

PAPER DETAILS

TITLE: Current Problems and Proposed Solutions for Testing Speaking: Opinions of EFL Teachers


AUTHORS: Devrim HÖL

PAGES: 27-36

ORIGINAL PDF URL: <https://dergipark.org.tr/tr/download/article-file/498707>

Current Problems and Proposed Solutions for Testing Speaking: Opinions of EFL Teachers

Devrim Höl¹

¹ Pamukkale University, Denizli, Turkey / devrimhol@gmail.com 

Abstract

This paper is an attempt to explore EFL teachers' opinions about the problems they encounter and proposals for solutions with regard to the speaking test implemented at a state university in Turkey. The participants were 22 EFL instructors teaching English at Pamukkale University in Turkey. The data were gathered through a semi-structured interview form developed by the researcher. The form was developed based on the review of literature. Among the content analysis methods, categorical analysis method and inductive analysis method were used to analyze the data obtained through semi-structured interview form administered to the teachers. The results were analyzed and discussed in terms of the speaking test and its implications for teaching and learning activities and testing. In the light of the study findings, majority of the teachers thought that time limitation is a factor that affects the validity and reliability of speaking test. In addition, teachers stated that formative assessment during the academic year would produce better results performances in testing speaking.

Keywords

Testing; speaking;
English as a
foreign language

Introduction

Speaking is a productive skill for the learners and it is doubtless a must to communicate for the learners learning English as a second language. As multilingualism gained importance throughout the world, testing speaking skills took its place in many institutions and universities for the past few decades; however, it requires a lot of difficulties to administer a valid and reliable test. As a result of this, sometimes having the real validity and reliability may be difficult in speaking tests

and may be differences in terms of appliance and assessment in every institution. It requires a good organization, planning and implementation process. In addition to all these, “teachers”, as the assessors and interlocutors during the test, is an important factor for a valid and reliable test, because testing speaking skills necessitates a lot of time, effort and objectivity. On the other hand, it is important to keep an anxiety-reduced atmosphere, and teachers have to perform their best whether they are giving the test for the first candidate or the last one, and, so there is no doubt that their perceptions and opinions for testing speaking skills is a great value to design a more effective and objective test. Heaton (2003) states that testing the ability to speak is the most important aspect of language testing, however, it is a very difficult skill to test, because it is far too complex and challenging. Testing speaking is widely regarded as the most challenging of all language tests to prepare, administer, and score (Madsen, 1983).

In our paper, we investigated a) what the opinions of EFL teachers who apply the speaking test are on testing speaking, b) what the current problems they confront while conducting the speaking test are, and, if there is, c) we wanted to find out the proposed solutions made by the teachers. The selection of the research site was motivated by the fact that the university is a medium-scaled university in Turkey, and it has been applying speaking tests in English as a foreign language for 10 years, which is enough duration to get insights of the test applying procedures and experiences of the EFL teachers. In our research site, speaking skills of the learners are assessed at the end of each module, which lasts eight weeks for all levels ranging A1 (beginner level), A2 (elementary), B1 (pre-intermediate), B2 (intermediate) according to CEFR (Common European Framework). The oral tests take between 5 and 15 minutes for each candidate. There are always two raters and one candidate throughout the test. The test specifications include three steps; a) introducing the candidates themselves, also a warm-up and anxiety reducing process, b) picture story, c) description or narration

Literature review

Challenges in speaking tests

Testing speaking is one of the skills to be tested in a valid and reliable way. It requires a much greater performance than those administered with paper-and-pencil test

(Fitzpatrick & Morrison, 1971). This difficulty not only comes from the nature of the test itself, which can be defined as a “live” test, but also the challenges before, during, and after the test make it haunting for both the candidates taking the test and examiners themselves. Madsen (1983) explains the speaking tests as the most challenging of all language exams, similarly, Luoma (2004) defines the assessment of speaking as challenging because “there are so many factors that influence our impression of how well someone can speak a language, and because we expect test scores to be accurate”.

Firstly, a test should be valid to produce reliable outcomes. If a test is to be goal oriented and yield the best results, validity is the most crucial aspect in test development (Luoma, 2004). Bachman and Palmer (1990) state that “the test to be useful, reliability, validity, authenticity, interactiveness, impact, practicality must be hand in hand”. Cohen (1994) sees validity as a problem and believes that “students may not speak in the class the way they actually would if performing in the real world”. Validity can be defined as the agreement between a test score or measure and the quality it is believed to measure (Kaplan & Saccuzzo, 2001). In other words, it measures the gap between what a test actually measures and what it is intended to measure.

Secondly, it is certain that “time limitation” is the most challenging and consuming part of the test. This is the most difficult one to arrange among the other tests because testing the learners orally takes more than pencil-paper tests. Each subject is tested individually or in pairs. In consequence, the tests should be prepared so appropriately that it should not take ages to test the candidates (Güllüoğlu, 2004).

In addition, another drawback of the test may be anxiety for the test takers. Second language anxiety has a negative effect on the oral performance of speakers of English as a second language (Woodrow, 2006). For the examiners, that is, interlocutors and assessor in the test, test anxiety students are having during the test may be a burden because it can be time consuming and also it may affect the performance of the students.

As the fourth dimension, physical conditions are of a great concern in testing speaking. It may not only affect the performance of the candidates during the test, but

also the objectivity of the raters. Hughes (1989) mentions that it is necessary to put candidates at their ease so that they can show what they are capable of. It is already a very stressful test for the students and from the beginning to the end, it is vital to have correct planning and decent atmosphere. He also states that tests should be held somewhere that is quiet and free from interruption. It may, however, not be the case in some institutions. Some students wait in the corridors while the candidate is taking the test in a classroom very near to the corridor, which is full of noise, and this may affect the performance of the test takers, and they will be disturbed by noise and other people. Anybody trying to hold a conversation will become irritated if repeatedly disturbed by noise or by other people. Not only for the students but also for the assessors is the same case; Hughes (1989) points out that it is important that scoring should take place in a quiet, well-lit environment.

In addition to factors affecting the assessment of speaking skills, there are some reliability concerns in testing speaking skills in English as a foreign language, and these concerns may create trouble and leads the test to be unreliable (Underhill, 1989). Indeed, student-related reliability is the weak or unsatisfying performance of the student because of some reasons which may be personal such as an illness or bad mood. His performance throughout the test may not reflect the student's level, and causes an unreliable result. His performance may also be affected by the examiners' behaviors during the test. Another important factor can be rater-reliability which reflects the human-errors during the test process. In many tests, the tester is also the rater. Throughout the whole interview process, the tester establishes the proper level for the interviewee while conducting the test (Kuo & Jiang, 1997).

As another factor, descriptions and criteria for the test should be used or developed. On a speaking test, having the student say something appropriate is only half the job, scoring the test is also challenging (Madsen, 1983). According to Underhill (1989), "a rating scale is a series of short descriptions of different levels of language ability". Its purpose is to describe briefly what the typical learner at each level can do so that it is easier for the assessor to decide what level or score to give each learner in a test. Harmer (2004) defines assessment scales as a way of specifying scores that can be given to productive skill work which is to use to create 'pre-defined descriptions of performance'.

Another important difficulty confronted in speaking tests is rater objectivity. It is vital that there should be stability for the oral tests when the same test is administered by the same scorers and there should be a consistency on different times and environments (Çopur, 2002). Therefore, subjectivity is inevitable in scoring procedure as long as human being is involved in the process. The best solution to this problem is having more than one rater for each performance. The scores of each rater can be correlated to see if the results are consistent. It would be difficult to expect the same performance from the testers from the beginning to the end if the test takes too long and the number of candidates is high.

Methodology

Study site and participants

The present study was conducted at Pamukkale University, School of Foreign Languages in Denizli during the 2015-2016 academic years. A total of twenty two full-time teaching staff⁰ teaching English courses to students with different levels from elementary to intermediate participated in this study. All participants were native speakers of Turkish and their ages range between 29 and 40, and they have expertise in teaching English between 6 and 18. The participants included 16 female and 6 male teachers. The participants reported that they have expertise in teaching listening, speaking, writing, and reading skills in English. Most importantly, they have at least 5 years experience in preparing and conducting speaking tests in English. They took part in the study voluntarily.

Procedure

The researchers conducted one to one semi-structured interviews with the participants in Turkish in the offices in the School of Foreign Languages. The questions for the interview were developed by the researcher and consisted open-ended questions. All interviews were audio-recorded and each interview lasted between 17 to 25 minutes. The interview was two-fold covering the issues below:

1. Current problem(s) they have encountered in implementing the speaking test,
2. Suggestions and proposals for the solution(s) of problem(s) they covered

Additionally, the participants were required to put the themes they mention to order from the most important to the least important in order to find out not only the problem or solution itself but also the degree of it.

Analysis

The audio-recorded interview protocols were transcribed by the researchers. The analysis of the interview protocols was qualitative and exploratory. Qualitative data analysis has three stages, which are data display, summarizing and interpreting, respectively (Büyükoztürk et al. 2008). First, the transcribed interview protocols were divided into themes, each of which represents a single part, which may represent the first part of the interview (problems encountered) or the second part (proposals for solutions). Secondly, data gathered were categorized and coded by the authors, and the codes were discussed carefully again to avoid the problems if there is a disagreement between the researchers. As the final step, frequencies and percentages were calculated.

Findings

The focus of this research is on investigating EFL teachers' opinions about the problems they encounter and proposals for solutions with regard to the speaking test implemented at Pamukkale University, School of Foreign Languages.

The results about the problems EFL teachers encounter for the Speaking test are as follows:

Table 1. Problems Encountered for the Speaking Test

Codes	f	%
P1 Time is limited	22	100
P2 High anxiety levels of the students during the test	13	59.09
P3 Subjectivity of the assessors during the test	11	50
P4 Test materials should be revised	10	45.45

As can be seen from Table 1, there were four listed codes related to problems EFL teachers have for the speaking test in their institutions. Firstly, all of the participants state that time in the speaking test is not enough or it is a time consuming test. As the second problem, more than half of the participants think that students have a high level of anxiety during the speaking test, which may be one of the factors that affect

their performance in the test. Thirdly, half of the participants state that the assessors or raters have a problem for conducting an objective test. Finally, nearly half of the participants believe that test materials used during the test should be revised.

Table 2 demonstrates the solutions proposed by EFL teachers related to the speaking test applied in their institution. Notably, all of the participants state that formative type of assessment, rather than a summative one, should be used for testing speaking skills of learners. As another solution, participants think that test materials used to test speaking skills of EFL learners should be revised/modified and be more authentic and up-to-date. In addition, nearly 40 % of the participants think that examiners should have training for testing speaking. Finally, and interestingly, more than 20 % of the EFL teachers who took part in the study think that “Speaking skills should not be given for levels lower than B1”.

Table 2. Solutions Proposed by the Participants

Codes		f	%
S1	Formative Assessment should be used	22	100
S2	Test materials should be more authentic and up-to-date	20	90.90
S3	Assessors should be trained for an objective test	9	40.90
S4	Speaking skills should not be given for levels lower than B1	5	22.72

Discussion and conclusion

This small-scale interview study conducted at a state university, in Pamukkale University School of Foreign Languages, in Turkey examined Turkish EFL teachers’ opinions about the problems they encounter and proposals for solutions with regard to the speaking test implemented. Important findings were revealed from the study. A notable finding was the strong emphasis on the time limitation of the speaking test. Participants state in-depth analysis that time they spend on testing speaking skills causes them feel tired and may affect the objectivity of the test. Furthermore, participants stated that learners have high-anxiety during the test. It should be also noted that test-takers may have never taken a speaking test before, so preparing the learners for speaking test and designing a syllabus which would familiarize the learners with speaking skills rather than just paper and pencil tests can be a way to decrease the level of anxiety. As another critical issue, participants think that the

assessors are subjective to assess the test takers, which requires an academic skill and which can be regarded as an important concern for all the stakeholders ranging the institution itself to the learners. Finally, it was reported in the study that test materials used in the assessment of speaking skills should be revised/modified. This finding can be important for the institutions, and although the test materials which are used during speaking tests are decided by the teachers, it seems from the findings that institutions must be more careful while choosing the test materials and when they have more authentic and goal-oriented materials, they will have a more valid and reliable test. Conducting the test with the materials which are to the target would doubtlessly improve the reliability of the test. On the contrary, it would be a test born dead.

To summarize the solutions proposed by the participants, first of all, teachers believe that formative assessment should be used rather than a summative one. This may be due to high anxiety level of learners during the test, and so they cannot produce as much as they can do during speaking classes, and they have more limited time during the test than in-class speaking assessment procedure. The hypotheses can be that when the speaking skills of the learners assessed throughout the academic term or academic year, anxiety level of the learners would decrease, which, in turn, increase their performance. Secondly, participants believe that test materials should be more authentic and up-to-date. This was particularly evident in their comments during the interview sessions, and they mostly believe that group work, pair work and role play activities should be inserted to the syllabus with other activities like discussions and presentations about a specific topic.

This research, however, has some limitations. Firstly, the findings cannot be generalized as it just represents a group of EFL teachers teaching English in university, which uses its own test materials, procedure and materials. The context in other institutions or schools may be different and produce different results. However, the results of the present study may be a guide way for them. Although the results give insights about the problems encountered by EFL teachers as assessors during the speaking tests and suggestions they have made for a more efficient speaking test system, further research in different institutions or groups would really produce different findings.

As the final word, key results of the study can help all the stakeholders including teachers and institutions execute a more careful and effective assessment of speaking skills, which, in turn, improve the performance of the learners speaking skills.

References

- Bachman, L. F. & Palmer, A.S. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.
- Büyüköztürk, Ş., Kılıç-Çakmak, E., Akgün, Ö. E., Karadeniz, Ş. & Demirel, F. (2008). *Bilimsel araştırma yöntemleri* [Academic research methods] (2nd ed.). Ankara: Pegem Akademi.
- Cohen, A.D (1994). *Assessing language ability in the classroom (Teaching Methods)*. Boston: Heinle ELT.
- Çopur, D. (2002). Testing first year FLE students' oral performance using four speaking test methods in spoken English II course and students' attitudes towards these speaking testing methods (Unpublished MA thesis). Middle East Technical University, Ankara.
- Fitzpatrick, R. & Morrison, E.J. (1971). Performance and product evaluation. In R. L. Thorndike (Ed.), *Educational measurement* (pp. 237-270). Washington DC: American Council of Education.
- Güllüoğlu, O. (2004). Attitudes and perceptions of the students at Gazi University towards testing speaking (Unpublished MA thesis). Gazi University, Ankara.
- Harmer, J. (2004). *The practice of English language teaching*. Longman.
- Heaton, J.B. (2003). *Writing English language tests*. London: Longman.
- Hughes, A. (1989). *Testing for language teachers*. Cambridge: Cambridge University Press.
- Kaplan, R.M. & Saccuzzo, D.P. (2001). *Psychological testing: Principle, applications and issues* (5th Ed.). Belmont, CA: Wadsworth.
- Kuo, J. & Jiang, X. (1997). Assessing the assessments: The OPI and the SOPI. *Foreign Language Annals*, 30(4), 503-512.
- Louma, S. (2004). *Assessing speaking*. Cambridge: Cambridge University Press.

- Madsen, H. S. (1983). *Techniques in testing*. New York: Oxford University Press.
- Underhill, N. (1989). *Testing spoken language: A handbook of oral testing techniques*. Cambridge: Cambridge University Press.