

Sundanese Speaking Test Model

U Kuswari¹, H S Nugraha²

{¹usep.kuswari@upi.edu, ²harissantosa89@upi.edu }

^{1,2} Universitas Pendidikan Indonesia Bandung, Indonesia

Abstract. This research focuses on the quality of speaking testing instruments for junior secondary students in West Bandung District. In general, the study aims at creating a model of Sundanese language speaking skills assessment which can provide information for schools in terms of contents, scope, delivery format and time, and can be effectively use Sundanese language teaching. Research and Development method is adopted to produce a model of class-based Indonesian language and literature learning assessment at the junior secondary level and a guide book of Sundanese language and literature assessment at this level. Statistical calculation of field tryout scores shows that speaking model is relevant to field data, namely, (P-value = 0, 72541 > 0,05; RMSEA = 0,000 < 0,05; GFI = 1,70 > 0,9; AGFI = 1,43 > 0,9; PGFI = 1,37 > 0,9) indicating that EPBI model is fit. The relevance of EPBI model and field data (P-value = 0, 72541 > 0,05; RMSEA = 0,000 < 0,05; GFI = 1,70 > 0,9; AGFI = 1,43 > 0,9; PGFI = 1,37 > 0,9) reveals that the quality of speaking testing model is fit is worth considering as an assessing instrument of junior secondary school speaking instruction as supported by theoretical concepts, expert judgment, and field tryout results., and judgment by experts, users, and practitioners shows that MPPBS model is reliable as a referred implementation model. Finally, school principals, Sundanese language teachers, schools, and District Office are recommended to apply the MPPBS Model in assessing the instructional processes of Sundanese language skills in West Bandung District.

Keywords: *Test, Integrated, Communicative*

1. INTRODUCTION

In an education system, assessment and evaluation are interrelated. The link lies in the measurement activities of students' competence conducted using test or non-test in accordance with their grades or levels of ability and development that result from the learning process that the students have experienced. Students' learning process of the Sundanese language can be evaluated based on the measurement and assessment of competence and language performance data. The evaluation can be performed formally or informally. All tests are formal evaluations, but not all evaluations are formal tests. Evaluation can be defined as a collection of information about the quality and quantity of changes that result from students' learning [1]. State that learning evaluation has the aims of (1) determining the level of students' knowledge and skills, (2) understanding the progress towards the learning objectives to help prepare an instructional program, and (3) providing data to consider the final level of the students' learning [2].

State that the quality of language learning evaluation is determined by the clarity of its objectives so that it enables the evaluator to get a genuine response from the students [3]. The choice and use of a suitable evaluation approach are particularly useful to learn what kind of language competence and performance to be used as evaluation criteria. The management of evaluation instrument types, i.e. test and non-test, including the recording of the test results and their presentation in numbers and fidelity in percentage; and the use of test results are necessary in the development of an evaluation instrument.

Input should also become a consideration in test administration. Input involves two aspects, namely (1) format and (2) the nature of language. Input format includes channel and presentation form and presentation language, problem identification and speed level. Input can be presented aurally or visually in receptive form, whereas the answer can be oral or written in productive form or mode.

Realities in the field show that Sundanese language tests have been developed based on discrete approach putting forward the linguistic and behavioristic theory, whereas in any curriculum, tests are expected to be able to measure language competence and performance for real life needs. Teacher-made tests are tests that put heavy emphasis on theoretically linguistic competence only.

Sundanese language teachers' lack of knowledge in speaking test management has resulted in the Sundanese language learning becoming less effective so that many SMP (junior secondary school) graduates are unable to communicate effectively in Sundanese. This can become a problem when they want to continue their study or find a decent job. Therefore, the researcher feels that it is necessary to discover the obstacles faced by the Sundanese language teachers. This has motivated him to develop a Sundanese language learning evaluation model despite Wilson stating that objective measurement has rarely used a model in the Sundanese language [4]. The researcher hopes that this Sundanese language learning evaluation model can be developed further into a Sundanese language learning model.

In general, this study aims to produce a Sundanese language speaking test instrument whose content, scope, format and time delivery can inform the school and can optimally benefit Sundanese language learning at junior secondary school level in West Bandung District.

Specifically, this study aims to describe:

- a. the Sundanese language speaking tests that are currently used in junior secondary schools in West Bandung District;
- b. the quality of integrative approach-based Sundanese language speaking test instrument that is being developed for junior secondary school students in West Bandung District according to expert judgement;
- c. the quality of integrative approach-based Sundanese language speaking test instrument that is being developed for junior secondary school students in West Bandung District based on tryout results; and
- d. the benefit of integrative approach-based Sundanese language speaking test instrument that is being developed for junior secondary schools in West Bandung District area.

Evaluation is a systematic process of determining the extent to which educational objectives are achieved [5]. This definition has two notions, namely (a) evaluation is systematic process, which means it consists of a series of activities that are conducted through and based on certain regulations; and (b) evaluation is always linked to a set of learning objectives. These objectives become the guidelines for the creation of an evaluation instrument [6].

Furthermore, evaluation can be interpreted as a systematic assessment of the value or appropriateness of several objects [6]. This idea is a systematic process for determining the level of objective achievement of learning material received by students [5] who in the process collects and interprets information about learning achievement to provide an even broader value than that [7] [8].

The term 'evaluation' is often misused especially in the context of learning. Sometimes 'evaluation' is equated with 'measurement' or used interchangeably with 'testing'. When a teacher manages a learning outcome test, he can use the terms 'achievement test', 'achievement measurement', or 'performance evaluation'. In other cases, evaluation is also used as a research method that is independent of measurement [9].

Test can be regarded as an instrument of evaluation. However, test can also be regarded as a technique of measurement. As Brown says, "A test will be defined as a systematic procedure for measuring a sample of an individual's behavior" [6]. Based on this definition, test has two underlying notions about it, namely it is a systematic procedure and it measures a particular sample behavior. The term systematic procedure means that a test has to be designed, administered and processed according to certain established rules. A test also has to be systematic, which means (a) its contents has to be designed and chosen systematically in accordance with the area or scope of behavior that will and have to be measured or tested so that the test's validity is really accountable, (b) its administration has to follow certain established procedures and conditions, and (c) the data that the test produces have to be processed and interpreted according to certain rules and norms. The phrase 'measuring a sample of an individual's behavior' means that the test measures only a sample of an individual's behavior that is being tested. A test cannot measure the entire behavior (the population) being tested, but only the content (test item) of the test.

A test contains questions and or problems that need to be answered and or solved by the individual being tested (teste). Such a test is called achievement test. This is in line with experts' opinion which says that "The type of ability test that describes what a person has learned to do is called an achievement test" [10].

This approach is based on the linguistic theory of structuralism which views language as a set of discrete elements that are organized by means of certain structures. This approach is also supported by a psychometric view that makes it possible for each element or language unit to be assessed separately.

This approach has several advantages, namely (a) it is easier to quantify as assessment can be made on smaller (atomistic) and discrete elements; (b) test items can be significantly extended or cover various language levels and skills; and (c) it is more efficient (in terms of cost) and easier to administer [11]. Additionally, this approach also produces more quantifiable data that cover a wide array of materials because what is being assessed is language elements, such as phonology, morphology, syntax, and vocabulary. Discrete approach is efficient and generally more reliable as an achievement test, but both the approach and the format have weaknesses to measure.

Despite its advantages, psychometric-structuralism approach also has its weaknesses. Among these weaknesses are that (a) discrete testing does not usually pay much attention to the interaction among language elements in the wider and more complex communication contexts; (b) this assessment becomes less affective because those important parts of language is missing when discretely analyzed; (c) grammatical competence is not a good predictor for communication skills, and (d) this approach is artificial, sterile, and irrelevant because it does not evaluate language performance in a real situation and context.

Maintained that an integrated assessment presents all language skills better than a combination or single test [12]. Cloze and dictation become very integrated because they integrate most or all language skills. There is a high correlation between cloze and other assessments.

Pointed out the fact that cloze as well as dictation offers the opportunity for a spontaneous production by the applicant and the language norms used are those of the examiner's, not of the student's themselves [13].

Communicative approach that Wier put forward is based on communicative competence. This competence can only take place in daily life activities. Therefore, assessment of language skills in communicative approach measures capacity [14], communicative language ability [15], and the evaluation of samples of performance in certain specific contexts of used created under particular test constraints [11].

Clarification of the terms used that are relevant to the formation of communicative competence. According to them, communicative competence includes grammatical competence (knowledge of grammatical rules), sociolinguistic competence (knowledge about rules of uses and rules of discourse) and strategic competence (knowledge about communicative strategies). This model was later modified by Canale, who proposed a four-dimensional model that consist of linguistic, sociolinguistic, discourse and strategic competence [16].

The learner's primary needs are not those of analytical or theoretical knowledge about the target language, but the need to understand and be understood in a language in context [17].

Speaking test is not only an oral test. It can also be a performance test, which is a non-verbal test. This means that speaking test does not only evaluate the act of speaking, but also the processes that lead to speaking. Therefore, observation technique is often considered to be part of speaking test technique. An examiner needs to use this technique to observe (not just listen) how a teste speaks. This is particularly applicable in a direct oral performance testing.

To elicit students' ability to speak Sundanese, pictures can be used as stimuli. Pictures are good stimuli for speaking skills evaluation for primary school pupils. However, picture stimuli can also be used with students with higher level language skills, depending on the pictures used. Good pictures are pictures that draw students' interest to speak or those that make it easier to reveal their speaking skills. The tasks can be given to the students in the form of questions or story telling task [18].

In the perspective of science, this research is an applied research. This research specifically deals with language competence (speaking and writing) in language teaching and learning. This research uses various theories of speaking and writing tests and how they relate to decision making about evaluating SMP students' competence in Sundanese language.

2. RESEARCH METHOD

The subjects of this trial study are the Sundanese language teachers of class VII, VIII, and IX of Junior Secondary Schools (SMP) and Sundanese language learning experts. The location and subjects of this study were chosen because the researcher has lived long enough in this district and is therefore very familiar with it. This has helped him in the data collection process. He has received a lot of assistance in conducting the tryout for the Sundanese language speaking and writing test instruments that were being developed through this development research. The subjects of this tryout study were derived from 3 SMPs and 20 state SMPs in West Bandung District of West Java Province.

The Research and Development method is used for this study, which aims to produce an integrated approach-based Sundanese language speaking and writing test instrument for SMP level. The education and development study are a type of research that is widely used to solve practical problems in the field of education. Education research and development is a process used to develop and validate an educational product, in this case the Sundanese language speaking test instrument [19].

The procedures being followed in this development research combine the procedures proposed by Borg & Gall with Kirkpatrick model of development procedures through four stages, namely: (1) preliminary stage; (2) design stage; (3) tryout and revision stage; and (4) implementation stage.

Quantitative data analysis with Confirmatory Factor Analysis (CFA) was used to figure out the validity of the Sundanese language speaking test instrument. The analysis was performed with the help of LISREL. CFA is used to examine the existing construct validity [20].

Qualitative data analysis is the analysis of validation result data received from experts (expert judgement) and users of the Sundanese language speaking and writing test instrument who have given input to improve the evaluation model and its instruments. Analysis was performed on the construct of the test instrument, namely the manual, the material, the language, the type of test instrument used, writing procedure, and scoring guidelines. Evaluation research aimed at finding information to understand a successful intervention mechanism [21]. The researcher should discuss the reasons for combining the techniques used. In this qualitative data analysis, the quantitative data obtained from the evaluation instrument were converted to qualitative data on the scale of 5, then described and the results of the description were made as the basis for evaluating the quality of the evaluation model that was being developed [22].

3. RESULT AND DISCUSSION

The validity test for the speaking test instrument was conducted three times: first tryout, second tryout, and third tryout (implementation). In each tryout, respondents were asked to give comments on the speaking test instrument.

The results of the tryouts were analyzed with *confirmatory factor analysis* (CFA) program LISREL 8.51. Item validity was based on the factor load value (λ) of each instrument's item, whereas the measurement model conformity was analyzed based on the significance value (*P-Value*) and *Root Mean Square Error of Approximation* (RMESA). During the implementation stage, the analysis of the measurement model conformity was based not only on the score of each instrument's item but also analyzed with composite score. The composite score was obtained from the sum of indicator score and the sub-aspect of each speaking and writing test aspect of the Sundanese language learning.

The number of model items analyzed in the Sundanese language speaking test was 23. The aspects analyzed include: (1) manual: clarity of scoring manual, (2) material: clarity of indicator, clarity of relationship between item type and indicator, relationship clarity between the instrument and the material being tested, and the clarity of evaluation material and testing instrument used; (3) language: use of standard language, definition of communicative language, use of easy-to-understand sentences and words; (4) writing procedures: font, font size, format or layout; and (5) grading guidelines of speaking test instrument.

Two types of analysis were performed on the Indonesian language test instrument, namely qualitative analysis and quantitative analysis. Qualitative analysis is the analysis of speaking

test according to expert judgment and user judgment, whereas quantitative analysis is performed by means of statistical method. Results of the analyses are described below.

The test manual for 23 test items of the Indonesian language speaking test was found to be the level of fair. This means that the existing manual needs to be improved. The directions need to be made in easy and quick to understand sentences, not too many, and readable. This is in line with the evaluation principle which considers that test or item directions are an important part of test development. Students will be able to do the test quickly if the directions are prepared and made well. No matter how good a test is, if the directions are unclear, the test will be of no good. Test directions will give directions to the students about what to do, how much time is given, the requirements that need to be fulfilled in doing the test, and what important thing the test takers need to know.

There are several important principles that need to be followed to make the test more effective. The first of these principles is certainty and clarity of what is being tested. A speaking test cannot be administered if the test objectives are not clearly and operationally defined. In order to evaluate students' learning progress, first we need to identify the learning objectives. Only then can we develop the evaluation instruments. Tests are the most common instrument of evaluation in education. A test should reflect the characteristics of the aspect to be measured. If we want to evaluate students' level of intelligence, then the components of intelligence and the learning achievement have to be clearly defined before a test is developed as an instrument of evaluation.

All the 23 items of the Indonesian language speaking test that were being developed were of fair quality. This means that the formulated indicators still needed to be modified. The indicator improvement that still needed to be made include the use of operational verbs, where some verbs were found to be less measurable and not an urgent indicator. For example in speaking test item 1, "*Mampu mendafta pokok-pokok cerita pengalaman yang mengesankan* (Able to register the gist of an impressive experiential story)". This sentence was not a relevant indicator of the basic competence "*Mampu menceritakan pengalaman yang paling mengesankan dengan menggunakan pilihan kata dan kalimat yang efektif* (Able to recount/tell the most impressive experience using effective diction and sentences)". The most appropriate indicator that reflects the basic competence is "*Mampu menceritakan pengalaman yang paling mengesankan berdasarkan pokok-pokok rangkaian cerita dengan menggunakan pilihan kata yang tepat dan kalimat efektif.*" (Able to recount/tell the most impressive experience based on the main points of the story using effective diction and sentences.) The main or *the most important indicator that should be included in the speaking test is, therefore, "Mampu menceritakan pengalaman yang paling mengesankan berdasarkan pokok-pokok rangkaian cerita dengan menggunakan pilihan kata yang tepat dan kalimat efektif"* (Able to recount/tell the most impressive experience based on the main points of the story using effective diction and sentences). If this indicator can be achieved well by students, the indicator "Able to recount/tell the most impressive experience based on the main points of the story using effective diction and sentences" is measurable and, therefore, no longer need to be tested.

The material tested was explained on the basis of compatibility analysis with the basic competence and indicator, material adequacy, being essential, real or realistic, and contextually appropriate in the use of Indonesian. The material being tested is reflected in the basic competence and indicator. An example from Speaking Test Instrument 1: The subject matter is "Story Telling" of the basic competence "Telling a story in a good order and appropriate voice, pronunciation, intonation, gesture, and expression". The test material that is relevant to the basic competence is therefore the aspects of order, voice, pronunciation, intonation, gesture and expression.

The scope of Indonesian speaking test material is reflected in the theoretical definition of speaking skill. Speaking skill is a complex linguistic skill which covers not only the question of pronunciation and intonation, but also the use of diction and other linguistic and non-linguistic elements. The aspects being evaluated in a speaking activity consists of linguistic and non-linguistic aspects. The linguistic aspects include: pronunciation, word stress, tone and rhythm, vocabulary or expression, sentence or sentence structure variations. The non-linguistic aspects include: fluency, material mastery, courage, friendliness, discipline, spirit and attitude.

Test material is linked to its context; therefore, context plays an important role in communicative language skills and in the integrative approach to assessment vis a vis decontextualize approach. Without (linguistic, discourse and sociocultural) context, language cannot have a full meaning. Higher level of language is a contextualized perception of language [18]. Achievement variability corresponds to the type of task involved, implied in the length of the test and type of text and adequacy of the format [23], [24].

In estimating the ability to interact orally, we need to try to reflect on the interactive nature of a normal conversation and try to ensure that reciprocity is considered in the test task involved. The task has to be given under normal time and the unpredictable element of interaction should be identified in order that a genuine communication with other participants may take place.

A speaking test instrument should be considered from three points, namely type of evaluation, type of evaluation instrument, and format of evaluation instrument. Types of evaluation consist of process evaluation and product evaluation; types of evaluation instrument can be test and non-test. The format of evaluation instrument are oral format, written format, performance, project, portfolio, attitude (character) assessment, self-assessment, and observation. An example from speaking test instrument 4: the subject matter is “Delivering a story with props”, the basic competence “Story telling with props”, and indicator “Able to tell a story with props based on story points”. A non-test process evaluation in the format of performance and observation is used to evaluate the material, basic competence, and indicator. These types of evaluation, instrument and format are appropriate to evaluate the basic competence “Story Telling with Props” because speaking test instrument 4 has a logical connection with the indicator, basic competence and test material.

The advantage of speaking test format should take into account (a) the test size – length of the text, productivity, involvement; (b) grammatical complexity and scope of the required coherence devices; (c) functional scope – illocutionary variation levels involved; and (d) reference scope – the depth of knowledge relevant to the dictionary needed to handle the activity.

The language of a speaking test instrument has to have a standard grammar system in accordance with the standard grammatical rules. In addition to being grammatical, a speaking test also has to be written in sentences that are easy to understand or, in other words, readable, and communicative. Because a speaking test uses command and order sentences, it is not too difficult to make the sentences. Here is an example from test item 21: *Berpidatolah dengan intonasi yang tepat serta artikulasi dan volume suara yang jelas!* (Make a speech with appropriate intonation and clear articulation and volume!) This test instrument sentence has covered the scope and aspects that students need to pay attention to.

A good speaking test instrument does not only have standardized language, it is also developed in a standardized format and writing system. The most dominant of this writing system is the spelling, such as the use of letters, the writing of words, and the use of punctuations and other written symbols. Because speaking test uses mostly command and

order in its instruction, the writing system and layout are relatively simple and easy, except when the test uses picture or sound stimulus [25].

A rubric is a description of achievement criteria for a teste. Therefor a rubric and its description must be clearly written so that students or their parents will know how much achievement their children have made in their learning. The rubrics provided in test items 1 until 23 are clear and easy to use that even students can use them on their own to evaluate their peers.

4. CONCLUSIONS

The twenty-three items of the Sundanese language speaking test are fit for use as a speaking skill evaluation instrument at Junior Secondary School (SMP) level. It is because (a) they conform to the theoretical construct that gives rise to the integrative test, i.e. a test that integrates linguistic competence with non-linguistic competence or competence and performance; (b) according to expert judgement, these test instruments have clear guidelines for scoring, clear relationship between the subject matter being evaluated and the evaluation instrument; (c) according to the tryout results, these speaking test instruments have been statistically shown to be consistent with the data found on ground.

REFERENCES

- [1] H. D. Brown, *Language assessment: Principle and classroom practices*. New York: Longman, Pearson Education, Inc., 2004.
- [2] D. W. Johnson and R. T. Johnson, *No Title Meaningful assessment: A manageable and cooperative process*. Boston: Allyn and Bacon, 2002.
- [3] J. M. O'Malley and L. V. Pierce, *No Title authentic assessment for English language learners: Practical approaches for teachers*. New York: Addison Wesley Longman, Inc., 1996.
- [4] M. Wilson, *Objective measurement: Theory into practice (vol.2)*. Norwood: Alex Publishing Corporation, 1992.
- [5] N. E. Grondlund, *Measurement and Evaluation in Teaching*. New York: Macmillan Publishing Comp, 1995.
- [6] U. Kuswari, *Evaluasi Pembelajaran Bahasa*. Bandung: JPBD FPBS UPI, 2010.
- [7] B. K. Lynch, *Language program evaluation: Theory and practice*. New York: Macmillan publishing, 1996.
- [8] D. Nunan, *Research methods in language learning*. New York: Cambridge, 1992.
- [9] A. R. A. Ghani, S. Hari, and S. (Ed), *Evaluasi pendidikan: Konsep dan aplikasi*. Jakarta: UHAMKA Press, 2006.
- [10] R. L. Thorndike and E. Hagen E, *Measurement and Evaluation in Psychology and Education*. New York: John Wiley and Sons Ltd.
- [11] S. Yusuf, "Pengembangan Model Ujian Nasional Berdasarkan Pendekatan Literasi: Kajian Struktur Soal dan Daya Serap Siswa SMP.MTs dan SMA/MA di Jawa Barat dalam Mata Pelajaran Bahasa Sundapada Ujian Akhir Nasional Tahun 2004," UPI, 2006.
- [12] A. Davies, "The construction of language tests," Allen, J.P.B. A. Davies, pp. 38–104, 1977.
- [13] K. E. Morrow, *Techniques of evaluation for a national syllabus*. London: Royal Society of Arts., 1977.

- [14] H. . Widdowson, "Teaching language as communication," Oxford: Oxford University Press, 1983.
- [15] L. F. Bachman, *Fundamental considerations in language testing*. Hong Kong: Oxford University Press, 1990.
- [16] M. Canale and M. Swain, "Theoretical basis of communicative approaches to second language teaching and testing," *Appl. Linguist.*, vol. 1, pp. 1–47, 1980.
- [17] B. J. Carroll, *Testing communicative performance: an interim study*. Oxford: Pergamon, 1980.
- [18] J. W. Oller, *Language tests at schools*. London: Longman, 1979.
- [19] W. R. Borg and M. . Gall, *Educational research: An introduction*. (4th ed). New York & London: Longman, 1983.
- [20] R. O. Mueller, *Basic principles of structural equation modeling: An introduction to lisrel and EQS*. New York: Springer, 1996.
- [21] J. M. Morse, *Critical issues in qualitative research methods*. London: Sage Publications, 1994.
- [22] Suhuri, "Model Evaluasi Pembelajaran Bahasa Inggris di SMA Palembang Tahun 2009," UNY, 2009.
- [23] D. Douglas and L. Selinker, "Principles for language tests within the 'discourse domains' theory of interlanguage: research, test construction and interpretation," *Lang. Test.*, vol. 2, no. 2, pp. 205–226, 1985.
- [24] P. Skehan, "Variability and language testing," in *Second language acquisition in context*, Englewood Cliffs, R. E. (ed.), Ed. NJ: Prentice Hall, 1987.
- [25] L.E. Rahmawati, S. Suwandi, K. Saddhono, B. Setiawan. 2019. "Need analysis on the development of writing competency test for foreign university students," *Humanities and Social Sciences Reviews* vol. 7 no. , pp. 467-471. 2019