency at advanced collegiate level for English majors should directly test the candidate's performance in authentic interactive tasks. Meanwhile, such a speaking test should also be a performance test that assesses test taker's productive skills rather than receptive skills. There is much literature discussing the relationship between task and performance testing. Although there are different views on the possible definition of authentic tasks used in speaking test and foreign language performance assessment can be conducted in a variety of contexts, it is increasingly recognized that the testing of spoken English should get students to take part in direct spoken language activities that should be as much authentic as possible. Many researchers have realized the importance of task authenticity. Weir (2004:108)

pointed out that "the test task need to reflect the target situation in terms of functional operations, and we need to try and ensure that the test score resulting from the task may be premised on tasks that exhibit both theory and context-based validity". Regarding the task-based approach, Wigglesworth (2008:117) stated that "A central tenet of task-based language assessments is that the tasks are designed to represent authentic activities which test candidates might be expected to encounter in the real world outside the classroom." But Wigglesworth also pointed out that: "the issue of authenticity is not a trivial one, and the extent to which specific tasks can represent authentic real world activity has attracted considerable debate and empirical investigation, using a variety of different approaches" (2008:117).

Meanwhile, it is necessary to briefly discuss the meaning of "task" for the purpose of this paper. According to Nunan (1989:10), we have the following definition of task: "A piece of classroom work which involves learners in comprehending, manipulating, producing or interacting in the target language while their attention is principally focused on meaning rather than form". This definition provides a general framework of task that can be used for classroom teaching. For the purpose of this paper, by task-oriented approach, the writer of this paper refers to a type of speaking test involving performance in a test of productive skills, whose main purpose is to elicit an examinee's English language ability. Specifically, a task refers to a given topic that an examinee is expected to talk about so that his/her language ability, rather than the completion of the actual task itself, can be assessed. Also by college English programs (including those for both English-majored and non-English-majored), the writer refers to general English proficiency courses of four skills, i.e. listening, speaking, reading, and writing rather than ESP programs. The rationale is based on the concept that a practically reliable and valid speaking test at colleges and universities should be developed according to the communicative approach.

An important point of task-oriented approach to the testing of spoken English involves the two issues for us to consider: to what degree that the test tasks can be formed as valid samples of oral tasks that students should be able to perform in the real world; and how authentic these tasks used in oral test should be. As Hughes (1989:101) pointed out: "We want to set tasks that form a representative sample of the population of oral tasks that we expect candidates to be able to perform. The tasks should elicit behaviour which truly represents the candidates' ability and which can be scored validly and reliably." However, it is difficult to be sure whether the collected speaking samples through a test can be considered as valid materials that represent the target domain of the required ELT context. Thus, we have two further issues to focus on. As Fulcher and Davidson pointed out:

> The first problem that we encounter is defining 'the real world', and trying to create that in the test. Then comes the problem of how to score the limited sample that we are able to collect in the test, so that the score carries meaning for the target domain and is relevant to any institutional context that is preparing learners to become competent communicators within that domain. (2007:249)

The two problems mentioned by Fulcher and Davidson are fundamentally important, and there is much literature providing various considerations. However, it is the researchers' own responsibility to adopt the "best" testing method according to a needs analysis. For the purpose of this paper, the writer would focus on the considerations of adopting the task-oriented testing method that is used to test senior English majors' oral ability in a university in Taiwan.

Considering oral test validity viewed from the aspect of authenticity, the writer holds that a good speaking test of English proficiency should directly test the candidate's performance in authentic interactive tasks. This concept will inevitably raise a fundamental question: "What is to be tested in oral English in terms of communicative competence?" For the purpose of this paper, the tentative answer can be generally linked with the testing of the ability to speak the target language fluently and properly. This means the focus of oral English testing, on the one hand, is to examine a testee's ability to express themselves in the target language in an acceptable manner in terms of both linguistic coding and decoding and mutual cultural understanding; but on the other hand, the focus is also on the ability of a testee's communicative ability.

A Small-scale Experimental Study of Task-oriented Testing

Based on the above discussion of speaking tests, the writer of this paper believes that the most decisive factor in successful testing of oral English ability relies on how the speaking test is conducted. There are many possibilities open to testers in terms of elicitation techniques for oral testing (Underhill, 1987:27). The experiment used in this study involves a one-year comparative study of one control group and one experimental group by assigning the 3rd-year English majors into the two classes at random. For the experimental group, i.e. Class 2, the oral tests were the task-based interview given to a class of 29 junior English-majored students, and a relevant course syllabus was designed for Class 2. Meanwhile, students of the control group (Class 1) keep its usual course syllabus, teaching method, and assessment approach. But at the end of the academic year 2009, two groups of students were tested with the same oral test — a set of mock-up IELTS Speaking Tests according to the IELTS test materials by Cambridge University Press, which is to ensure the proficiency range in terms of test validity. Like what is required in the IELTS Speaking test, all students of two classes were given the same three-part speaking test with different task topics. Part 1 is a 4-5 minute introduction. For Part 2, a task card is given to the student who has then one-minute preparation time for the topic on the test card. The student should talk about the issue continuously for one to two minutes. Finally, related questions were asked in Part 3.

Considerations were given to categories for assessment, which focused on communicational effectiveness, and were mainly based on the "performance criteria" proposed by Underhill (1987:96), i.e. length of speech, speed, complexity, accuracy, appropriacy, repetition, etc.. But for the purpose of this study, the writer adopted the four criteria of *Fluency and Coherence*, *Lexical Resource*, *Grammatical Range and Accuracy*, and *Pronunciation* (ESOL, 2006) according to the criteria used in the IELTS Speaking Test. It is believed that the IELTS Speaking Test is a highly valid test that can be used for a mock-up test for the purpose of this study. However, the 9-band marking system used in IELTS was replaced by a 100-point marking system for this study. The mock-up IELTS test was given to all the students of the two classes. Accordingly, the results of the mock-up IELTS Speaking Test revealed a significant difference in their scores, which is shown in Table 1 and Table 2.

Table 1   **Descriptive Statistics of Test Results of a Mock-up IELTS Speaking Test**   (Full score = 100)

	N	Range	Minimum	Maximum	Mean	Std. Deviation	Variance
Class 1	29	45	50	95	77	10.4	108.07
Class 2	29	18	80	98	86	4.1	16.83

Table 1 shows that the mean scores of the two classes based on a mock-up IELTS test are 77 and 86 respectively, and the experimental class performed much better than the control class. In addition, in order to find more empirical proof, the writer of this paper conducted an Independent Samples Test between the usual final oral test of Class 1 and final task-based oral test of Class 2 so as to see if these two kinds of tests are fundamentally different. In order to meet the pre-condition for an Independent Samples Test, the writer first conducted tests of normality of the scores of usual final oral

Table 2   **Independent Samples Test**
(Usual Final Oral Test of Class 1 and Final Task-based Oral Test of Class 2)

		Levene's Test for Equality of Variances		t-test for Equality of Means						
								Std. Error Difference	95% Confidence Interval of the Difference	
		Sig.		t	df	Sig. (2-tailed)	Mean Difference			
		Lower	Upper	Lower	Upper	Lower	Upper	Lower	Upper	Lower
Scores C1C2	Equal variances assumed	16.663	.000	-4.420	56	.000	-9.172	2.075	-13.33	-5.015
	Equal variances not assumed			-4.420	36.51	.000	-9.172	2.075	-13.37	-4.966

Note. p > .05

test of Class 1 and task-based test scores of Class 2 with the help of computer software SPSS, which provided One-Sample Kolmogorov-Smirnov Test and the results are .573 and .342 of Asymp. Sig. (2-tailed) for Class 1 and Class 2 respectively, indicating both test scores distributions are normal. Next, the result of the analysis of the Independent Samples Test between the usual final oral test of Class 1 and final task-based oral test of Class 2 can be reflected in Table 2, which reveals that these two kinds of oral tests are very different in nature. There is a statistically significant difference between the two types of tests, i.e. the usual final oral test of Class 1 and final task-based oral test of Class 2, (value $p < .05$ at +.000 sig.), which is a very clear index to claim that these two tests are dissimilar to each other. This means different test methods may have different washback effects on teaching; but different teaching methods may also bring about different performance results.

Thus, as Table 1 and Table 2 illustrate, the control group which was instructed with usual teaching method performed poorly than the experimental group on the same mock-up IELTS Speaking test. In addition, the final oral tests of the two groups are significantly different, which in turn may assert that students had made progress in significant manner that would not have been made had the test and the teaching procedures been different. However, the test approach suggested in this study is to design task-based oral tests of communicative competence according to the learning needs, so that these tests will have positive washback effect on the courses. It is worthwhile for college ELT instructors to realize that appropriate test method may bring about strong washback effect, and the effective task-based test method should be used to test students English communicative ability. However, on the other hand, the writer also acknowledged that more detailed studies on the task-based method should be conducted. Next, further considerations for the task-based test method will be discussed.

Considerations in Adopting Task-based Approach for Oral Tests

As proposed at the beginning, the writer believes that task approach is a sound and dependable testing method for college ELT instructors to adopt for testing oral English at the advanced level for English-majored university students in Taiwan. Task approach requires test candidates to use their language knowledge and competence in actual or real-life communication. The test candidates need to take the initiative in the test in order to complete the task, which differs from many other test methods by which test candidates just demonstrate their receptive skills rather than productive language skills. Furthermore, the "real-life" or "authentic" performance in a task approach test reflects prevailing view that knowing a language includes not just knowledge of the formal features of the language but also knowledge of how to use the target language appropriately for communicating in particular situation (Bachman & Palmer 1996). In speaking test, task approach includes tasks as adopting a specific role in a role play, describing the advantages and disadvantages of an issue, presenting an argument to a small group of peers, etc. Meanwhile, we should be aware that this type of task, which often attempts to replicate real-life tasks, involves more than one skill; making a report, for example, may involve reading and even listening to questions from audience, as well as speaking. Underhill (1987) pointed out that other factors also need to be considered when designing an oral test: adequate local knowledge; a human approach; a suitable balance; and the ability to adapt and improve the test.

More specifically, conducting a task approach test requires teachers or examiners to do much careful preparation. There are a number of things that language teachers should bear in mind (Mousavi, 2002: 696):

1. Directions for the task approach test should be so clear that both the examiner and the student should know exactly what the student must do. The task itself may be anything that the students have to do with the target language. The point is task instructions should be very clear so that the students know exactly what is expected of him or her and thus they will not talk off the subject.
2. The task should be reasonably narrow in scope so that the student(s) can complete the given task within the allotted time.
3. Teachers should work out their scoring scale to ensure objectivity. Teachers can use either analytic scoring method or holistic scoring method. If they choose analytic scoring method, more attention should be paid to each category so as to reduce or avoid halo effect if possible.

Advantages of the Task-oriented Approach for Speaking Tests

Traditional oral tests are often conducted through teacher-student interview, role-play, picture description, prepared speech, or tape-recording, etc. Although some of them have special features for certain purposes, they are criticized for lack of authenticity, interviewee's poor initiative, one dominating role over the other, low reliability and validity (Hughes, 1989; Li, 2001; Nunan, 1989; Wigglesworth, 2008). Compared with other test methods, task approach suggested in this paper has the following advantages (Fulcher and Davidson, 2007; Liu and Hu, 2000; Mousavi, 2002; Underhill, 1987; Wigglesworth, 2008).

- It improves the relationship between the oral examiner and the test candidate(s) as the student needs to take initiative in completing his/her task while the role of the teacher or examiner becomes less dominating when he observes students to give productive and interactive performance. In this respect, it differs from the traditional speaking test that the teacher asks questions, and the student just passively gives answers.
- The task-oriented approach involves real-life simulation of potential activities on the part of test candidates who are required to do something that looks like authentic examples of language use in real world. The test assessment can be used to predict the candidate's ability to perform that or similar tasks in the future. Therefore, the test conducted by task approach directly examines the candidate's performance in authentic interactive tasks, and this method can be considered to have relatively high face and content validity.
- As the test scores can provide detailed information of actual performance, which can be obtained by audio or video recording, this method has positive washback effect on teaching and learning. This is also one of the main advantages of this approach.
- In practice, the task approach test can be conducted with two or more students at the same time, which is less time-consuming when compared with other oral test such as face-to-face interview test.
- As the task approach test can be conducted with two students or a small group of students, it is likely for the examiner to reduce interviewer bias so as to lower the unreliability as much as possible. (By interviewer bias, it refers to the interviewer's own biased feelings, attitudes, opinions, or expectations of the test candidate. Similarly is halo effect.)

Problems of Task Approach

After analyzing the advantages of the task-oriented test method, the weaknesses inherent in its application have to be acknowledged. The weakness of the task-oriented method can be viewed from the following aspects: samples of real world activities, test reliability and validity. Fulcher and Davidson pointed out: the first problem is defining 'the real world', and trying to create that in the test (2007:249). There are many factors and conditions for college ELT teachers to consider while designing the task-based testing system. In other words, inappropriate samples of real-life tasks may influence the test reliability and validity. It is difficult to be sure whether the collected speaking samples can represent authentic real world activity, and whether they can be considered as valid materials that represent the target domain of the required ELT context (Wigglesworth, 2008).

Regarding the reliability problems of the task-oriented approach, examiners also need to guard against a tendency for severity or leniency because this type of test, like many other speaking tests, heavily relies on the examiners' subjective judgments of students' language performance. Therefore, it is important for the examiners to be aware of this reliability problem. Examiners will need to be trained to interpret and apply assessment criteria in a consistent way.

As far as test validity is concerned, although the task approach tests are said to be based on authentic or real-life language use, critics point out that it does not mean that the sampling of language involved is of a sufficient amount for assessment purposes (Wigglesworth, 2008:117). In other words, a task may be authentic, but may be impoverished in terms of what it reveals about the learners' language, which means authenticity cannot guarantee validity. However, the writer of this paper would argue that such a problem of validity exists in other speaking tests too, and it could be even more serious. To reduce the negative effect of this validity problem, teachers need to establish a larger sample of language activities for task approach tests. This refers to an establishment of a test

task bank which provides adequate tasks for teachers to use in testing students spoken English ability. Task items in the task bank can be selected by experienced teachers so as to increase the face, content, expert, and construct validity.

The last possible weak point of this task approach test is that this kind of test could be time-consuming for teachers because they need to spend much time in preparing for every possible detail of a task-based speaking test. However, it is also believed that once a systematic testing system is established, the task-based test could become time economic for most teachers. In this aspect, the testing format and tasks of the IELST Speaking tests might be valuable for teachers to consider in their teaching practice.

Discussion of Key Aspects in Successful Task-based Testing for English Majors

In order to deliver a successful task-based speaking test, it is necessary for us to take a brief look at some key relevant issues. Generally, a successful speaking test is not only linked with test designing on one hand based, but also with test implementation on the other. But the writer of this paper would consider the following four aspects have priority over other concerns, i.e. test organization, method, content, and scoring regarding a successful task-based speaking tests.

Test organization involves coordinating various sections that an oral test needs. For a relatively large-scale oral test, for example, when the number of test takers is 100 or more, there should be careful consideration of all the relevant or possible factors, such as test sites, number of oral examiners, inter-rater training, the time required, etc. There are more factors to be taken into consideration especially when the number of test takers is large.

As far as the *test method* is concerned, a reliable and valid oral English test is also connected with how the speaking test is conducted. In this respect, direct test, or real-task test, has been highly recommended by many test experts and organizations. This is because direct test is intended to measure the productive skills of speaking ability (Liu Renqing, 1991; Bachman, 1990; UCLES, 1997; Council of Europe, 2001), which is a trend for language testing to move from measuring test takers' receptive language skills to productive skills in terms of the use of language for communicative interactions, or authenticity. Theoretically, direct test involves authentic tasks on the part of test takers, which provides a close relationship between test performance and future use. Therefore, direct test is more likely to have higher face validity. In practice, it means if we want to test how well our students speak English, we should give a test that requires students to speak directly rather than ask students to choose a choice of conversation indirectly on a multiple-choice test. In other words, much care should be taken into consideration so as to avoid using indirect tests or semi-direct tests ("tape-based" tests).

In practice, the most commonly used test methods for English majors are interview, interaction with peers, and response to tape-recordings. But as Hughes (1989,104) has pointed out: "The relationship between the tester and the candidate is usually such that the candidate speaks as to a superior and is unwilling to take the initiative." Therefore, it is important for teachers to realize that task-oriented test approach may provide more opportunities for students to take the initiative than traditional methods, not to mention the problems of response to tape-recording, or a prepared speech. So, the task-oriented method suggested in this paper can provide an effective option for college ELT teachers.

Test content is also a crucial factor to be taken into consideration in speaking test for English majors. The content of classroom achievement tests is closely related to what has been taught in a textbook or syllabus. However, if the syllabus is badly designed or textbooks and other teaching materials are poorly chosen, then the results of these classroom achievement tests can be misleading. The concern of test content is content validity, which may include two broad aspects: relevance to and coverage of the complete syllabus. In an achievement test, which is mostly the case at universities, content validity is achieved by means of an adequate sample of the target domain. But in a proficiency test, content validity would be based on the needs analysis. In short, designers task-based oral tests should be fully aware of the content validity while choosing samples of the real world activities.

Test scoring is also an important factor to consider in speaking tests. In general, test scoring plays an important role in ensuring that a speaking test is fairly conducted as good tests have positive washback effect on teaching and learning. In practice, no matter whether examiners use holistic scoring and analytic scoring (Hughes, 1989; Li, 2001; Underhill, 1987), most oral tests are scored by

every individual examiner. Therefore, training of examiners is extremely important considering certain test deficiency such as halo effect, interview bias, poor intra-rater reliability, poor inter-rater reliability, etc. In classroom teaching, it is desirable if the oral test at midterm and final should be conducted by an instructor who doesn't teach the class. The writer of this paper would like to point out that cross testing should be used as long as a standardized scoring criterion is used and proper training of examiners is conducted. In addition, in order to increase test reliability, it would be the best if more than one examiner gives scores to every student's performance whenever it is possible. So far, we have briefly discussed the four key aspects of test organization, method, content, and scoring with respect to major concerns in task-oriented approach to oral English tests.

Conclusion:

This paper has attempted to show that the task-oriented approach is an effective option for college teachers to adopt in their oral tests of senior English majors. The writer has briefly presented theoretical and practical considerations of this test method, i.e. task approach. It is suggested that this task approach method can be used as a practical approach to the assessment of college English majors' oral English proficiency in conjunction with the needs of productive communicative ability. The small-scale experimental study in this paper shows that task-oriented approach has positive washback effect on students learning and can result in better test performance. This paper has shown that there are good reasons for language teachers to adopt a communicative approach and use the task approach to test speaking by asking students to give productive and interactive task performance. It is hoped that the task-oriented approach can be a valid method for the assessment of students' productive communicative ability, especially for English-majored college students at advanced level. It is hoped that this study has provided a new starting point for further studies and efforts in terms of putting more desirable testing method of oral English ability into practice.

REFERENCES

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

Bachman, L. F., (1990). *Fundamental Considerations in Language Testing*. Oxford: Oxford University Press.

Council of Europe. (2001). *Common European Framework of Reference for Languages: Learning, teaching, assessment*. Cambridge: Cambridge University Press.

ESOL. (2006). *IELTS Scores Explained DVD*. Cambridge: Cambridge ESOL Publications.

Fulcher, G. and Davidson F. (2007). *Language Testing and Assessment*. London: Routledge.

Gong, Byron. (2009). *Successful National English Proficiency Testing Systems in Mainland China and Taiwan*. Paper presented at the 2009 IAEA Annual Conference in Brisbane, Australia.

Hughes, A. (1989). *Testing for Language Teachers*. Cambridge: Cambridge University Press.

Liu. Renqing and Hu, Zhuanglin. (2000). *Language Testing and Methods (Revised)*. Beijing: Foreign Language Teaching and Research Press.

Li Xiaoju. (2001). *The Science and Art of Language Testing*. Changsha, China: Hunan Province Educational Press.

Mousavi, S. A. (2002). *An Encyclopedic Dictionary of Language Testing 3rd ed*. Taipei: Tung Hua Book Co. Ltd.

Nunan, David. (1989). *Designing Tasks for the Communicative Classroom*. New York: Cambridge University Press.

Underhill, N. (1987). Testing Spoken Language. Cambridge: Cambridge University Press.

Weir, C. (2004). Language Testing and Validation: An Evidence-based Approach. Basingstoke: Palgrave.

Wigglesworth, Gillian. (2008). Task and Performance Based Assessment. In Elana Shohamy and Nancy H. Hornberger (Ed.), *Encyclopedia of Language and Education, 2nd Edition, Volume 7: Language Testing and Assessment* (pp. 111-122). Springer Science+Business Media LLC.

Zhou Shen. (1998). *A Summative Report on the Validity of TEM (Test for English Majors) Test. In English Language Testing*. Shanghai: Shanghai Foreign Language Education Press.