

Vol. VI No. 2 (Desember 2008)

ISSN: 1693-7996

STUDIA AKADEMIKA

JURNAL PENELITIAN ILMU-ILMU SOSIAL

THE ASSESSMENT OF STUDENTS' LISTENING,
SPEAKING, READING, AND WRITING
ACHIEVEMENTS BY THE ENGLISH TEACHERS
OF MADRASAH ALIYAH NEGERI 2 BATUSANGKAR

Suswati Hendriani

VALIDATING THE RESULT OF A TEACHER-MADE
TEST IN ASSESSING ORAL LANGUAGE
PROFICIENCY

Suyono

IMPROVING STUDENTS' CRITICAL LITERACY IN
READING SHORT STORIES THROUGH QUESTIONS

Yulnetri

STUDIA AKADEMIKA	Vol. VI	No. 2	Halaman 109-206	Batusangkar Desember 2008	ISSN 1693-7996
---------------------	---------	-------	--------------------	------------------------------	-------------------

**SEKOLAH TINGGI AGAMA ISLAM NEGERI (STAIN)
BATUSANGKAR**

2008

STUDIA AKADEMIKA

JURNAL PENELITIAN ILMU-ILMU SOSIAL

**THE ASSESSMENT OF STUDENTS' LISTENING,
SPEAKING, READING, AND WRITING
ACHIEVEMENTS BY THE ENGLISH TEACHERS
OF MADRASAH ALIYAH NEGERI 2 BATUSANGKAR**

Suswati Hendriani

**VALIDATING THE RESULT OF A TEACHER-MADE
TEST IN ASSESSING ORAL LANGUAGE
PROFICIENCY**

Suyono

**IMPROVING STUDENTS' CRITICAL LITERACY IN
READING SHORT STORIES THROUGH QUESTIONS**

Yulnetri

STUDIA AKADEMIKA	Vol. VI	No. 2	Halaman 109-206	Batusangkar Desember 2008	ISSN 1693-7996
-----------------------------	----------------	--------------	----------------------------	--------------------------------------	---------------------------

**SEKOLAH TINGGI AGAMA ISLAM NEGERI (STAIN)
BATUSANGKAR
2008**

STUDIA AKADEMIKA

JURNAL PENELITIAN ILMU-ILMU SOSIAL

ISSN: 1693-7996

Vol. VI No. 2 (Desember 2008)

❖ Ketua Penyunting
Suswati Hendriani

❖ Wakil Ketua Penyunting
Arwadi

❖ Penyunting Ahli
Helmi

Anas Yasin

Suryadi

Rita Erlinda

Syukri Iska

Muhammad Fazis

❖ Penyunting Pelaksana

Yusri Akhimmuddin

Irwan

Ulya Atsani

Suharmon

❖ Pelaksana Tata Usaha

Chardinal Putra

Marhen

Alamat Redaksi dan Tata Usaha:

Jl. Sudirman No. 137 Kuburajo Lima Kaum Batasangkar ☎ 27213 📠 (0752) 71150
e-mail: jurnal@stainbatasangkar.ac.id

Studia Akademika: Jurnal Penelitian Ilmu-ilmu Sosial, terbit pertama kali tahun 2003, berdasarkan SK Ketua STAIN Batasangkar Nomor ST/2/HM.02.1/896/2003. Jurnal ini selanjutnya terbit dua kali setahun dan dibiayai dengan dana Proyek Peningkatan Perguruan Tinggi Agama/STAIN Batasangkar. Jurnal Penelitian ini secara khusus hanya memuat tulisan berupa hasil penelitian tentang ilmu-ilmu sosial (*social sciences*). Penyunting menerima kontribusi tulisan dari berbagai kalangan dan berhak menyunting tulisan sepertiunya tanpa merubah maksud. Pedoman penulisan dan transliterasi dapat dilihat pada halaman akhir. Dilarang mengutip atau memperbanyak tulisan dalam jurnal ini tanpa izin dari redaksi atau minimal menyebutkan sumber aslinya. Hak Cipta dilindungi undang-undang.

STUDIA AKADEMIKA

JURNAL PENELITIAN ILMU-ILMU SOSIAL

DAFTAR ISI

- ≠ **Suswati Hendriani**
the Assessment of Students' Listening, Speaking, Reading, and Writing Achievements by the English Teachers of Madrasah Aliyah Negeri 2 Batusangkar..... 109 - 129
- ≠ **Suyono**
Validating The Result of A Teacher-Made Test in Assessing Oral Language Proficiency..... 130-140
- ≠ **Yulnetri**
Improving Students' Critical Literacy in Reading Short Stories Through Questions..... 141-162
- ≠ **Rafsel Tas'adi**
Konseling di Sekolah Studi Deskriptif pada SMAN dan MAN di Kabupaten Tanah Datar..... 163 - 187
- ≠ **Radiatul Hasnah**
Validitas Hadis-Hadis dalam Buku Kebenaran Al-Quran dan Hadis Jilid II Madrasah Aliyah Berdasarkan Kurikulum Berbasis Kompetensi Tahun 2004..... 66 - 86

THE ASSESSMENT OF STUDENTS' LISTENING, SPEAKING, READING, AND WRITING ACHIEVEMENTS BY THE ENGLISH TEACHERS OF MADRASAH ALIYAH NEGERI 2 BATUSANGKAR

By: Suswati Hendriani*

Abstract

The aims of the research are to find out the techniques used by the English teachers of MAN 2 Batusangkar in assessing their students' listening, speaking, reading, and speaking achievements. The population of the research was 7 English teachers of MAN 2 Batusangkar. The sample was taken by applying purposive random sampling technique. As a result, four English teachers were selected. The data of the research were collected by using questionnaire, interview, and documents. The research results showed that to assess the student's listening, speaking, reading, and writing ability, the respondents used tests and non tests

Key words: assessment, listening, speaking, reading, writing, achievement

A. INTRODUCTION

Assessing students' achievements is not easy since from the result of the assessment, many decisions and conclusions will be made. A tiny mistake in doing an assessment may lead to wrong conclusions and wrong decisions. Therefore, the assessment should be done carefully, continuously, comprehensively, and accurately so that the result of the assessment will be reliable. To assess students' achievements is usually concerned with the process of getting information about their progresses after experiencing learning process at a given time. In other words, the result of the assessment should show students' true ability. For English learning, especially, the assessment of students' language achievements should be directed toward their true ability in using the language. Therefore, a teacher as the evaluator should try to notice every aspect of the students' language learning, and their mastery of the language components and skills. Thus, their true ability will be truly reflected.

To do this, however, good instruments for the assessment are needed. The instruments used by the teacher to reflect the students' true ability in using the language should be valid and reliable. If not, what is expected from the assessment cannot come true.

* Penulis adalah Lektor Kepala dalam Mata Kuliah Bahasa Inggris pada STAIN Batusangkar

Therefore, the teacher should try to find appropriate instruments which can help them to find out their students' true achievements. Thus, what they indicate will really be the true conditions of the students.

To assess the students' achievements, there are some instruments that can be used, either in the forms of tests or non-tests. When using a test to assess the students' achievements, there are three criteria that should be met. They are validity, reliability, and practicality. If the three criteria cannot be met, the students' achievements will be questionable. Moreover, using non-tests to evaluate their achievements needs some criteria, too, to be met. They are whether the instruments used are well prepared or not, consider each aspect of the students' learning or not, have met the criteria previously set or not, etc.

Ideally, every teacher, including an English teacher, should try to master and combine the two forms of instruments to dig the students' true ability. Using test as the only instruments may not be adequate since the tests tend to be product-oriented. They cannot show and dig the processes the students have already experienced and the result might be strongly influenced by the situation and condition while the students are having the tests, such as, by their sickness and/or lack of concentration. Therefore, other kinds of instruments which can show the students' progresses are needed. Related to the Curriculum being implemented nowadays, the assessments should be done by applying classroom-based assessment, which involves a written/paper test, a performance test, assignments, project evaluation, product evaluation, attitude assessment, and portfolio assessment.

Unfortunately, in doing assessment of students' learning achievements nowadays, some teachers tend to use specific tests, such as in the form of Multiple Choice, True/False, Cloze Procedure, Essay, and so on. Whatever the types the tests used, they should be first made valid, reliable, and, if it is possible, practical, too. A valid test will test what is supposed to be tested. A reliable test will produce consistency in the students' achievements. A practical test will be easily administered. Yet, At schools often disobey these criteria. They tend to do assessment which focuses on materials and forgets the language skills. Even though in the curriculum it is stated that there are four language skills that the students should master, some teachers tend to test just certain skills. It means the achievements of the students cannot yet be well reflected. The techniques the teachers use do not show the students true ability, yet. The results of the students doing tests cannot show their true abilities in using English in real life situations. Conver-

struments
achievements.
itions of the

some instru-
s or non-tests.
nts, there are
reliability, and
the students'
g non-tests to
o, to be met.
epared or not,
have met the

teacher, should
nts to dig the
nts may not be
i. They cannot
y experienced
situation and
ch as, by their
ther kinds of
es are needed.
owadays, the
-based assess-
ormance test,
tion, attitude

ents' learning
specific tests,
se, Cloze Pro-
ests used, they
ble, practical,
ted. A reliable
chievements. A
schools often
ch focuses on
though in the
skills that the
at just certain
not yet be well
the students
cannot show
ations. Conver-

sely, there are students who got low English achievements, fortunately can show their real capability in using the language. It seems that the techniques the teachers use do not match with the curriculum being implemented ,i.e. Competency-Based Curriculum of 2004. Based on the curriculum should be done continuously and comprehensively.

In fact, in assessing the students' achievements, some teachers tend to use tests as the only instruments of assessment. This is supported by the result of the interview done with three English teachers. One of them who was interviewed on March 23, 2006 said that to assess her students' achievements, she used just tests. She does not use portfolio yet.

Another phenomenon of English assessment is that the instruments used are not valid. The reasons are the skills tested are not the ones taught. Based on the researcher preliminary observation in the field, the researcher recognized that to determine students' marks is based on the English final examination which did not test all skills taught. Most of the items in the final examination are aimed at measuring the student' reading comprehension and speaking ability. Unfortunately, to test the students speaking ability it was done indirectly, viz. through written/paper tests.

An English teacher who was interviewed on February 14, 2006 informed that the tests used for final examination were constructed by those working at *Kantor Wilayah Departemen Agama* (Religious Affairs Regional Office) and tested only certain language skills and language components. Therefore it is possible to say that the tests are not valid in terms of their objectives of English teachings at Islamic Senior High Schools, because they did not test all skills taught at those schools.

Another English teacher who was interviewed on March 27, 2006 informed that she had tried to use not only tests but also non-tests to assess her students' achievements. She, then, continued to say that she assessed all aspects of her students' learning, except for speaking. She just saw it from cognitive and psychomotor domains. Yet, she said, other English teachers at her school just used tests to assess the students' achievements.

The next phenomenon related to the English assessment is that different English teacher tended to use different techniques of assessments even though teach the students of the same level using the curriculum. This may be caused by the lack of coordination among them.

Based on the phenomena mentioned before, some questions arise. Do the English teachers of MAN 2 Batusangkar follow the same ways in their English assessments? How do they assess their students; abilities? These questions make the writer interested in finding out the techniques the English teachers use to assess their students' English achievements. Therefore, she would like to conduct a research about it at *Madrasah Aliyah Negeri 2 Batusangkar*.

There are two main reasons for choosing the school. First, the school was the best *madrasah* among Madrasah Aliyahs in Indonesia in terms of its achievements in the year 2004. Second, all of the English teachers of that school had been trained about the implemented curriculum. Based on the two reasons, the researcher thinks that the school is appropriate to choose as the place to conduct her research.

B. RESEARCH METHOD

The method applied was descriptive. There were four English teachers as the respondents of the research. The sample was taken by applying purposive sampling technique. To collect the data of the research, the researcher used a questionnaire, did interviews, and used related documents. The data of the research were analyzed quantitatively and qualitatively.

C. DISCUSSIONS

1. Review of Related Literature

a. Nature of Learning Achievement Assessment

According to *Departemen Pendidikan Nasional* (2003: 16), "Assessment is an activity to get information of the students' achievements as stated in the curriculum. Based on the curriculum, the assessment should be classroom-based. There are some principles that need to be considered when doing assessment. They are valid, fair, open, continuous, comprehensive, meaningful and educative.

Brown (2004) states "assessment is an ongoing process that encompasses a much wider domain. Tests are subset of assessment. Tests are relatively time-constrained and draw on a limited sample of behavior." Furthermore, he states that there are two functions of assessment: formative and summative assessment. Most of classroom assessment is formative assessment which evaluates students in the process of "forming" their competencies and skills with the goal of helping them to continue that growth

process. While summative assessment aims to measure or summarize what a student has grasped, and typically occurs at the end of a course or a unit of instruction. Final exams and a general proficiency exams are examples of summative assessment.

Achievement assessment, according to Hamalik (2003) the assessment of all activities (collecting data and information), analyzing, interpreting, and making decisions about the level of the students' achievement after doing learning activities in order to achieve the goal of learning already set.

b. Assessing Students' Achievements

Ansyar (2003: 331) states "students' achievements are assessed based on student's individual performance using criterion-referenced testing rather than norm-referenced testing".

Moekijat in Mulyasa (2005) the domain to be assessed involves knowledge, attitude and skills. Knowledge is assessed through written and oral test, and questionnaire, skills are assessed through practice test, skill and task analyses and the assessment by the student themselves, while attitude is assessed through a self-attitude check-list, the check list is matched with the goal of the program, and semantic differential scale. It means the assessment is a means to get information about their success in learning.

Furthermore, Mulyasa(2005) points out that the assessment of learning can be done through tests and non-tests. Tests can be forms of an oral test, a written test, and performance test. The tests involve a pre-test, a process evaluation, and a post test. A pre-test is conducted to anticipate the learning process to be conducted. The process evaluation is done to assess the quality of students' learning and the shaping of basic competencies in them, and also to see how the goals of learning are actualized. While a post-test is very useful mainly to see the learning success.

Tests used for diagnosis and feedback provide critical information about strengths and weaknesses in the learned abilities to the student, teacher, and administrator that should make the learning process more efficient. Another important use of tests is use assist in the decision of who should be allowed to participate in a particular program of instruction. In the area of language testing, according to Carrol in Henning (1987) a common screening instrument is termed an aptitude test. This kind of test is used to predict the success of failure of students prospective in a language program. The tests of the third kind, placement tests, which are used to identify a particular performance level of the student and place him or her at an appropriate level of instruction. Another common use of

information about purpose of tests.

According to Harmer or failure as a good item types that it states to perform the direct test items, on knowledge and ability by and productive skills. It is stated that to know it can be directly tested in

1996:14) states "an specific reference to a program usually means based on the objectives. It means that if a test is based on the objectives

There are three objectives which are: 1) Developing language, either spoken or written, listening, speaking, and awareness of the nature of language as a main tool for understanding about the world as well as to enlarge their horizons. It will have cross-cultural awareness of cultures.

It is stated that the aims of the program are to make the students able to use English to grow their awareness of the world, and to appreciate

It is stated that English language speaking, reading and writing, linguistics competency, and reading and writing ability, and reading and writing attitude toward the world, and that the student should have the necessary competencies, as

At the end of English learning program, the standard of competency that students should meet is that they are able to do spoken and written communication by using appropriate language diversities fluently and accurately in interactional discourses and monologues.

Based on the ideas above, it can be said that at the end of English learning program, the students should be tested based on the materials taught and the skills trained since Cameron (2001: 220) states "assessment should fit comfortably with students' learning experience". In other words, the assessment should be done based on their learning experiences. In addition, Green (1975) suggest that the measurement instrument or procedure must fit the specific objectives to be assessed; thus, there is no "all-purpose" test or procedure. In line with him, Brown (1996:269) points out "although tests should never be treated as though they are somehow divorced from the language teaching and learning process that are going on in the same context".

To make a result of the tests reflect what students have experienced in their classroom, the tests should be constructed based on some practical steps. They are: a) test toward clear, unambiguous objectives, b) from the objectives, draw up test specifications, c) draft your test, d) revise your test, and e) final-edit and type the test.

According to Finochiaro and Sako (1983) language test result are used principally for 1) measurement of students' progress toward specific objectives, 2) continual evaluation of the instructional program and of the language curriculum, 3) selection and placement of students into classes or groups, 4) diagnosis of strength or weaknesses, and 5) exclusion or certification.

In assessing students' achievement, a teacher as a test developer as suggested by Code of Fair Testing Practices in Education in Anderson .et. al., (1995:244) should follow 5 guidelines. The test developers should:

1. Define what each test measures and what the test should be used for.
2. Describe the population (s) for which the test is appropriate accurately represent the characteristic, usefulness, and limitations of test for their intended purposes.
3. Explain relevant measurement concepts as necessary for clarity at the level of detail that is appropriate for the intended audience (s).
4. Describe the process of test development. Explain how the content and skills to be tested were selected.

Based on the guidelines above, it can be concluded that to test students' achievement, a teacher should know, first, what to measure. It means in the area of language testing, a teacher should know what domains and skills to be measured.

According to Bachman (1990) there are two advantages of a test. They are as a source of information to make decisions (about the student and the teachers) in the educational context, and as an indicator of students' ability. The decisions about students are related to the process of selection (students' entrance and readiness), the process of placement, diagnosis and progress, as well as decision about the quality of students' achievements. While for the teachers, the test results can be used to design the next program of student learning.

Yasin (2000) implicitly states that to assess students' achievements, using a test only at the end of a learning program is not enough since the result can be influenced by the situation and condition when the students are having the tests, such as, by sickness and/or lack of concentration. Therefore, he suggests that teachers use portfolio assessment.

In line with Yasin, related to the curriculum being implemented nowadays, that is, Competency-Based Curriculum or Curriculum 2004, Surapranata and Hatta (2004) state that the assessment should be done by using classroom-based assessment (CBC). Which involves a written/paper test, a performance test, giving assignment, project evaluation, product evaluation, attitude assessment, and portfolio assessment.

Mulyasa (2005) adds that to assess students' achievement can be done through a classroom-based assessment, a basic performance test, a school based examination, a benchmarking, a program evaluation, and portfolio assessment. Based on the ideas above, it can be concluded that to assess students' achievement, using test only is enough. A teacher should also use non-tests.

In addition to some ideas proposed above, learning achievement assessment based on the Curriculum 2004 according to Mulyasa (2005) can be done through classroom-based assessment, basic ability test, school-based examination, benchmarking, program evaluation, and portfolio. Classroom-based assessment can be done through post/daily-tests (after finishing a learning process for certain standards of competencies), final-tests (at the end of each semester) and tests conducted at the end of a school program.

cluded that to test
ow, first, what to
g, a teacher should

two advantages of a
ke decisions (about
context, and as an
about students are
entrance and readi-
progress, as well as
ments. While for the
the next program of

ess students' achie-
ning program is not
the situation and
tests, such as, by
re, he suggests that

um being implemen-
curriculum or Curri-
ate that the assess-
d assessment (CBC).
rformance test, giving
tion, attitude assess-

nts' achievement can
t, a basic performan-
marking, a program
n the ideas above, it
achievement, using
non-tests.

ove, learning achie-
n 2004 according to
m-based assessment,
tion, benchmarking,
m-based assessment
finishing a learning
s), final-tests (at the
t the end of a school

Surapranata and Hatta (2004:5) say " classroom-based assess-
ment is the assessment that a teacher does in to assess the learn-
ing process conducted". The goal is to know students' progress and
their achievements, to diagnose their learning difficulties, and to
give feedback/improvement of the learning process, class promotion
and to motivate them.

In line with them, Mulyasa (2004) points out that classroom-
based assessment is done to know the progress of the students, to
diagnose their learning difficulties, to give feedback of their learn-
ing and to decide their class promotion.

Furthermore, Supranata and Hatta (2004) state that there are
several principles to follow to do classroom-based assessment. They
are motivating, valid, fair, open, continuous, meaningful, and
educative. The assessment can be done through written tests,
actional/performance test, assignments, project evaluation, product
evaluation, attitude assessment, and portfolio assessment. Written
tests can be administered when the students have post-tests or
final/semester tests. The tests are in the forms of multiple choice,
matching, true-false, cloze procedure, and essay. Actional test is
done while learning process is taking place through observation
toward learner's behaviors. Assignment is for all subject matters
based on the materials taught and learners' growth by considering
the quantity of the assignments (don't give too many assignments),
kinds and materials (should be essential, train, enrich learners'
horizon, and the development of their creativity, responsibility, and
independence. Project evaluation is classroom-based assessment
toward assignment at a given time. It is done by starting from
collecting, organizing, evaluating, up to data presentation.

The assessment of learners' products is classroom-based
assessment toward learners' skill mastery in making a product
(process), and the quality of students' certain work (product). The
evaluation involves three steps, namely: preparation, production,
and evaluation.

Attitude assessment is directed toward various objects, i.e.,
attitude toward learning materials, toward content teachers, toward
learning process, and the like. This assessment is done through
various ways, like observation, direct question, personal report, and
the use of affective scale.

Basic competence tests are done to know students' basic
competencies, primarily in reading, writing, and counting needed in
improving learning program (remedial teaching). School-Based
Assessment (SBE) can be done at the end of each school level, to get
complete and comprehensive description about students' learning

mastery at a given time, and the success of a school in general. Benchmarking is a standard to measure ongoing performance and processes, and results to achieve satisfying excellence. Program evaluation is conducted continuously by the *Departmen Pendidikan Nasional* and Dinas Pendidikan. While, portfolio assessment is the evaluation of all tasks done by the learners in certain subject matters. Brown (2001: 237) states "a portfolio is a collection of examples of work that, as a collection, reveal both the capability and the progress of learner". The characteristics of portfolio assessment according to Yasin (2000) are comprehensive, planned, and systematic, informative, patterned, and authentic.

Some points that need to be considered in conducting portfolio assessment according to Mardhapi et.al., in Mulyasa (2005) are as follow:

1. The artifacts collected should be the learners' true artifacts;
2. Deciding what tasks to be done;
3. Collecting and keeping the samples of the artifacts;
4. Deciding the criteria for portfolio assessment;
5. Asking learners to continuously evaluate their portfolio;
6. Planning a meeting with the learners being evaluated;
7. Involving parents and society in doing portfolio assessment.

Portfolio assessment is classroom-based assessment toward the collections of learners' artifacts systematically ordered and organized. Barton and Collin in Supranata and Hatta (2004) point out that the object of the portfolio can be divided into four, i.e.,:

1. The learners' artifacts are those produced in the classroom;
2. The reproductions are those produced outside of the classroom;
3. Attestations are the statements and the result of observation performed by teachers or others about the learners;
4. Productions are those of the learners' artifacts specially prepared for portfolio.

c. Assessing Language Skills

There are four language skills that all language users perform. Brown (2004: 117) says "all language users perform the acts listening, speaking, reading, and writing. If a teacher wants assess a student's ability, in one or combination of the four skill he or she assesses that student's competence, but he or she observe the student' performance". Moreover he says "in assessing

learner's competence, a teacher considers two principles. The first one is a fallibility of the results of a single performance, such as that produced in a test. ...The second one is that we teachers must rely as much as possible on an observable performance in our assessments of students".

It is the obligation as a teacher to validate measurements by considering two (or more) performances and/or contexts before drawing a conclusion. That could take the form of one or more of the design below:

- several tests that are combined to form an assessment;
- a single test with multiple test tasks to account for learning styles and performance variables;
- in-class and extra-class graded work;
- alternative forms of assessments (e.g. journals, portfolio, conference, observation, self-assessment, and peer assessment).

The techniques of assessment based on CBC include quiz, oral interview, post-test, formative test, individual task, group task, practice, and project report. While the types of instruments that can be used can be in the forms of tests and non-tests, which include multiple choice, closed-ended question, open-ended question, short answer or cloze procedure, matching, performance, portfolio, interview, inventory, or observation.

1. Assessing Listening Ability

To assess one's listening ability, Brown (2004: 121-130), Hughes (1989: 134-135), and Weir (1993: 98) implicitly state that an evaluator should follow two steps: determining objectives or specifying what the candidate should be able to do, and designing or setting assessment tasks. Brown (2004: 119-120) states "designing appropriate assessment tasks in listening begins with the specification of objectives, or criteria'. He, then, continues that objective may be classified into four types of listening performance, i.e. intensive, responsive, selective, and extensive. In the performance, micro-and macro skills are included.

Richard as quoted by Brown (2004: 121) classifies a listening skills into microskills, and macroskills which provide 17 different objectives to assess in listening. The microskills include:

1. Discriminate among the distinctive sounds of English;
2. Retain chunks of language of different lengths in short-term memory;

3. Recognize English stress patterns, words in stressed and unstressed positions, rhythmic structure, intonation contours and their role in signaling information;
4. Recognize reduced forms of words;
5. Distinguish word boundaries, recognize a core of words, and interpret word order patterns and their significance;
6. Process speech at different rates of delivery;
7. Process speech containing pauses, errors, corrections, and other performance variables;
8. Recognize grammatical word classes (nouns, verbs, etc), systems (e.g. tense, agreement, pluralization), patterns, rules, and elliptical forms;
9. Detect a sentence constituents and distinguish between major and minor constituents;
10. Recognize that a particular meaning may be expressed in different grammatical forms;
11. Recognize cohesive devices in spoken discourse.

The macroskills include:

1. Recognize the communicative functions of utterances according to situations, participants, goals;
2. Infer situations, participants, goals using real-world knowledge
3. From events, ideas, and so on, described, predict outcomes, infer links and connections between events, deduce cause and effect and detect such relation as main idea, supporting idea, new information, given information, generalization, and exemplification;
4. Distinguish between literal and implied meanings;
5. Use facial, kinesics, body language, and other non verbal clues to decipher meanings;
6. Develop and use a battery of listening strategies, such as, detecting key words, guessing the meaning of words from context, appealing for help, and signaling comprehension or lack thereof.

Related to the objectives that need to be assessed, Hughes (1989: 133-134) suggests to specify what the candidate should be able to do which contains macro- and micro-skills. The macroskills would be directly related to candidate's needs or to course objectives, and might include: listening for specific information, obtaining gist of what being said, following directions, and following

instructions. While the macro-skills might include: interpretation of intonation patterns (recognition of sarcasm, etc), and recognition of function of structures (such as interrogative as request). At the lowest levels are abilities like being able to distinguish between phonemes (for example /w/ and /v/).¹²

Based on the ideas of Brown and Hughes above, it can be concluded that to assess one's listening ability, he or she should be tested based on some macro- and micro-skill as the objectives. In other words, the assessment should be specified on what the candidate should be able to do.

Hughes (1989) proposed some techniques to use to assess one's listening comprehension. The techniques are multiple choice, short answer, information transfer note-taking, and partial dictation.

In line with Hughes, Weir (1993: 104-126) proposes some techniques for testing listening. They are short answer questions, information transfer techniques, dictation, and listening recall.

Based on the Curriculum 2004, listening should be assessed through performance tests. The tests can be in the forms of oral interview, retelling, or discussion. While the criteria of the assessment will include fluency, vocabulary, structure, comprehensibility, and listening comprehension.

2. Assessing Speaking Ability

To assess students' speaking ability, Hughes 1989: 101-110) suggests that a teacher, first set the tasks by specifying all appropriate tasks, obtaining a sample that properly represents each candidate's ability and which can be reliably judged, and obtaining valid and reliable scoring.

Brown (2004: 141-142) five speaking performance tasks that need to be assessed are imitative, intensive, responsive, interactive, and extensive (monologue). Imitative is the ability to simply parrot back (imitate) a word or phrase or possibly a sentence (pronunciation). Intensive is the type of speaking frequently employed which is the production of short stretches of oral language designed to demonstrate competence in a narrow band of grammatical, phrasal, lexical, or phonological relationship (such as prosodic elements—intonations, stress, rhythm, juncture). The examples of this are directed response tasks, reading aloud, sentence and dialogue completion limited picture-cued tasks including simple sequences, and translation up to the simple sentence level. Responsive assessment tasks include interaction and test comprehension

but at somewhat limited level of very short conversations, standard greetings and small talk, simple request and comments, and the like. Interactive speaking is longer and more complex than responsive, which sometimes includes multiple exchanges and/or multiple participants. Interactive can take the two forms of transactional language, which have the purpose of exchanging specific information, or interpersonal exchanges, which have the purpose of maintaining social relationships. While extensive oral production tasks included speeches, oral presentation, and story-telling, during which the opportunity for oral interaction from listeners either highly limited or ruled out altogether.

Moreover, based on the skills and tasks to be mastered by the students, Brown (2004: 144-183) proposes some skills to be assessed. They are imitative speaking, intensive speaking, responsive speaking, interactive speaking, and extensive speaking. Imitative speaking can be in the forms of repetition of words, phrases, and sentences. Another kind of test that uses imitative (as well as intensive) production task is PhonePass.. Intensive speaking includes directed response tasks, read-aloud tasks, sentence/dialogue completion tasks and oral questionnaires, picture-cued tasks, and translation (of limited stretches of discourse). Assessment of responsive tasks includes question and answer, giving instructions and directions, and paraphrasing. Interactive speaking is assessed through interview, role play, discussion and conversations, games, and oral proficiency interview (OPI). Every effective interview, according to Brown (2004:168) contain four stages warm-up, level-check, probe, and wind-down. While for extensive speaking, the assessment is done through oral presentations, picture-cued-story telling, retelling a story, news event, and translation (of extended prose)

Based on the Curriculum 2004, speaking is assessed through performance tests in the forms of oral interview, retelling, or discussion. Moreover, the criteria of assessment will include fluency, vocabulary, structure, comprehensibility, and listening comprehension.

To assess speaking ability of a student, there are several micro- and macro-skills to become the objectives of an assessment task. The skills are as follow:

Micro-skills include:

1. Produce differences among English phonemes and allophonic variants;
2. Produce chunks of language of different lengths;

versations, standard
 comments, and the
 more complex than
 exchanges and/or
 two forms of transac-
 exchanging specific
 have the purpose of
 nsive oral production
 and story-telling, during
 from listeners either

to be mastered by the
 some skills to be
 nsive speaking, respon-
 nsive speaking. Imita-
 on of words, phrases,
 ns imitative (as well as
 Intensive speaking
 ud tasks, sentence/
 nnaires, picture-cued
 of discourse). Assess-
 n and answer, giving
 g. Interactive speaking
 discussion and conver-
 (OPI). Every effective
 rtain four stages warm-
 le for extensive speak-
 presentations, picture-
 ent, and translation (of

ng is assessed through
 w, retelling, or discussi-
 include fluency, voca-
 ening comprehension.

ent, there are several
 tives of an assessment

emes and allophonic

lengths;

3. Produce English stress patterns, words in stressed and unstressed positions, rhythmic structure, and intonation contours;
4. Produce reduced forms of words and phrases;
5. Use an adequate number of lexical units (words) to accomplish pragmatic purposes;
6. Produce fluent speech at different rates of delivery;
7. Monitor one's own oral production and use various strategic devices—pauses, fillers, self-corrections, backtracking—to enhance the clarity of the message;
8. Use grammatical word classes (nouns, verbs, etc.), systems (e.g., tense, agreement, pluralization) word order, patterns, rules, and elliptical forms;
9. Produce speech in natural constituents: in appropriate phrases, pause groups, breath groups, and sentence constituents;
10. Express a particular meaning in different grammatical forms;
11. Use cohesive devices in spoken discourse.

Micro-skills include:

1. Appropriately accomplish communicative functions according to situations, participants, and goals;
2. Use appropriate styles, register, implicature, redundancies, pragmatic conventions, conversation rules, floor-keeping, and yielding, interrupting, and other sociolinguistic features, in face-to-face conversations;
3. Convey links and connections between events and communicate such relations as focal and peripheral ideas, events, and feelings, new information and given information, generalization, exemplification;
4. Convey facial features, kinesics, body language, and other non-verbal cues along with verbal language;
5. Develop and use a battery of speaking strategies, such as emphasizing key words, rephrasing, providing a context or interpreting the meaning of words, appealing for help, and accurately assessing how well your interlocutor is understanding you.

3. Assessing Reading Ability

According to Brown (2004: 186), there are three types or genres of written text. They are academic reading, job-related reading, and personal reading. The content validity of an assess-ent

procedure is largely established through the genres of texts. Aside from attending to genres of text, according to him, reading skills and strategies are also important for assessment of reading ability. The micro- and macro-skills below represent the spectrum of possibilities for objectives in the assessment of reading comprehension.

The microskills include:

1. Discriminate among the distinctive graphemes and orthographic patterns of English;
2. Retain chunks of language of different lengths in short-term memory;
3. Process writing at an efficient rate of speed to suit the purpose;
4. Recognize a core of words, and interpret word order patterns and their significance;
5. Recognize grammatical word classes (nouns, verbs, etc.), systems (e.g., tense, agreement, pluralization) word order, patterns, rules and elliptical forms;
6. Recognize that a particular meaning may be expressed in different grammatical forms;
7. Recognize cohesive devices in written discourse and their role in signaling the relationship between and among clauses.

The macroskills include:

1. Recognize the rhetorical forms of written discourse and their significance for interpretation;
2. Recognize the communicative functions of written texts, according to form and purpose;
3. Infer context that is not explicit by using background knowledge;
4. From described events, ideas, etc., infer links and connections between events, deduce causes and effects, and detect such relations as main idea, supporting idea, new information, given information, generalization, and exemplification;
5. Distinguish between literal and implied meanings;
6. Detect culturally specific references and interpret them in a context of the appropriate cultural schemata;
7. Develop and use a battery of reading strategies, such as scanning and skimming, detecting discourse markers, guessing the meaning of words from context, and activating schemata for the interpretation of texts.

Moreover, according to him, a variety of performance is derived more from the multiplicity of types (the genres listed above). Nevertheless, for considering assessment procedures, several types of reading performance are typically identified. They are perceptive, selective, interactive, and extensive.

For perceptive reading, assessment of basic reading skills may be carried out in a number of ways: reading aloud, written response, multiple choice, picture-cued items (picture-cued word identification, picture-cued sentence identification, picture-cued true-false sentence identification, and picture-cued matching word identification).

For selective reading which involves lexical and grammatical aspects of reading, some possible tasks that can be used to assess reading ability are: multiple choice (multiple choice vocabulary-grammar tasks, contextualized multiple-choice vocabulary-grammar tasks, and multiple choice cloze vocabulary-grammar task), matching tasks (vocabulary matching task, and selected response fill-in vocabulary task, editing task, picture-cued tasks (multiple-choice picture-cued response, and diagram-labeling task), filling and tasks

To assess interactive reading, some following tasks are used: cloze tasks (cloze-procedure fixed ratio deletion (every seventh word), cloze-procedure, rational deletion (preposition and conjunction) C-test procedure and cloze elide procedure, impromptu reading plus comprehension questions, short-answer tasks, editing (longer tasks), scanning, ordering tasks, and information transfer (reading charts, maps, graphs, and diagrams).

For extensive reading, some tasks that can be used are skimming tasks, summarizing and responding, and note-taking and outlining. Whereas based on the Curriculum 2004, the assessment of reading ability can be done by using tests: oral interview, multiple choice, close-ended question, and so on. For reading aloud, especially, the test is conducted by asking students to read a text loudly by seeing the criteria: spelling accuracy, intonation, performance, etc.

4. Assessing Writing Ability

Genres of written language or writing according to Brown (2004: 219) include: academic, job-related, and personal writing. While, written performance can be categorized as imitative, intensive (controlled), responsive, and extensive writing. In imitative writing, there are some fundamental, basic tasks of writing letters, words, punctuation, and very brief sentences. This category includes the ability to spell correctly and to perceive phoneme-grapheme