

DEVELOPING INSTRUMENTS FOR EVALUATING THE AUTHENTIC ASSESSMENT FOR SPEAKING SKILL AT JUNIOR HIGH SCHOOL

Hidayanti Rohimah Nurdin¹, M. Zaim², and Refnaldi³

¹FBS Universitas Negeri Padang, Padang, Indonesia, ✉ hidayantirohimahnurdin@gmail.com

²FBS Universitas Negeri Padang, Padang, Indonesia

³FBS Universitas Negeri Padang, Padang, Indonesia

Abstract

Authentic assessment which is applied in 2013 curriculum has been used as an assessment method in assessing students' language competence. It engages the students in a meaningful learning experience for the real world situation. A guidebook to implement authentic assessment for speaking skill for junior high school has been developed by some experts. However, it needs to evaluate this product to see if the assessment already meets the criteria of an authentic assessment, based on the requirement of 2013 curriculum, related to content and authenticity. Moreover, the appropriate instruments are needed to evaluate the authentic assessment for speaking skill. This study aimed to develop the instruments that will be used to evaluate the implementation of authentic assessment for speaking skill at junior high school. The researcher used evaluation research to evaluate the implementation of authentic assessment for speaking skill at junior high school. The evaluation instruments were developed by using theoretical based approach and were validated by the expert in language assessment and English language teachers at Junior High School. Based on the validation from the experts and English teachers, it can be concluded that the evaluation instruments are valid and can be used for evaluating the implementation of authentic assessment for speaking skill at junior high school.

Key words: Authentic assessment, speaking skill, junior high school

Introduction

Since the 2013 curriculum has been enacted in the early of 2013, the implementation of authentic assessment as a method of assessment used in the 2013 curriculum has obtained many obstacles in the field. Several researchers found some obstacles in implementing the authentic assessment. First, Fitriani (2014: 1152) found a number of difficulties in the implementation of authentic assessment in the 2013 curriculum. They are the purpose of assessment, fairness and grading. The teacher gets some difficulties in determining the purpose of the assessment, taking students' fairness and to make an objectivity and reliability of scoring students' work. It takes longer time since the teacher has to assess many factors as authentic as possible.

Second, Ewing as cited in Linh (2016) found that authentic assessment requires high cost and great efforts, but lacks public support, and it may not be appropriate in all cases. Moreover, the teachers are confused to use authentic assessment because they do not have enough knowledge about authentic assessment, thus teachers have to know the characteristics of authentic assessment before applying them in the teaching-learning process. Then, they also have to be able to design authentic assessment as suitable as possible to students and teaching-learning process objectives. Later, designing authentic assessment requires considerable work prior to the beginning of a subject so that teachers are suggested to be cooperating with schools, parents, and administrations in order to gain the main purpose of the assessment.

The aim of authentic assessment is to measure skills in various contexts that reflect real world situations. It assesses real students' ability as closely as their real life experience in a subject. This kind of assessment will assess students' achievement based on the reality they get in the learning process. It is a comprehensive assessment of all instructional activities covering both process and product of learning. By measuring the process and product during the learning, from the beginning to the end of the learning process, it is expected to help the teachers monitoring the students' progress, achieving the learning purpose and making decision toward students' knowledge, attitude and skill.

Furthermore, Authentic assessment used to measure how students are able to apply their knowledge and skills to be more meaningful in their life. As Gulikers (2004) defines authentic assessment as assessment which the students have required to use the knowledge, skills and behaviors they possessed, applying to which will depend on the level of similarity of the situation they will face in the real world. Using authentic assessment requires learners not only to answer the test correctly, but also to apply their knowledge, skills and attitude to solve their problems of everyday life or professional life in the future (Jaedun, 2014: 16).

Later, a guidebook to implement authentic assessment of speaking skill for junior high school has been developed by some experts. This product aligns with the 2013 curriculum and suitable to assess students' speaking skill. It has been designed systematically by the experts following the standards of designing the speaking assessment. In order to know whether these assessments appropriate or not to use to assess students' speaking skill in Junior High School, it is necessary to evaluate this product. It is conducted to make sure whether these assessments are good or not to used in assessing students' speaking at Junior High School. Moreover, the appropriate instruments are needed to evaluate the authentic assessment for speaking skill. However, it needs to evaluate this product to see if the assessment already meets the criteria of an authentic assessment.

In order to develop the instruments for authentic assessment. It is necessary to see if the assessment already meets the criteria of an authentic assessment, based on the requirement of the 2013 curriculum, related to content and authenticity. First, it is related to the requirement of the 2013 curriculum. It deals with the criteria that authentic assessment for speaking have to fulfill in line with the 2013 curriculum. Second, it is related to the requirement of content/types of assessment used in assessing speaking. The requirement of content/types discussed about the varied types of authentic assessment used in assessing speaking that in line with the topics material. Third, it is related to the requirement of authenticity. The authenticity is dealing with the characteristics needs to fulfil to claim an assessment as authentic.

First, related to the requirement of the 2013 curriculum, based on Permendikbud No. 23 on 2016 about education assessment standards. There are several principles of assessment. They are; (1) Authentic, it means that the assessment is based on data reflecting the measured ability. Assessment categorized as authentic when learners are asked to present a real task or situation demonstrating meaningful application of essential knowledge and skills. (2) Objectives, it will evaluate based on clear procedures and criteria (3) Fair, means that the assessment unprofitable judgment or disadvantage of learners due to special needs and differences in religious background, ethnicity, culture, customs, socioeconomic status, and gender. (4) Integrated, that assessment is one of the integral components of learning activities. An assessment conducted continuously during the learning takes place and after the end, through various types of test (daily test, semester midterm test and semester test). It is conducted to get a complete picture of the development of students' learning outcomes, monitoring the process, progress, and improvement of the results. (5) Daily review is done after completing one basic competency (KD) or more, integrated into the learning process in the form of repetition or assignment. (6) Transparent, the assessment procedure, criteria, and decision-making basis can be known by the stakeholders. (7) Universal and continuous, it means that the assessment covers all aspects of competence by using appropriate assessment techniques, to monitor and assess the development of learners' abilities. (8) Systematic, means the assessment is done in a planning and phases by following the standard steps. (9) Criteria referenced, means the judgments are based on the defined competence attained. And (10) Accountable, the judgment can be accountable, either from mechanism, procedure, technique, or outcome.

Second, related to the content of authentic assessment. In order to make sure that this product are suitable to the content or the textbook used in the learning process, it is necessary to adjust the products of speaking assessment to the textbook. Tomlinson (2003) proposes eleven criteria in conducting the textbook analysis. (1) Brainstorming a list of universal criteria. (2) Subdividing some of the criteria. (3) Monitoring and revising list of general criteria. (4) Categorizing the list, it is very useful to rearrange the random list of general criteria into categories which facilitate focus and enable generalization to be made. Possible categories for general criteria would be cultural perspective, learning principles, topic contents, teaching points, texts, activities, methodology, instructions and design. Then, (5) developing media specific criteria. This criteria is about asking questions which are specific relevance to the medium (book, cassette and video) used by the materials being evaluated. (6) Content specific criteria, this criteria related to teaching points and topics of the materials that will be analyzed. (7) Developing age specific criteria, this criteria focuses on age of target learners because they would like to use the textbook. The textbook provided should be suitable to the learners age, whether they are children or adults. This is also related to their cognitive, affective, their psychomotr, experience, needs and interest. (8) The next one is developing local criteria. It is related to the potential environment of use. It includes the syllabus, class size or type of institution, objective of the course, the target of examination, the number of exposure to the target language outside the classroom, and the intensity and extent of the teaching time available. (9) Developing other criteria, it can include cultural specific, administrator specific, teacher specific and gender specific. And (10) Trying the criteria, it must be

done because it is very important to ascertain those criteria are effective, efficient, answerable, reliable, sufficient, beneficial, and useful.

Third, related to the authenticity. In order to make a claim for authenticity of a test, it can be seen from the task is likely to be enacted in the real world. As Bachman (1990 : 303) states that authenticity as real life language use. It is defining authenticity essentially considers the extent to which the test performance replicates some specified non test language performance. Further, it is also defined authenticity as the degree of correspondence of the characteristic of a given language test task to the features of a target language task.

Wiggins in Koh (2017: 7) proposes the criteria to define the authenticity of an assessment. (1) The authentic assessment is realistic. It means that the assessment can be categorized into authentic assessment when they are appropriately public, involving an actual audience, client, panel, and so on. It also has to replicate how a students' knowledge, skills, and disposition are assessed in real world context. (2) The authentic assessment requires students to make a good judgement, be creative and innovative in solving problems. (3) The authentic assessment enables students to deeply engage in subject through critical thinking and inquiry. (4) In authentic assessment, teachers give chance for students to rehearse, practice and give feedback to improve their quality, the students have chances to present their work. (5) The authentic assessments look for multiple evidences of student performance over time and the reasons or explanations behind the success and failure of a performance. (6) The authentic assessment used multifaceted scoring system, and scoring criteria must be transparent. They must be scored with reference to authentic standards of assessment that have their own rubric score. The teachers need to share the scoring criteria with the students, in order they can understand and internalize the criteria of success. To ensure fairness and equity, the teacher must be provided with informative data of students' strengths and weaknesses at the end of each assessment. It will ensure that the teachers' feedback is aimed at helping all students to make progress toward the standards. (7) Student self-assessment plays a pivotal role. (8) The authentic assessment is defensible. The reliability or defensibility of teachers' professional judgment or scoring of student performance or work is achieved through social moderation, in which the teachers of the same subjects gather to set criteria and standards for scoring, and to compare their scores. This kind of assessment do not focus to judge on right or wrong answer, the aim tends to enable the students to show off what they can do. Authentic assessment balanced between honoring achievement, progress, native language skills and prior fortunate training.

There are several research results related to the developing instruments for evaluating the authentic assessment. Rahmawati and Ertin (2014) recommend the teacher to develop their speaking assessment which is appropriate and contextual. In order to develop assessment for speaking, it is important for the teacher to be familiar with the issues of practicality, validity, reliability, authenticity and washback effect. Shatrova et al (2017) found that English learners have their own specific needs, in case of group of learners such as students in preparatory language schools, their needs can be reflected in the number of tasks, topics and scoring rubrics. Moreover, Ferdiant (2016 : 93) notes that developing speaking test is not as easy as other tests because a test developer has to prepare the mechanism or direction and instruction well in order to keep the test valid in which the test developer used content validity to prove that the test was valid.

Based on the explanation above, this study aims to develop the instruments that will be used to evaluating the implementation of authentic assessment for speaking skill at junior high school.

Research Method

The type of research used in this research was evaluation research to evaluate the implementation of authentic assessment for speaking skill at junior high school. According to Gay and Airasian (2009; 17), "evaluation research is the systematic process of collecting and analyzing data about the quality, effectiveness, merit or value of programs, products, or practices". This research was conducted at Junior High School in West Sumatera. The evaluation instruments were developed by using theoretical based approach and were validated by the expert in language assessment and English language teachers at Junior High School. The data were collected through distributing the evaluation instruments for requirements of the 2013 curriculum, related to content and authenticity to the English language teachers and experts in language assessment. There were three documents analyzed in this research, they were format evaluation for requirement of the 2013 curriculum, format evaluation related to content and format evaluation related to authenticity.

Research Findings and Discussion

The findings discussed in this research are related to the instruments used in evaluating the authentic assessment in assessing students' speaking skill. The instruments used for evaluating the authentic assessment for speaking skill at junior high school was the format evaluation. There are three kinds of format evaluation used in evaluating the authentic assessment for speaking skill, they were the format evaluation for

the requirements of the 2013 curriculum, format evaluation related to content and format evaluation related to the authenticity. Based on the validation from the experts and English teachers, it can be concluded that the evaluation instruments are valid and can be used for evaluating the implementation of authentic assessment for speaking skill at junior high school.

The format evaluation for the requirements of the 2013 curriculum.

There are 9 points need to consider in fulfil the criteria of authentic assessment based on the requirement of the 2013 curriculum as stated in Permendikbud No. 23 on 2016 about education assessment standards. The format evaluation for the requirement of the 2013 curriculum can be seen in the table.1.

Table 1 . The format evaluation for the requirement of the 2013 curriculum.

No	Indicators	Sub-indicators
1	Authentic	a. The authentic assessment assesses what it is meant to be assessed.
2	Objectives	a. The authentic assessment have a clear and standardized procedure. b. The authentic assessment have clear criteria of what is being assessed.
3	Fair	a. The authentic assessment consider student with disability b. The authentic assessment consider students' background (<i>religion, ethnic, culture, custom, economy status and gender</i>).
4	Integrated	a. The authentic assessment include what students have in the teaching and learning process.
5	Transparent	a. The evaluation process of the authentic assessment is acknowledge by parties involved. b. The criteria what being assessed in the authentic assessment acknowledge by parties involved. c. The principle in making judgement of the authentic assessment acknowledge by parties involved.
6	Universal and Continuous	a. The authentic assessment cover all the competencies. b. The authentic assessment use various techniques to monitor and assess students' progress.
7	Systematic	a. The authentic assessment planned by systematic rules. b. The authentic assessment do in phases of sytematic rules.
8	Criteria referenced	a. The authentic assessment based on the defined competency achieved.
9	Accountable	a. The authentic assessment can be guaranteed by mechanism. b. The authentic assessment can be guaranteed by procedure. c. The authentic assessment can be guaranteed by technique. d. The authentic assessment can be guaranteed by results.

The format evaluation related to content.

In order to make sure that the authentic assessment are suitable to the content or the textbook used in the learning process, it is necessary to adjust the products of speaking assessment to the textbook. Tomlinson (2003) proposes eleven criteria in conducting the textbook analysis. The format evaluation related to the content can be seen in the table.2.

Table 2. The format evaluation related to the content.

No	Indicators	Sub – indicators
1.	The content of assessment is suitable with the text book objective	a. The content is suitable with the text book objective
2.	The authenticity and accurancy of the assessment	a. Quote of material source from TV or radio b. Using magazines, newspapers or books as the source of materials

3.	The assessment should have tight relationship with students' social condition	a. Relate to the students' social condition
4.	The materials should help students to develop confidence	a. Help students in develop their confidence
5.	The materials should provide the learners with opportunities to use the target language to achieve communicative purpose	a. Give the students opportunities to communicate in English
6.	The materials should be in variative	a. Use variative materials
7.	Assessment materials should help learners to feel at ease	a. The materials helps students in the learning process
8.	Assessment materials should match with the syllabuls provided	a. The materials used suitable to the syllabus

The format evaluation related to the authenticity

Further, Wiggins in Koh (2017: 7) proposes the criteria to define the authenticity of an assessment. The format evaluation related to the authenticity can be seen in the table 3.

Table 3. The format evaluation related to the authenticity

No	Indicators	Sub-indicators
1	Realistic	a. The authentic assessment provide a real world tasks.
2	Universal	a. The authentic assessment assesses students' knowledge. b. The authentic assessment assesses students' attitude. c. The authentic assessment assesses students' skills.
3	Engaged	a. The authentic assessment engaged students to have critical thinking and inquiry to their subject or discipline.
4	Applicative	a. The authentic assessment require students to investigate their work. b. The authentic assessment require students to present their work
5	Various	a. The authentic assessment provides variation tasks
6	Transparent	a. The authentic assessment provides information of scoring system. b. The authentic assessment provides information of giving feedback .
7	Students' centered	a. The authentic assessment supports students' centered activity.
8	Trustworthy	a. The authentic assessment provides a standardized scoring.

Conclusion

Based on all of data analyze about the instruments for evaluating the authentic assessment for speaking skill at junior high school, researcher developed the format evaluation instruments which is reliable, valid and usable. There were three types of the format evaluation used in evaluating the authentic assessment for speaking skill at junior high school. They were :

1. The format evaluation for the requirement of the 2013 curriculum
2. The format evaluation related to the content
3. The format evaluation related to the authenticity

Based on the validation from the experts and English teachers, it can be concluded that the third of evaluation instruments are valid and can be used for evaluating the implementation of authentic assessment for speaking skill at junior high school.

References

Bachman, L. F. 1990. *Fundamental considerations in language testing*. Oxford: Oxford University Press.
 Gay, L.R, Geoffrey E.Mills & Peter Airasian. 2009. *Educational Research Competencies for Analysis and Application*, New Jersey : Prentice Hall.
 Ferdiant, Ghufran. 2016. Developing The Assessment Instruments of Speaking. *OKARA Journal of Languages and Literature*, Vol.1 Tahun 1.

- Fitriani. 2014. The Challenges in Implementing Authentic Assessment In Curriculum 2013. *The 61 TEFLIN Internasional Confrence*, UNS.Solo.
- Gulikers et al. 2004. A Five-Dimensional Framework for Authentic Assessment. *Educational Technology Research and Development*, Vol. 52, No. 3.
- Jaedun, Haryanto & Nuryadin. 2014. An Evaluation of The Implementation of Curriculum 2013 at The Building Construction Department of Vocational High Schools in Yogyakarta. *Journal of education*.
- Koh.H.Kim. 2017. Authentic Assessment. *Oxford Research Encyclopedia of Education*. Online Publication Date: Feb 2017 DOI: 10.1093/acrefore/9780190264093.013.22
- Rahmawati Yenny & Ertin. 2014. Developing Assessment for Speaking. *IJEE*. Vol. 1. No. 2.
- Shatrova et al. 2017. English speaking assessment: developing speaking test for students in a preparatory school. *International journal of english language teaching*. www.eajournals.org
- Peraturan Menteri Pendidikan dan Kebudayaan Nomor 23 Tahun 2016 Tentang Standar Penilaian Pendidikan. Jakarta : Departemen Pendidikan dan Kebudayaan.
- Tomlinson,B. (ed). 2003. *Developing materials for language teaching*. London: continuum.