

LANGUAGE TEACHING AND EVALUATION: AN ANALYSIS ON STUDENTS' SPEAKING PERFORMANCE ASSESSMENT

Putri Novitasari

UNIVERSITAS WAHIDIYAH KEDIRI

princess.putrinovita@gmail.com

Abstract

Assessment in classroom activities can be derived by using various sources, within oral or written. Related to the assessing students' performance, this study will discuss on the assessment of students' speaking performance. The data collected by giving questionnaire to the three different English Department lecturers from three different private universities. The questionnaire consists of eight points which is used to get the information related to speaking assesment. The items of question are; 1) the criteria of a good speaking performance, 2) the way of developing scoring rubric, 3) the common errors of learners, 4) The common activities in speaking performance, 5) the ice breaking activity, 6) the most contributed activity of learners' score, 7) the percentage of final test score, and 8) the problem in assessing speaking performance. A descriptive qualitative design was use to conduct this research. The research results reveal two points which are in lining among three lecturers. First is the scoring system. Generally among three lecturers use certain number criteria for scoring learners' speaking performance. Each lecturer use their own scale criteria but, actually they use same method, maximal and minimal scales. The second similarity is among three lecturers actually have same argument or maybe that problem is generally happened in assessing speaking performance. The problem is time consuming while assessing learners' presentation, speech, or debate. Finally, more specific focus on extensive speaking performance assessment may give broader finding and detail information.

Keywords : *evaluation, assessment, speaking.*

EVALUASI DAN PEMBELAJARAN: ANALISA TERHADAP ASSESMENT PADA KEMAMPUAN SPEAKING SISWA

Abstrak:

Assesmen dalam kegiatan belajar mengajar dapat dilakukan dengan berbagai cara, baik lisan atau tertulis. Sehubungan dengan assesmen yang dilakukan pada penampilan siswa, penelitian ini akan membahas tentang assesmen pada kemampuan speaking siswa. Data pada penelitian didapatkan dengan metode kuisisioner yang diberikan kepada 3 orang tenaga pengajar yang berasal dari 3 universitas swasta yang berbeda. Kuisisioner terdiri dari 8 item, yang berisi antara lain; (1) kriteria kemampuan speaking yang bagus; (2) rubrik penilaian; (3) kesalahan yang umum dilakukan siswa; (4) kegiatan pembelajaran dalam speaking; (5) kegiatan pembuka; (6) kegiatan yang berpengaruh banyak pada nilai siswa; (7) persentase nilai akhir siswa; (8) kendala dalam proses assesmen. Hasil dari penelitian ini memunculkan dua simpulan yang didapat dari hasil kuisisioner yang dilaksanakan. Hasil tersebut pertama dari penyusunan rubrik penilaian speaking, dari ketiga narasumber masing-masing menggunakan rentang nilai tersendiri tetapi metode penilaian yang mereka gunakan sama yaitu menggunakan standar nilai tertinggi dan terendah. Kedua, kendala dalam proses assesmen, ketiga narasumber mengungkapkan bahwa kendala yang muncul adalah waktu terutama pada kegiatan presentasi, pidato, dan debat.

Kata kunci: assesmen, speaking

A. INTRODUCTION

Knowing and measuring students' achievement on studying language takes the important role in classroom activity. The term of language assessment means assessing and checking students' progress on their performance, and it is supposed to be done by the teacher. Assessment which done by the teacher is divided into two, formal and informal assessment. In the term of formal assessment, it is related to exercises, usually within a relatively short time limit, systematic, planned sampling techniques constructed to give teacher and student an appraisal of students' achievement (Brown, 2004). While informal assessment Harris and McCann (1994) says that informal assessment is a way of collecting information about students' performance in normal classroom conditions. Having assessment on students' speaking performance will be varying in some activities. Because of that reason, the writer limits the activity on classroom presentation or academic presentation, since this kind of presentation hold by college students. The objective of this study is to identify the lecturer's technique on assessing students' speaking performance, the

good assessment criteria of students' performance, the way and problem of assessing speaking in classroom activity. The discussion and finding of the data will be interpreted and answered those objectives of the study.

B. REVIEW TO RELATED LITERATURE

Speaking is considered as one of the most important skill of learning language that should be mastered for the ability to communicate in foreign language. It helps the students to communicate with other people. Because of the importance of English, it is taught widely in formal school start from elementary schools up to the university. Speaking as one of productive skills is commonly defined as the use of language orally. It is the ability to express something into spoken language and concern ideas into words to make other people grasp the meaning. Harmer (2001: 269) says that speaking is actually the ability to speak fluently presupposing not only knowledge of language features, but also the ability to process information and language 'on the spot'. Through speaking, someone can produce the language and can express their ideas, thoughts, feelings, opinions, emotions, and reaction to other person" (Evans and Green, 2006). Meanwhile, Fulcher (in Tamimi and Attamimi, 2014: 5) defines speaking as the verbal use of language and a medium through which human beings communicate with each other. According to Chastain (in Castillo, 2007: 78) speaking is a productive skill and it involves many components. Speaking is more than making the right sounds, choosing the right words or getting the constructions grammatically correct. To master speaking skill, students should focus on some of indicators such as; vocabulary, pronunciation, grammatical structure, fluency, and comprehension.

Among four language skills, speaking is a difficult one to assess with precision, because speaking is a complex skill to acquire. There are components are generally recognized in analysis of speech process such as pronunciation (including the segmental features, vowel and consonants; and the stress and intonation), grammar, vocabulary, fluency, comprehension (Brown, 2004). In addition, Brown (2004) stated that each type of classroom speaking performance should be practiced as usual in order to make the speaking good in function. In intermediate or advanced level of English, speaking performance of the students can be assessed formally and informally, and both should be supported by clear scoring rubric. Since extensive speaking focus more on variations on monologues with minimal verbal interaction, and is more complex, relatively lengthy stretches of discourse (Brown, 2004), the criteria of assessment is also more complex than intensive speaking. The forms are varied from scoring rubric, checklist, and etc.

According to Thornbury (2009), there are two main ways in assessing speaking: either giving it a single score on the basis of an overall impression (holistic scoring) or giving a separate score for different aspects of the task (analytic scoring). Holistic scoring (e.g. giving an overall mark out of, say, 20) is adequate for informal testing

and progress while analytic scoring is more complicated way of scoring, but it is assumed to be more reliable and fairer since it forced the testers to take a variety of factors into account. In assessing speaking informally, it can be done by observing oral performance of the students. Oral performance here is a certain speaking activity that covers both class performances as a whole and individual performance (For example, all students are working in groups discussing a particular issue and solving problem).

Assessment for oral production needs more attention to be done. It is because an oral assessment not only just give score than report it but also how it can built a good communication and interaction. "It has become viewed as an essentially interactive process, in which the teacher can find out whether what has been taught has been learned, and if not, to do something about it." (William, in Ounis, 2017). Assessing speaking's performance has some fundamental issues which make this kind of assesment sometimes feel reluctant to do. Those issues are a) whether or not the test is used as intended, and (b) what its consequences may be (Bachman & Purpura, in A. Ghufrant, 2016). According to Shepard (in Ounis, 2017), "classroom assessment refers to the kind of assessment that can be used as part of instruction to support and enhance learning." By those statements, writer concludes that assessing students' speaking needs to be moved from traditional method, students' achievement to build a good interaction and communication.

The general objective of speaking assessment is assessing students' way of expressing their ideas orally. Their content of oral activity should cover some criterias of oral assessment. They have to fullfill them because they will be scored if they can achieve the criteria such as; "(1)clear content, (2)well organized, and (3)good language in terms of: intelligible pronunciation, appropriate grammar, appropriately chosen words" (A. Ghufran, 2016). The way of students' speaking activity has to cover all speaking competence. Those speaking comptence are; (1) Content The content should be relevant to the topic given in the test. It means that in conveying the spoken text, the whole content of the text should refer to the topic stated by the raters. (2) Organization,the test participant should organize his/her sentences in systematical organization; and (3) Language" (A. Ghufran, 2016). Based on those statements, the writer concludes that assessing students' speaking can not be separated from some components. Those components are used to measure and score the students' speaking performance. They are how the students deliver their idea, their idea, and attitude.

C. METHOD

The data collected by giving questionnaire to the three different English Department lecturers from three different private universities. The questionnaire consists of eight points which is used to get the information related to assessing speaking in extensive level. The items of the question are; 1) the criteria of a good

speaking performance, 2) the way of developing scoring rubric, 3) the common errors of learners, 4) The common activities in speaking performance, 5) the ice breaking activity, 6) the most contributed activity of learners' score, 7) the percentage of final test score, and 8) the problem in assessing speaking performance. A descriptive qualitative design was use to conduct this research.

D. THE FINDINGS AND DISCUSSION

In order to analyze the result of questionnaire, writer use eight points of questionnaire than describe them. The first point is about the criteria of a good speaking performance. For this point, lecturer stated that some important criteria in order to be said that a learners have good speaking performance are content or material of presentation, learners' fluency, and their accuracy while delivering their presentation. This is as stated by (Djiwandono,in Ghufran, 2016) that content, organization, and language must get more attention in speaking. The next point is the way of developing scoring rubric for assessing speaking performance done by the lecturer. In constructing scoring rubric the lecturer said that he adopt the criteria from some resources than adapt them based on the learning need and classroom condition. According to (Louma, in Ghufran, 2016) One way to elicit the construct of speaking ability for a certain context is through a scoring rubric which informs test users what a test aims to measure. The third point that will be analyzed is the common errors of learners while they are having presentation. Based on the questionnaire responded by the lecturer, the common errors in performing their presentation are weak of pronunciation and grammatical error. The fourth point of questionnaire is the common activities used to derive score in speaking performance. The common activity used by the lecturer is presentation. By this activity, learners supposed to present certain material or topic divided by the lecturer. The term of ice breaking activity will be discussed by the fifth point. In this point lecturer stated that he sometimes give this kind of activity in order to break the learners' anxiety, lead the learners to the main material, also check their readiness to get presentation.

The sixth point is about the activity which contributes most in learners' score. Responding this point, first lecturer stated that his common activity for assessing learners' speaking is classroom presentation which done in group. In extensive speaking, one of the activities which can be assessed is oral presentation. Lecturer stated that the score of learners affected by group project likes presentation and individual assignment. That kind of activity also will be calculated with their final test at the end of semester. Lecturer also stated that learners' final test contributes about 30% to overall learners' score instead of presentation, individual assignment, and lecture attendance. And the last point of questionnaire being described is the problem in assessing learners' speaking performance. By this point writer want to know about the lecturer's problem while he is assessing learner speaking performance in context of classroom presentation. The second lecturer adopts

scoring rubric from some resources and adapts it based on the students' needs. There are some criteria in scoring rubric such as topic, opening, content, closing and delivery. Then, the rating scale consists of five categories. The point 5 is for excellent, the point 4 is for good, the point 3 is for average, the point 2 is for fair, and the point 1 is for poor. There is no zero point, the highest point is 5 and lowest point is 1. All points are calculated for getting the last score. From the total score, there is different calculation in content. The content is times two because the lecturer assumes that the content is the important part in speech.

The first criterion in scoring rubric is topic. This criterion consists of two items that are content and time allotment. The lecturer assess the topic which is delivered by the learners, is the topic suitable or not with the theme given by lecturer. The second criterion in scoring rubric is opening. This criterion consists of two items that are attention getter and review. The third criterion is content. This criterion consists of two items that are main point, sub point, supporting fact-illustration-example, and appropriateness. The lecturer assesses those items whether those items relate to topic or not. How those items support each other and those items of the content flow naturally. The fourth criterion is closing. This criterion consists of two items that are review and memorable statement. The last criterion is delivery. This criterion consists of some items that are good smile, confident, interaction, vocal variety, body language and eye contact. Those items relate to the manner, how the learners deliver speech in personality.

Based on the scoring rubric used by the lecturer, she is very details in scoring the learners' speech. From opening until closing, she assesses all items in speech even the learners' manner in speech. She adopts some resources as her guidance to make scoring rubric. Then, she sees the learners' need, what should add and what should omit. In other word, she adapts scoring rubric. According to Brown (2004), there are two criteria of scoring rubric in assessing the learners' presentation performance. They are content and delivery. Rating scale consists of four categories that are excellent, good, fair and poor. The highest score is four and the lowest score is zero. It seems similar with the lectures' scoring rubric. Only some items make different. In addition, the scoring rubric of the second lecturer can be said that is good enough.

Assessing students' performance is not easy. There are some problems that lecturer faces. Based on questionnaire of the lecturer, the most problem is time allotment. In assessing speech, the lecturer needs a lot of time because the learner should perform one by one. This is in line with (Ludenberg, in Issaacs, 2016) that "aural/oral skills are less measurable because they are less tangible, more subject to variation, and probably will involve the cumbersome and time-consuming expedient of the individual oral examination." Then, the other problem is in the learners. Most of the learners feel nervous in speech. They do not confident in delivering their speech. It makes the content of speech does not deliver in smoothly and clearly.

Furthermore, this assessment is not to be the final learners' score. Based on questionnaire, this assessment is only 40% of all score. The score will calculate more with other score, for example the daily activities' score. In conclusion, this assessment is quite good just need added activities that make learners more enjoy because based on questionnaire the lecturer did not commonly use the game.

Based on the data, the third lecturer provides the scoring rubric of debate that contains 4 criteria of assessment. Each criterion consists of scale ranged from 1 to 4. The highest scale 4 and the lowest scale is 1 (There is no zero score). The lecturer assesses the learners viewed from; 1) Organization & Clarity, 2) Use of Argument, 3) Use of cross-examination and rebuttal, 4) Presentation Style. Organization and clarity means that the learners can deliver the main arguments and responses well and clearly in order. The use of arguments is how the learners provide strong, persuasive and relevant arguments to support their idea. The use of cross-examination and rebuttal means how the learners analyze the weaknesses of the opposite group and defend their own argument from the opposite group. Presentation Style relates to the learners' body language starting from tone of voice, clarity of expression, precision of arguments, etc to keep audiences' attention and persuade them of team's case, and learners are assessed on how they used the style convincingly.

Because the description is need to be clear in order to derive reliable and objective perception (Brown, 2004), the lecturer should explain more what does actually the term *convincingly* means? What are the criteria in order the learners are said to have convincing style of debate? Is there any correlation between convincing and the errors produced? Etc. The total score is determined by the total of the scales. The gap of score among the learners with one different scale is 6. It means, the learners who have close total of points (16, 15, 14) can achieve the total score that quite contrast among them (100, 94, 88). The question that comes up is 'is reliable and valid?' or 'is it fair enough?' If we think logically, the level of the ability and performance of the learners with only one different gap of point is not too contrasted. They may only differ from some certain criteria, but the skill that they have in doing debate may almost the same since the gap of the point is only 1. Unfortunately, the gap of score according to the total point is quite contrast. It is obvious, for example, when we compare learner who gains 15 points with learner B who gains 10 points. The score must be contrast since the gap of total points is clearly different. It also may be assumed that the learner A has higher competence in doing debate than learner B because the contrast gap between the points.

E. CONCLUSION AND SUGGESTION

1. Conclusion

Assessment for speaking performance seems different with the other language skills. The problem in time limitation is the most common problem stated by the three lecturers. Assessing speaking performance focused on extensive speaking has more specific elements and need more detail criteria. Since this kind of speaking done in college level which has advance assumption so, the lecturer should assess them by any point of views. In assessing learners' speaking performance among presentation, speech, or debate, lecturer tend to take into account to their fluency, accuracy, performance, delivery, and their manner while having performance in front of the audience. In case of make the assessment being objective, lecturers used certain scoring rubric for it. Among three lecturers stated that they are actually adopting the scoring rubric than adapt it based on his or her learners' need and condition. Why condition? Since each university has their own background level of learners so, that is why adaptation needed to be done by lecturers. The observation fulfilled by the writer raises some errors commonly produced by the learners. Generally their common errors are pronunciation, grammatical use, and mental disturbance likes nervousness and lack of confidence.

The way how the learners or speakers manage their time and floor turn taking also being the focus assessment by the lecturers. Even though, extensive speaking tend to monologue, at least their friends and lecturer is their audiences. That is why learners or speakers should be aware to the audiences. In scoring the learners, among three lecturers stated that assessing learners' presentation, speech and debate is not as the final assessment.

The first lecturer stated that final score contributes 30% of overall learners' score instead of their daily project, presentation, and lecture attendance. The second lecturer also stated that learners' speech only contributes 40% for overall score instead of daily activity score likes assignment or attendance. The third lecturer tends to focus on scoring learners' speaking performance in debate by assessing their technique in delivering their argument and rebuttal flow. The other aspect which assessed by the lecturer in debate are manner and method. Method here means that how the debaters raise their issue, argue their idea, prove their argument by the data, rebut the argument, defend the argument, and conclude the team statement. While manner here means that how the debaters behave while having debate. Eventhough they are standing on different side and possible to have "war" argument but, keep try to behave well is being a good point. Straight to the technical scoring seems to have wide and illogical range. The level of the ability and performance of the learners with only one different gap of point is not too contrasted. But, unfortunately, the gap of score according to the total point is quite contrast. It is obvious, for example, when we compare learner who gains 15 points with learner B who gains 10 points. The score must be contrast since the gap of total points is clearly different. It also may be assumed that the learner A has higher

competence in doing debate than learner B because the contrast gap between the points. That is the weaknesses point of that kind of scoring criteria.

Focusing in this analysis, the writer concludes that there are two points which are in lining among three lecturers. First is the scoring system. Generally among three lecturers use certain number criteria for scoring learners' speaking performance. The second similarity is the problem in assessing learners' speaking performance. Among three lecturers actually have same argument or maybe that problem is generally happened in assessing speaking performance. The problem is time consuming while assessing learners' presentation, speech, or debate.

2. Suggestion

By this study, it will be more beneficial if the writer also give the suggestion for further study in same field. Since this study only takes at the general activities of extensive speaking, writer expects that every single activity will be analyzed in deeper aspect. Furthermore, more specific focus on extensive speaking performance assessment may give broader finding and detail information.

Bibliography

- A. Ghufran Ferdiant. (2016). Developing the assessment instrument of speaking. OKARA Journal of Languages and Literature, Vol. 1, Tahun 1, Mei 2016.
- Al-Tamimi, Nasser Omer M. And Attamimi, Rais Ahmed. 2014. Effectiveness of Cooperative Learning in Enhancing Speaking Skills and Attitudes towards Learning English. *International Journal of Linguistics*, Vol. 6, No. 4, page 27-45.
- Brown, H. Douglas. 2004. *Language Assessment-Principles and Classroom Practice*. San Fransisco: Longman.
- Castillo, Claudia Yanive Prieto. 2007. *Improving Eleventh Graders' Oral Production in English Class through Cooperative Learning Strategies - Mejora de la producción oral de estudiantes de grado undécimo en clase de inglés, a través de estrategias de aprendizaje cooperativo*. Profile 8, 2007, Page 75-90.
- DOROBĂȚ, Dumitru. 2007. *The Methodology of Evaluation and Testing*. Ministerul Educației și Cercetării Proiectul pentru Învățământul Rural.
- Evans and Green. 2006. *Cognitive Linguistics and Introduction*. Britain: Edinburgh University Press.

- Harmer, Jeremy. 2001. *The Practice of English Language Teaching Third Edition*. Cambridge: Longman
- Harris, Michael and McCann, Paul. 1994. *Handbook for the English Classroom Assessment*. Macmillan Publishers Ltd.
- Isaacs, T. (2016). Assessing speaking. In D. Tsagari & J. Banerjee (Eds.), *Handbook of second language assessment*. Berlin: DeGruyter Mouton
- Thornbury, Scoot. 2009. *How to Teach Speaking*. Pearson Education.
- Wisconsin Education Association Council. 1996. *Performance Assessment*. Education Issues Series, May 1996.
- Wren, Douglas G. (2009). *Performance Assessment: A Key Component of A Balanced Assessment System*. Department of Research, Evaluation, and Assessment March 4, 2009 Number 2.