An Analysis of English Tests at Pilot International Standard School in 2010/2011 – 2011 /2012 Academic Year

Rika Afriyanti, S.Pd., M.Pd.
STKIP PGRI Sumatera Barat
E-mail: afriyantirika@yahoo.com

Abstract

This research was based on the phenomenon happened in one of the Pilot International Standard Schools (RSBI) in Padang in which most of the students' score at grade VIII could not fulfill the minimum standard of achievement (KKM) in English semester test designed by the Central government in 2010/2011 academic year. Therefore, this research was aimed at analyzing whether the test designed by the Central government contributes to the students' failure or not. However, the researcher did not only analyze the test designed by the Central Government but also analyse the tests designed by the teachers. The researcher did qualitative item analysis which focused on analyzing the format of the test items (stem, options and distractors), content validity and face validity. This research was conducted based on descriptive method and there were two sources of data of this research: the document of English tests and the respondents of interview. The result of this research is the test designed by the Central Government in 2010/2011 academic year is the better one, because almost all the items could fulfill all the criteria in designing multiple-choice test. Meanwhile, the Mid-test designed by the teachers in 2010/2011 academic year indicated as the poor one. There are some criteria that could not be fulfilled by all the tests. They are; the curriculum demand, the relevance between the assessment plan/blue-print with the items, writing the stem and many items are not paraphrased. Therefore, it can be interpreted that the teachers still do not have enough competence in designing the test which based on National standard and the government also, because the researcher still found the test items designed by the central government that could not fulfill the criteria in writing a good multiple-choice test.

Keywords: English Tests, Pilot International Standard School

INTRODUCTION

One of the strategies to increase the quality of education is to develop the schools that have international standard. The government (in Education law No 20, 2003) has stated that central or regional government should develop at least one school level for all grades to be an International standard school. It is clearly stated that each province in Indonesia should have at least one International standard school. It is started from Elementary School up to Senior High School.

Furthermore, the Directorate General of Management of Primary and Secondary Education (2010:5) states that ISS (International Standard School) is defined as the school which fulfils National Educational Standards and also accomplishes the education standard from economical developed countries. It explains that SBI is a school which should have

an excellent quality because it has an international qualification for the curriculum, the teachers, students, the facilities. In other words, they have equivalent standard to the schools from other countries.

Furthermore, to achieve National Education standard, one thing that must be fulfilled by SBI is evaluation standard. Evaluation is an activity to get, analyze and interpret data about the students' achievement to the teaching and learning process and it becomes a continuation activity in order to make judgment for the students and the teacher. It means that the main purpose of evaluation is to measure students' achievement toward the teaching objectives and also the teacher's success in the teaching process. The process of evaluation is done by giving the tests to the students. There are two kinds of tests that are usually given at schools; formative tests and summative tests. Formative tests are done to measure students' achievement for one or two units of lesson, for instance; daily test and Mid-test. Then, summative tests are given at the end of semester to evaluate students' achievement within one semester, for example is semester test.

To know whether the students can pass the test, each school has own standard that is called as KKM (Minimum Standard Achievement). Especially, for International Standard Schools including SMP Negeri 8 which is as piloting of International Standard School in Padang, the minimum standard or score for the English test is 80. This score is based on the evaluation standard for International Schools in which the central government has stated that the minimum score for the students' recruitment is 75. Consequently, the Minimum Standard Achievement of the teaching and learning process must be above the recruitment score.

Considering the explanation about SBI, there is an assumption that the students from International Standard School will have good score for the tests because they are selected students. Nevertheless, the researcher found different fact about the students' score at International Standard School for their English semester test. This fact is based on the researcher's preliminary research at SMP Negeri 8 Padang which is as one of Pilot International Schools in Padang. It will be explained as the following.

The researcher conducted the preliminary research at grade eight students of international class. There are eight International classes: VIII A – VIII H. All of the classes are taught by two teachers. They followed the same English semester test which is designed by the Central Government. This is the first time for the students to get English semester test which is designed by Central Government. The Minimum Standard Achievement (KKM) that must be achieved by the students is 80. However, the data of students' scores showed that the students failed to achieve that score. It can be seen from the average of the students' mean scores for each class are; VIII/A 59.2, VIII/B 61.5, VIII/C 68.4, VIII/D 68, VIII/E 47.6, VIII/F 55, VIII/G 54.5 and VIII/H 57.8.

The data above show that the lowest average of the English semester test is 47.6 and the highest score is 68.4. It means that the mean scores of the English test are under the minimum standard achievement that is 80. Hence, It can be concluded that all the students fail in English semester test.

The fact above attracts a big question about the quality of International schools. Ideally, the students could pass the test because they are selected students, but the fact is different from the expectation. The data of the students' scores indicated that the International students do not show significant different quality from Non-International Schools. There are some possible indications that cause the problem above. The first

possible cause is the process of students' recruitment. The students at International School have to pass some tests and get minimum score 75. Then, the teaching and learning process may be the second possible indication of the problem above. Ideally, the teachers who teach at International schools are challenged to prepare the best and up to date teaching materials, because the curriculum used in these schools require it. The third possibility that causes the problem above is the test itself. It is assumed that the test is too difficult for them.

Dealing with the causes above, the researcher assumes that the test contributes to the students' failure, because the researcher's informal interview with the English teachers showed that most of the students got difficulty in doing the test. It may be happen because of some factors: first, the construction of the tests designed by the central government is different from the tests designed by the teachers even though they use the similar curriculum. The differences can be from the vocabulary or the word choice, the context of the tests and so forth. Second, this is the first time for the students to have semester test from the Central Government, because they usually have the test from their teachers. Sometimes, the test designed by the teachers have different face validity from the test designed by the Central Government, for example the directions of the test and the number of items in the test. The directions and the items of the test from the Central Government is more than the test from the teachers. In addition, the tests designed by the central government usually consist of 60 items, but the tests made by the teachers have 50 items. These facts may contribute to the students' failure, because the students may be little bit shocked with the test. That is why an investigation about the tests designed by the Central government is needed to do. However, the researcher needs to analyse the tests designed by the teachers, because they give information about the teaching and learning process indirectly. Hence, the researcher knows the specific causes of the students' failure. One of the ways to do it is analysing the tests. Analysing the tests can be done by doing item analysis. Sudjana (2009:135) explains that item analysis refers to the study of the test items in order to decide whether the test items are good or not. It means that item analysis will give information to the test makers about the construction each item of the test. Because the type of the test item that will be analysed is multiple-choice test, item analysis will provide information about the content validity of the test, face validity, stem, options and distractors. Finally, the result of analysis determines which item is perfect, less perfect and imperfect.
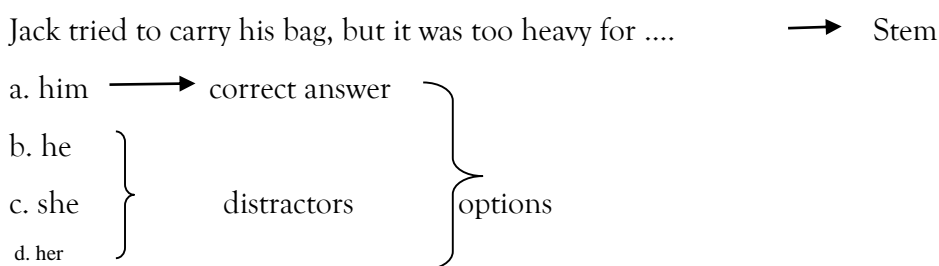
Multiple-Choice Tests

One of the types of the test items that used at school is multiple-choice test. This test has a special characteristic that is the students have to cross of the correct answers from the alternatives given. It is similar with McMillan theory (1997: 338) who states that multiple-choice test is a test which asks the test-takers to write their response by crossing one of the correct answers from the options given. It means that this test offers some options to the test-takers, but only one is correct. It can be said that the other options are only as the distractors and they are used to make the test-takers feel doubt in choosing the answer. Indirectly, multiple-choice test can develop the test-takers critical thinking, because they have to analyze and think hard to find the best answer from the options available.

Furthermore, Nitko adds that the test makers have to know the format in designing a multiple-choice test. He points out that there are five formats of a multiple-choice test. They are stem, alternatives, keyed alternative, distractor, and interpretive material. The following is some explanation about all of multiple-choice formats.

The first one is stem that is the part of the item that asks question. Therefore, one item consists of the problem that will be questioned to the students. The second format is alternative that is the options that will be chosen by the test takers. The alternatives should be arranged in a meaningful way (logically, numerically, alphabetically, and so forth). If no logical or numerical order exists among them, the alternatives should be arranged in alphabetically order. The third format is keyed alternative and distractors. The correct or the best answer from all the options suggested is called as keyed alternative. Then, distractors refer to incorrect alternatives. They have a purpose to present plausible but incorrect answers to the question. The last format is interpretive material that is additional information to make a question clearer or more authentic, such as picture, graph, and table. If adding this kind of information makes the stem more than one or two sentences long, then the information is placed in a section that comes before the stem. The following is the example of all the elements above:

For example:

Jack tried to carry his bag, but it was too heavy for ....　　　⟶　　Stem

a. him ⟶ correct answer

b. he

c. she　　　distractors　　options

d. her

The sample of multiple-choice item with interpretive material

For example: Publishing schedule

Magazine	Frequency of Publishing
Hello	1st every month
Smart	1st and 5th every month

The table shows that Hello is published ... month
A. Once　　　　　C. Three times
B. Twice　　　　　D. Four times

From the explanations above, it is known that there are four formats of multiple-choice test; stem, options, distractors and correct answer. All the formats must be had in each item of the multiple-choice test. However, to determine each item fulfills all the formats, the test-makers have to check the each item before testing them. Analyzing the test items can be done in quantitative and qualitative. Quantitative refers to numerical data, such as analyzing distractors, level of difficulty, and index discrimination. Meanwhile,

qualitative item analysis refers to non- numerical data, such as the content validity and the format of the test or this analysis based on the theories from the experts. It is supported by Borich & Kubiszyn (2003: 197) who divide the item analysis into two categories; quantitative and qualitative analysis. Quantitative item analysis consists of analyzing distractors, difficulty index and discrimination index and qualitative item analysis focuses on checking an item's content validity and inspecting it for technical faults.

However, the researcher focuses on qualitative item analysis, because it is considered more important than quantitative item analysis. It is supported by Borich & Kubizsyn (2003: 2), "Quantitative item analysis is useful but limited. It points out that quantitative item analysis could not explain the problems of the item. In other words, quantitative item analysis is fallible". Then, this theory also emphasizes that qualitative item analysis is more effective than quantitative item analysis. It does not need the students' response, because it only analyzes the test objective, content validity and technical item quality. Test objective is a part of content validity, because developing content validity is done by matching the test items to the test objectives. Then, technical item quality is assumed as the quality of the test items based on the procedures in writing the items of the test itself, in this case is multiple-choice test. Therefore, the guidelines of designing multiple-choice test can be used to analyze the test items.

In the following part, the researcher will discuss more parts of qualitative item analysis; the guidelines in writing multiple-choice test and validity of the test items.

Writing the Multiple-choice Test

Dealing with the explanation above, there are three elements of multiple-choice tests: stem, options, and distractor. To write those of them, the test-makers have to know some guidelines to construct them. The following is the detailed explanation each of element. The first one is writing the stem. Stem refers to the problem that should be solved by the students or the test-takers in an item. To write and analyze the stem of the item, some experts proposed the similar theories about writing the stem. They are Nitko (1996:153), Haladyna, et al. (2002: 4), Borich & Kubiszyn (2003:116). The following is the explanation.

There are ten guidelines to write the stem. They are: (a) writes direct question. It means that the stem consists of a clear problem to be solved by the students. (b) If an incomplete sentence is used, be sure that it implies a direct question and the alternative come at the end (rather than in the middle) of the sentence. (c) Control the wording. It points out that the test-makers have to choose vocabulary for item which is suitable with the level of the students. (d) In items testing definitions, place the word or term in the stem and use definitions or descriptions as alternatives. (e) Avoid window dressing. It means that the test-makers must avoid extraneous, superfluous and non-functioning words or phrase in the stem. (f) Avoid negative worded stem. For example, "which of the following is not ….". However, if the test-makers use negative word, it is only used in the stem or only in an option (not both) and then, negative word must be written in capital letters. (g). avoid to ask personal opinion, because multiple-choice test needs the students to choose one of the correct answers not the students' opinion. (h). avoid textbook wording. It explains that the stem in an item will be functioned well if it is not quoted directly from the text and the effect is the item will have poor design. The better one is the test-makers have to paraphrase

the question. (i). avoid "cluing" and "linking" items. It emphasizes that the test-makers have to avoid linking the correct answer from one item to the correctness of the answer to a previous item. (j) Using negative questions or statements such as NOT or EXCEPT in the stem if the knowledge being tasted requires it, but it is must be written in capital letters and bold type.

From the experts' explanation, it is known that the stem must consist of the specific problem to be solved. Then, the other important points should be considered are the vocabulary and the level of grammar used are based on the students' level. Shortly, a good stem is understandable and challenging.

Furthermore, writing the alternatives and distractors also very essential in designing the multiple-choice test. To write them, Nitko (1996: 153), Haladyna, et al. (2002: 4) propose their ideas about them. The following is the detailed explanation.

There are fifteen guidelines in writing the options and distractors. They are: (a) strive for creating three to five functional alternatives. It means that an item cannot have the alternatives less than three or exactly two, because the test-takers will be easy to eliminate it. In addition, each alternative attracts at least one of the students who does not have the knowledge that you are expected and this is the meaning as functional alternatives. (b) all alternatives should be homogeneous and appropriate to the stem. It explains that the homogeneity of the alternatives is used as the distractors. Lack of homogeneity is a primary reason why the distractors do not function. Also, the alternatives must appropriate to the stem. As the example, if the stem requires the test-takers to identify the someone's name, the alternatives should be name. The homogeneity is also used as the control for difficulty of an item. (c) Put the repeated words and phrases in the stem. It emphasizes that it is better to put in the stem the words or phrases which are repeated in each alternative. (d) Use consistent and correct punctuation in relation to the stem. (e) Arrange the alternatives in a list format rather than in tandem. It explains that the alternatives are easier to read if they are arranged one below the other in a list form than beside one another. (f) Arrange the alternatives in a logical or meaningful order. For example, the alternatives are arranged chronologically or alphabetically. (g) All distractors should be grammatically correct. it means that if there is ungrammatically distractor, it will be easy to the students to eliminate incorrect or correct answer. (h) Avoid overlapping alternatives. (i) avoid making the alternatives a collection true-false items. (j) Avoid using "not given" none of the above" etc. (k) avoid using "all of the above". (l) Avoid using verbal clues in the alternatives. (m) Avoid using technical terms, unknown words or names and silly terms or names as distractors. (n) Avoid making it harder to eliminate a distractor than to choose the keyed alternative. All the points must be avoided in writing the alternatives. (o) Change the location of the correct answer according to the number of choices.

Based on the explanation above, the researcher concludes that the theories proposed by the experts above are almost similar, but there some main points that should be underlined in writing alternatives and distractors. First, the options must be homogeneous and appropriate to the stem. Second, there is only one correct answer in the option. Third, Avoid to use non-of-the-above or *all-of-the-above* as the options. Fourth, avoid giving clues for the correct answer. Fifth, the distractors should be plausible and they are based on the students' common errors.

Dealing with the theories about writing the stem, the alternatives and the distractors above; the researcher concludes all of them to be the guidelines in analyzing the format of

multiple-choice test. However, the researcher chooses the theory proposed by Nitko because he has made the guidelines above into the complete one but simple to be used. Nitko turns all of the guidelines in writing the multiple-choice format becomes eleven criteria and turns them into questions. The criteria are: (1) the test assesses an important aspect of the unit's instructional targets. It refers to the ability of the test to fulfill the curriculum demand. (2) The exercise matches your assessment plan in terms of performance, emphasis, and number of points. It refers to the table of specification or the blue-print of the test in which this question will answer about the appropriateness between the test and the blue-print of the test.(3) The stem of the item asks a direct question or set a specific problem. (4) The item based on paraphrase (rather than lifted directly from a textbook). (5) The vocabulary and the sentence are relatively low and nontechnical level (6) Each incorrect alternative (foil) is plausible so that it cannot be viewed as absurd or silly by a student who lacks knowledge of the correct answer. (7) if possible, every incorrect alternative (foil) is based on a common student error or misconception. (8) The correct answer to this item is independent of the correct answer to other test items. (9) All the alternatives are homogenous and appropriate to the content of the stem. (10) Avoiding to use "all of the above" and "none of the above" as much as possible?. (11) There is only one correct answer or best answer to the item?. All of those criteria are complete enough, because they have questions about the stem, options, distractors and also the curriculum achievement. Therefore, the researcher uses them as the guidelines to analyze the multiple-choice format.

Testing Validity

The success of the teaching and learning process in the classroom or in an education institution does not only depend on the process itself but also it is determined by the quality of the instrument used, in this case is the test. One of the qualifications of a good test is measuring what should to be measured or it is called as the validity of the test. There are some experts propose the concept of validity.   Weir (1990: 22) explains the definition of validity of the test. He says that validity refers to the ability of the test to measure what it is intended to measure. Then, the same statement also comes from McMillan (1997:55) who explains that validity is the extent to which a test measures what it is supposed to measure. From both of the experts' points of view, it can be said that validity means the test can accomplish the objective of the assessment, or in other word the test truly measures what it claims to measure. For instance, a valid test of reading comprehension measures the students' ability in understanding the content of the text. Some of the tests can be given to the students, for instance; asking the students to answer the questions related to the text or summarizing the content of the text. On the contrary, the reading comprehension test is invalid if it only asks the students to read the text loudly. Why it is invalid?, because the students only decode the text with correct pronunciation and punctuation without comprehending the text, hence reading aloud does not fulfill the aim of assessment.

However, It is not as simple as our imagination to create a valid test because there are several criteria that must be fulfilled in deciding whether a test has sufficient validity or they are commonly known as types of validity. Mcmillan (1997: 56) and Borich & Kubiszkyn (2003: 299-302) have same points of view in categorizing types of validity. However, Hughes (2003: 26-33) divides validity into five categories by adding two more

types of validity. They are validity in scoring and face validity. Hence, there are five types of validity according to both of the experts. They are content validity, criterion-related validity and the last is construct validity, scoring and face validity. However, this research only focused on content validity and face validity. Therefore, the following part will discuss more about them.

As stated previously, content validity of a test is determined by matching the test items to instructional objectives. It is supported by Hughes (2003: 16) who states that a language testing has a good content validity if the test items represent the language skills, structure, etc. From this theory, it is obvious that a speaking test, for instance, will have content validity if the test evaluate student's speaking ability or the students have to speak up in order to know their speaking ability. However, it does not give a guarantee that the test has fulfilled a content validity of a speaking test. The test will have content validity if it includes such as the sample of relevant expression and linguistic features that must be used by the students in speaking.

Furthermore, to analyze whether the test items have distributed the specific learning objectives or the domain that will be covered. The test-makers have to check the curriculum and write the table of specification or test blue-print which consists of the list of materials that will be evaluated. It is supported by Hughes (2003:16) who states that to determine the test has content validity; the test-makers have to write the specification of the skills or topics of the test items. Then, Borich & Kubizskyn (2003: 300) also add that content validity can be determined by matching the test items with instructional objectives. From both of theories, it is known that to fulfill content validity of a test, speaking test for instance, the test-makers have to know the instructional objectives that is stated in curriculum in order to help the test-makers in writing the test-specification or blue-print which provides some guidelines of topics that will be evaluated.

To sum up, content validity refers to the test questions represent the area that want to be covered. To analyze content validity can be done by matching the test items to the instructional objectives which comes from the curriculum. Then, the test makers have to write the test-blue print before designing the test items because the blue-print of the test reflects the achievement of the teaching and learning process toward the curriculum. However, to make the analysis of content validity complete, it is needed to know whether the test looks valid from the surface or the test does not make surprised. For example, the teaching material, the instructions, the number of items, the pictures or diagrams used in the test. all of them are part of face validity. Therefore, the discussion about face validity is considered important. The following paragraphs provide the discussion about face validity.

Face validity is the way to determine that the test items complete the instructional target and the objective of the test from the appearance. Does the test, on face of it, look like an adequate measure or from the surface of the test, it does not make the test-takers surprised. It is supported by Mcmillan (1997:56) who states that the test seems to a reasonable measure the objectives and domain. It is clearly stated that by observing the form of the test, it can be decided that the test fulfills the instructional target or not.

In addition, Gronlund (1998 : 210) explains that the extent to which the students view the assessment as fair, relevant and useful for improving learning is called as face validity. This theory reveals that the first judgement about the validity of the test can be done by checking the test from the face of the test. Then, determining face validity of a test is done by questioning the student's opinion about the test. for instance, are the test items

relevance with the teaching material in the classroom? Or do you understand the how to do the test and so forth.

Moreover, Mousavi in Brown (2004:26) also emphasizes that face validity is the degree to which a test looks right and appears to measure the knowledge or skills that it wants to be measured based on the objectives of the examinees or institution who take it. This theory explains more about the importance of face validity to provide general information about the validity of the test before doing a deep analysis.

To sum up, validity of the test can be determined by looking the appearance of the test it is known as face validity. Then, the researcher also concludes that face validity is a part of content validity, because it is used to check whether the test items are relevant with the objectives of the test.

METHODOLOGY

The type of the research in this study was classified into descriptive one. This research has a purpose to illustrate the existed condition in the field, in this case the phenomenon happened in one of international standard schools that is SMP N 8 Padang by analyzing the test items used in that school. Then, there were two sources of data in this research: documents of English tests and the respondents of interview. Documents of English tests consist of four kinds; they were English semester tests designed by the central government in 2010/2011 – 2011/2012 academic years , then English semester test designed by the teachers in 2011/2012 academic year and English Mid-test designed by the teachers in 2010/2011 academic year. Then, there were two instruments in this research: interview and evaluation format. For interview, the researcher used structured interview. The followings are the blue-prints both of the instruments.

No	Indicators	Sub-Indicators
1	Content and Face validity	1. Using curriculum and syllabus in constructing the test. 2. Having/designing the blue-print of the test. 3. Knowing about the language skills 4. Knowing how to assess the language skills (listening, speaking, reading and writing). 5. Giving example how to assess one of language skills (e.g speaking). 6. Asking about the category of language skill for item number 1. 7. Asking about the reason of testing speaking skill in written.
2	Multiple-choice format	1. Asking about the definition of multiple-choice test. 2. Asking about the guidelines in writing multiple-choice test. 3. Asking about "stem" and how to design it. 4. Asking about "options" and how to design them. Asking about "distractor" and how to design it.

Then, the researcher used the blue-print of interview and the evaluation format of multiple-choice test items above to collect and analyze the data. For the interview, the researcher interviewed the teachers and the evaluation format was used to analyze the multiple-choice items format. In addition, analyzing the English tests needed one more data that was the analysis of content validity. To analyze it, the researcher checked the curriculum and the blue-print of the tests. After analysing the data, the researcher interpreted the data. For analyzing the multiple-choice items format, there were three classifications. Perfect item; less perfect item and imperfect item. Perfect item was the item that could fulfill all the criteria stated in evaluation format. Then, less perfect item was the item that could fulfill some of the criteria and imperfect item was the item that could not fulfill all the criteria. Meanwhile, for content validity, it had two classification; valid item and invalid item. Valid item was the item was relevant with the curriculum demand and the blue-print of the test, and then it was vice versa for invalid item. The final result of the multiple-choice format analysis was combined with the result of interview.

The indicators	The sub-indicators
1. The curriculum achievement	1. The exercise assesses an important aspect of the unit's instructional target.
2. The blue- print of the test	2. The exercise matches your assessment plan in terms of performance, emphasis, and number of points.
3. Writing the stem	3. The stem of the item asks a direct question or set a specific problem.
	4. The item is based on a paraphrase (rather lifted directly from a textbook).
	5. The vocabulary and sentence structure are at relatively low and nontechnical level.
4. Writing the alternatives and distractors	6. Each incorrect alternative (foil) is plausible so that it cannot be viewed as absurd or silly by a student who lacks knowledge of incorrect answer.
	7. If possible, every incorrect alternative (foil) is based on a common student error or misconception
	8. The correct answer to this item is independent of the correct answer to other test items.
	9. All the alternatives are homogeneous and appropriate to the content of the stem.
	10. Avoiding use "all the above" and "none of the above" as much as possible.

FINDINGS AND DISCUSSION

Findings

The first analysis was focused on the multiple-choice format. The resarcher used the evaluation format in chapter III to analyze the data. The result of analysis can be seen in the following table.

Kinds of the tests	Category		
	Perfect items	Less perfect items	Imperfect Items
Semester test 2010/2011 designed by Central Government	37	10	13
Semester test 2011/2012 designed by Central Government	15	32	13
Semester test 2011/2012 designed by The teachers		38	12
Mid-Test 2010/2012 designed by the teachers	-	36	14

The findings above show that the English semester test designed by the central government in 2010/2011 academic year has a better format than others. Meanwhile, the Mid-test designed by the teachers in 2010/2011 academic year is the poor one.

Furthermore, the result of analysis for the content validity can be seen in the following table.

Kinds of Tests	The total items	Valid items	Invalid items
Semester test 2010/2011 designed by Central Government	60	37	23
Semester test 2011/2012 designed by Central Government	60	16	44
Semester test 2011/2012 designed by The teachers	50	-	All
Mid-Test 2010/2012 designed by the teachers	50	-	All

The table above shows that the English semester test designed by the central government has a better content validity and the English tests designed by the teachers show poor content validity.

In addition, the result of interview shows that the teachers known about the multiple-choice format, such as writing the stem, options and distractors. However, the problems in the teachers' made test is related to speaking test items in which the curriculum demand asks the students to do oral performance to show their speaking skill, but the teachers did it in written test. However, after doing interview, the researcher found that the teachers also asked the students to do oral performance, but the teachers made the written test because they saw the examples of the tests form the central government and the teachers' association (MGMP). Hence, it can be said that it is not the teachers fault only, but also the central government give an incorrect example. This facts show that the teachers still need more information about curriculum demand and how to construct the test based on the curriculum demand. Then, the government must be careful in writing the test items , because the valid information comes from the government.

Discussion

The findings above show that the test designed by the central government is the better one. However, all the tests have problem. At least three problems found in all the tests. The first one is the appropriateness of test items with the instructional target in this case is curriculum demand, syllabus or lesson plan and the tests does not have the table of specification. Specifically, curriculum demand and the blue-print are related to content validity. Nitko (1996:142) states that one of important skills in crafting the multiple-choice format is the test-makers have to make sure that the items assess specific learning target. The second problem is writing the stem. Many items used incomplete sentence in which the alternatives come in the middle of the sentence. This is not allowed because it makes the stem is not clear and it is difficult to predict the answer. As stated by Nitko (1996:142) if incomplete sentence used, be sure that the alternatives come at the end not in the middle. The third problem is the best item is paraphrased (rather than lifted directly from a textbook). It means that in writing the test item, it is suggested to write the items which have comprehension question rather than quoted directly from the text book. Nitko (1996:145) there are some disadvantages if the items are not paraphrased; (1) the sentence loses its meaning when you take it out of the context. (2) it only encourages rote memory of textbook material rather than comprehension. (3) it only produces awkwardly worded items with implausible distractors. (4) it is very simple for the learners with a superficial understanding to find the answer because they obtain the clues from the item. In other words, it is better to use implicit questions rather than explicit questions.

To sum up, there is no test that has perfect items. Even though, the tests made by the teachers are the poor one, but it is not the teachers' fault only because the result of interview shows that the teachers have enough competence to design the test and knowing the curriculum demand. This is caused by the teachers just imitate the test designed by the central government. Therefore, the problem at SMP N 8 is not only the teachers' responsibility but also the central government.

CONCLUSION AND SUGGESTION

Conclusion

After doing research, it can be concluded that the Mid-test designed by the teachers assumed having contribution to the students 'failure in doing the English semester test designed by the central government in 2010/2011 academic year. However, we cannot blame the teachers only, because in designing the Mid-test, the teachers imitate the test items from the central government. Moreover, the tests designed by the central government are not perfect also. The main problems in designing the tests are: the relevance of curriculum demand, the blue-print with the test items. Then, writing the stem and paraphrasing the items.

Suggestion

Dealing with the information before, it is suggested to the teachers to pay attention to the guidelines in writing the test and do not only copy the available tests. Then, the

teachers should do item analysis: quantitative and qualitative item analysis to know the problems of their tests and directly revise them.

Notes

REFERENCES

Borich,B., & Kubiszkyn, T. 2003. *Educational Testing and Measurement*. New York: John Wiley & Sons.

Brown, H. Douglas. 2004. Language Assessment: Principles and Classroom Practice. New York: Pearson Education, Inc.

Haladyna, Thomas M., Downing, Steven M., Rodriguez, Micheal C. 2002. A Review of Multiple-Choice Item-Writing Guidelines for Classroom Assessment. Accessed on February 11th, 2012 from http://nflrc.hawaii.edu/rfl/.

Hughes, Arthur. 2003. *Testing for Language Teachers*: Cambridge: Cambridge University Press.

Mcmillan, H James & Schumacher Sally. 1997. *Classroom Assessment: Principles and Practice for Effective Instruction*. Needham Heights: A Viacom Company.

Johnson. T. Rogert & Johnson. W. David. 2002. *Meaningful Assessment. A Manageable and Cooperative Process*. Boston: A Pearson Education Company.

Nitko, Anthony J. 1996. *Educational Assessment of Students*. New York: Prentice Hall, Inc.

Weir, J. Cyril. 1990. *Communicative Language Testing*. New York: Prentice Hall.

JIGSAW: Helping Students to Comprehend Linguistics Material in Introduction to Linguistics Class at English Department of STKIP PGRI West Sumatra

Lili Perpisa, S.S., M.Pd.
STKIP PGRI Sumatera Barat
E-mail: liliperpisa@gmail.com

Abstract

This study aims to improve students' understanding of linguistics by using a jigsaw strategy. It is constituted as a linguistic study about language in overall. Language is a communication tool for a person which has several systems. Thus, the language itself is important to be learned and understood, especially for the students of the language. Another reason is the existence of several phenomena that researchers have encountered in the classroom in English language courses of STKIP PGRI SUMBAR. The study of

language that is learned is English or English linguistic. First, English is a foreign language which is not easy to be learned. The second, the material on this language is a complete knowledge that cannot be separated from one another. Meanwhile, the time provided only 2.5 hours per week. Students are expected to understand the concept of language, how to communicate, sounds, words, meanings and applications. Third problem is learning strategies used by lecturers. Lecturer asks students to learn in groups. Each group will present a different material each week. Then the others respond, causing discussion. This strategy is less effective against the student's understanding. Discussion only occurs for students who had studied. Most of them are just silent. Likewise, group presentations, it will dominate only one or two students. The others are just as followers. So in this study the researcher wants to apply the jigsaw to solve the above problems.

Type of research is a Classroom Action Research. It would increase the students' understandings of linguistic of STKIP PGRI SUMBAR. Participants that would be used were the student of English at STKIP PGRI SUMBAR who were taking the course Introduction to Linguistics. The data were collected through observation would also be assisted by collaborator. The instruments to be used are notes, interviews, and tests. The data is analysed by classifying based on the characteristics. Then, the data would be interpreted to see whether the jigsaw can increase students' understanding of linguistics or not.

Keywords: JIGSAW, Linguistics Material, Linguistics Class

INTRODUCTION

Language is a tool for human communication. Through language, one could express their ideas, feelings and thoughts. We cannot imagine how this world without language. Maybe, all that happened was a misunderstanding because the sound meaningless. Thus, this language is very important in everyday life. It also occurs in the world of education, as indicated by Delahunt and Garvey (8) that "language is central to education", meaning languages are as an important role in the world of education because education is the meaning of communication: Communication between students and teachers, students and other students, as well as with people who are in their social life. In addition, the language is also the student's identity. Lastly, language is also an object of a subject. Thus, language is key in transferring knowledge and values in human.

There is diversity of languages in the world. It is sometimes based on educational background, socio-economic, and geographic. Each region has its own language, dialect and different accent. To communicate between others, they use the national language. Likewise, for the world level, people used English as an international language that will be used for existing between countries for that is required for each student / students learn the English language in order to rate the world of globalization. As the students who learn a language, especially English, they must know the roots of the language itself. It starts from knowing the vocabulary, how to pronounce, the process of word formation, the process of getting someone language, grammar, meaning word for word, phrases, clause, sentences, use of the language in the society until conjunction with other scientific language. The insight is called as linguistics.

According to Ipsen, Guido (2006: 11) linguistics is the study of human language with all aspects of existing. This linguistic aspect is very broad in a scope, and then Refnaldy split into two, macro and micro linguistics. Macro linguistic is the study of language itself, such as phonology (the knowledge of dealing with sound and pronunciation, morphology (the study of the smallest part of the language / word / morpheme), lexicology (the study of the process of forming a new word), syntax (the knowledge that examines how the incorporation of the word-for-word, phrase up into a sentence), semantics (the study which deals with meaning) and pragmatic (linguistics which relates meaning with context) and many others. Meanwhile, micro-linguistics is the knowledge of language that has been linked to other disciplines, such as sociolinguistic (linguistics with social culture, dialect, accent, etc.), psycholinguistics (knowledge of language with Psychology) and most recently the neuro-linguistics (language of the psyche). For studying this all takes a long time. for several colleges that have majors languages, above all the treasures of study studied separately and, especially for students who have chosen education for Linguistics, but for those who choose a literature of the language, they learn the language in generally which combines everything. Likewise, those who choose for education language, they learn overall in the dressing courses. So also in STKIP PGRI SUMBAR, English Study Program, the knowledge of the English is learned in one course; Introduction to Linguistics, 3 credits. Although in the course of English STKIP PGRI SUMBAR courses studied linguistic frontage with Introduction to Linguistics, this does not mean that students are only introduced to the existing terms, they remain in introducing with all the knowledge related to the English language. With a time of 2.5 hours per week, this is a challenge for faculty and students to be able to achieve the learning targets at each meeting. Based on the experience of researchers as one of the team teaching this course, research institute, found that there are several obstacles. Barriers that researchers have found in the course of which this is a matter that is too much to be learned in every meeting, while the time is only 2.5 hours / week of 16 meeting. So there are some parts that have not been discussed in the classroom or abandoned due to lack of sufficient time.

To overcome this, the technique has been carried out by means of discussion groups. One group each week to present the material in front of the class, the other listened. This way is also less efficient, because only the group as a speaker who reads the material while another group was silent. If no one has responded or asked, the question is not the existence of the analyst is looking or understanding of the material in question. Finally, more lecturers who asked to explore the ideas of the group. Such problems also occur in groups of presenters, between 4-5 students, not all of them are able to respond or explain the material appropriately. Diligent student, she/he would dominate the discussion. Thus, one-way communication (one-way communication) appears between the questioner with the answer. The other group members as if just go with because they did not understand the materials that are discussed. In the end when in reviewing the material again, both orally and in writing, many students are not able to answer or describe. This is because they do comprehend the material well though this course is very important because it is also related to other subjects, like pronunciation and grammar.   Related to the above phenomena, it required a solution to overcome these linguistic subjects as it is very useful for life to communicate. Therefore, researchers intend to apply the jigsaw strategy in helping students to understand the material linguistics in English language courses of STKIP PGRI SUMBAR.

REVIEW OF RELATED LITERATURE

Linguistics

Linguistics is the study all about language. According to Ipsen (2006: 11), "linguistics is the scientific inquiry into the human language with all its aspects. All its aspect: These are many ". He explained that the basic knowledge of linguistics is to learn all aspects of the human language. Aspect is very much proficiency level. It is also similar to other experts, such as Loretto Todd. Tod (2005: 5) also found linguistics as "the scientific study of language."

All linguistic experts have the same opinion that the study linguistic has many things that should be known;

a. Definition of the language and its characteristics.
    On this topic, it discusses various concepts of the language and the characteristics of the language itself which language is a communication tool for people who are systematic and productive.
b. Human communication and communication the animals
    Here, in linguistics, the way people communicate with animals is different. English only possessed by humans, while animal sounds produced by animals is merely a voice, not the language.
c. Phonology
    Phonology is the study of how to produce sound, place of articulation or organ - the organ for speaking. Part of this study is called phonetics. Organs are functioning within the study penology; jaw, lips, teeth, tongue, palate, uvula, vocal cords, nose, etc. While the manner of articulation in English in familiar with; plosive, fricative, trills, laterals, nasals, affricative, semi-vowels. Place of articulation in English is called as bilabial, labiodental, alveolar, palate, palatal, and velar (Todd). Additionally phonology also studied English vowel sounds and consonant.
d. Morphology
    Morphology/morphology is referred to the study of morphemes or the smallest part in the unit of grammar or grammar in English.
e. Lexicology
    Lexicology in English linguistics is the study of words and the formation of new words. It is also about the class or kind word, such as verbs, nouns, adjectives, adverbs, conjunctions, and many others.
f. Syntax
    If the lexicology patronizing words, the syntax related with how the incorporation of the word - the word into phrases, clause, and sentences.
g. Semantics
    Semantic relates to the meaning of the words, phrases, and in clause sentence.
h. Linguistics with other studies, such as psycholinguistics, sociolinguistics, and applied linguistics in the teaching process.

Knowledge - the knowledge relates to the above linguists also the subject matter at introduction to Linguistics in English education study program STKIP PGRI WEST

SUMATRA. In this study the authors will try jigsaw strategy in improving students' understanding of linguistics in the English materials.

Jigsaw Strategy

Jigsaw is a cooperative learning strategies (cooperative learning). Thus, in the application of this jigsaw strategy takes small groups in the classroom. Some experts are promoted on jigsaw strategy is Ann Browne in his book Developing Language and Literacy 3-8, Mel Silberman, and others.

According to Silberman (2003: 111) Jigsaw is a strategy that groups will change. The important thing in this strategy, one student teach something, he/ she also confirmed that the material is suitable in this jigsaw is using the parts of the materials. This is also confirmed by Browne (2009: 12); "The children work in groups on the class topic. Each member of the group is allocated one specific aspect of the topic to find out and is designated as an expert on this. The next stage is for all those who were designated experts on the same topic to leave Reviews their original group and meet together to discuss what they discovered. The children then return to Reviews their original group and reports on their child Click or call now part of the group investigation." Brown also explained that it should be done with a jigsaw strategy group work. One group has a different section or topic to another group. They will discuss the initial group. After that they will be merged into a new group, where each group consists of all existing topic. Each member of a different group will report the results of their discussion. Thus, those all students receive information about all the topics of the day.

For the procedure, Silberman also explains in his book Active Learning with multiple steps;
1. The material is divided into several sections
2. Calculate the number of students, for according to the amount of material that has been divided. For example, the number of students 12, the material can be assumed to be three, then students at the three as well. Then ask students to read, discuss and understand the material in their respective groups
3. After some time, the form of the "Jigsaw learning". Ask the students to count in the group, 1,2,3,4. After that create a new group based on the same number. Then there will be four groups consisting of groups with 1, 2,3 depend on different topic. This can be described as number of students, study group, cooperative learning groups.
4. After that, students are asked to explain what they have learned in the previous group.
5. Finally, review all the material and do Q & A with the students to see their understanding of the material. While Glosgow and Cathy (2003: 3) explained there are six steps in the application of the jigsaw in the classroom, minister;
   1. Divide the material into 5 parts
   2. Then, divide students into five groups, for example the number of students in the class of 25 people, each group containing 5 students. Here, they confirmed that each group should be heterogeneous both seen from the background, gender and level of achievement that they already can. Thus, each group of students should not contain the same good level as well as gender. This group is called the base group.
   3. Each group gets a different part of the materials for them to understand.

4. After that, members of the base group is divided, each joined with other groups. This group is called "expert group" because they are experts in different topics. They study and discuss on the topic that they have brought from the base group.

5. Students return to the base of the group to discuss what they can from the expert group.

6. Teachers will check the comprehension of the students about what and how much they understand.

Overall, both procedures above have a common step in the use of the jigsaw. Silberman explains the theory is not just a phrase; he also describes the figure easier to understand. While Glosgow and Cathy describe without the map, they are also very clear explanation. The most important thing in their theory is that they assert in terms of the division of the group so that each group consists of the diversity that exists in the classroom. Thus, in this study, the researcher will use the steps outlined by the experts above.

METHODOLOGY

The method used in this research was Classroom Action Research (Classroom Action Research) because this study seeks to improve the teaching and learning process in the classroom. Creswell (2012: 577) argues that action research is research conducted by teachers to raise issues that are in the class, and then find the solution that is aimed at the achievement of a better. Creswell is in line with the opinion of Cohen (2005: 226) argues, that the class action is a tool or a very powerful way to change and improve quality at the local level. In this study, researchers focus on improving the ability of students' understanding of the linguistic materials in English language courses STKIP PGRI SUMBAR.

In this study, researchers would also collaborate with another lecturer who teach introduction to Linguistics as well as described by Kemmis and Mc.Taggart in Cohen, that this study is a collaborative research.

Procedures for Research

This study had four phases as described by Kemmis and Mc.Taggart in Creswell (2012: 577), namely; (1) Planning, (2) Acting, (3) Observing and (4) Reflecting.

1. Planning
   This stage is the planning, where researchers will analyze the problems in the classroom, then planning actions would be applied to solve the problem.

2. Acting
   Action at this stage is the application of the jigsaw strategy on subjects Introduction to Linguistics.

3. Observing
   At this stage, researchers would observe the actions that have been granted. This step was key to the next stage. This observation was not only done after a course of action, but from the beginning of the action took place for record everything.

4. Reflecting
   Final stage, the team of researchers would analyse and interpret the data gathered in order to deduce what was happening in this process. The result of this process was the input to perform an act further.

Based on preliminary observations on the first and second meetings, the researcher obtained the information that students are still having problems in understanding the linguistic material. This case appeared when the students had discussing the first and second materials which were about language and Communication, Human and Animal Communication. The material was about the nature of language of some linguists, the nature of human communication and communication differences and communication of animals. When discussing, only three students who wanted to talk to express an opinion about what the language was. Then when asked about animal communication, all of the students stated that the animals have a language. By using examples and analogies supposed that animals had the language, the researcher asked what the language of cats which come from Indonesia use with what the language of cats from Australia use. All the students were not able to answer. Therefore, it came to the conclusion that it was agreed that animals do not have language but sound. However, in the pre-tests perform, only one person who replied that the animals do not communicate with the language but through sounds and other means such as gestures, smells, colours, etc. The others were still the same answer when the researcher asked before getting the material. One of the factors they were not the focus of the current discussion, there was also a serious face but dispersed, thinking of others. As for the questions on the first test;

1. Explain about the concept of language based on your understanding!
2. What is communication?
3. Do you think that animals have a language? Why?
4. Differentiate how human and animal Communicate?
5. Please, explain about your language in your daily life!

From five questions only question No. 1 was answered by the students with the right but not on the basis of their own language, but based on the quote experts without discussing with their language. Question No. 2 is more than 50% responded beat in English yet all of these had been implemented in everyday life. It was unfortunate that the question no. 3 only two students had the correct answer where in the discussion they have been agreed that animals do not have language, but sound. Question No. 4 and 5 did not answer them more. Some people seem perfunctory answer.

Thus, for the first cycle, the researcher had design to solve those problems;

A. Cycle I
    a. Planning
        1. Divide the material into sections in accordance with sub heading.
        2. Ask the students to have all the material that would be discussed on the day.
        3. Divide students into groups according to the amount of material
        4. Give time for discussion on each group in accordance with their respective material
        5. Then merge them into a new group consisting of groups or earlier origin.
        6. Ask them to discuss
        7. Provide several questions related to the material that they had discussed.

Based on the syllabus or Lesson Plan One semester (RPKPS), the third meeting discusses Phonetic and Phonology. Phonetic is the study which deals with sound by sound. In this case, it also talks about the organs that produce sound and how sounds are pronounced in these organs. To improve students' understanding of these materials, the material at the first meeting in the first cycle of phonetic divided into 4

parts, Articulatory Phonetic, Organs of Speech, Manner of Speech and Place of Speech. While in the second meeting is divided into four parts namely, the material on phonology study which deals with sound, what is phonology, phonemes, all phones, English and Consonant phonemes in English. Meeting to discuss the three cycled the vowels in English. Students are divided into 4 groups. Each group gets vowel sounds. After that, the researchers tested their understanding of the material.

b. Action (Action)
    In this section, there were several actions in conducting discussion.

c. Observation
    Observations made by the lecturer (research) in the classroom in several ways;
    1. Looking at the current students during discussion through the records of all the activities by the students
    2. Provide questions relating to the material on each of lecturing
    3. giving the test for the third meeting.
    Based on the observation of the activity of the students during the discussion, a lot of students talked in each group. On average, there was one person in every group that dominates. One group of the group IV, all students were passive students. When the lecturer asked to switch to the new group, lot of students were concerned about their seats noisily. Thus, the misunderstanding happened.

    Other observations were the first discussion with time of 30 minutes which was not enough for students to discuss because the first 10 minutes they still look confused. 75% of them were working on their own to understand the material and record it into the book. The others just listened and recorded sometimes recorded in accordance with the book source.

    When melted to new group, they seemed to be surprised because it was responsible for explaining the material they discussed earlier to the new group. The new group was forced to re-read the material before the group. Given 5 minutes is not enough. On average they use 10 minutes to deliver it to the new groups because they reread again.

    At each end of the discussion, the students were asked to explain the material obtained has not been able to explain in their own language. In average, everything just re-read a book or notes. At the 3rd meeting because less material and they knew there would be a written test, then the concentration of dispersed discussion started they did not listen to their friends who shared their reading material but the material previously because there was no repetition learning at home and when the discussion at the previous meeting they only focused to the book.

    In tests at the third meeting with three questions;
1. Explain each items of Phonetics, phonemes and allophones?
2. Write the phonetic transcript of this song?
3. Differentiate of sound below based on Manner and place articulation?
    Based on the rank values used here, there are an A = 81- 100, B = 66-80, C = 56-65, D => 55, E => 45. So, 1 student has a grade of B, 5 C, 15 D, and E. Thus, 7 people 22 people or more 50% can be said to fail or do not understand this material. Meanwhile, still low 5 people understanding and a person just understood it.

d. Reflection

Based on the above observations on the first cycle, discovered the problems that must be resolved in the next cycle;

1. Problems of time.
   The first discussion with a time of 30 minutes, it's still not enough for the students. Discussion into two new groups is still lacking which each member of the group gets 5 minutes to deliver results.
2. Problems classroom management
   Classes can be noisier when the exchange group from the origin moves to the new group.
3. When the discussions are still many who have not been able to express their opinions, so that they receive from one person, who just recorded or still passive.
4. There is no fear when sharing with a new group, fearing that one of the express.
5. At the third meeting of the student is more concentrated on the face of the test compared to the discussion.
6. From the results of the first test seen many students who are not able to Understanding the material that were being discussed.

From the problems above, the students still require other measures that will help them improve their understanding of the linguistic material. As a lecturer who teaches this course required more attention to students when discussing and teaching material deeper in the end of lecturing. So it can be concluded that the first cycle has not been able to reach the targets in expect it then, researchers (faculty) conduct its review and revise to the next cycle.

B. Cycle II

Second cycle executed three meetings as well. The material discussed at the meeting was about the morphology and lexicology. Based on the reflections in the first cycle, there was still the issue of students' understanding of the linguistic material both in terms of process and in terms of results, then it is designed for II plan as follows;

a. plan
   1. The division of the group is no longer based on the order list of attendance; the first group was too large. Then, it is subdivided into smaller groups so that there are two groups of materials.
   2. Time to discuss of the first group was added to 40 minutes, and the second time the group each member of the group to 7 minutes
   3. Then, at the turn of the group, the lecturer has set seating instruct each group to move the group without a voice.
   4. Lecturer more monitors every discussion group
   6. At the end of lecturing, the lecturer also opened question and answer session or reviewing material using a tool if they were bored like projectors
   7. At the 3rd meeting of no more material there is only discussion or review and test

With the improvement in the second cycle, the students were very enthusiastic, because at the first meeting of the group division is done by way of fun based on the

colour of clothes, clothes that are red and pink merge into one group. Then the black shirt and grey into one group: shades of green clothes group. Blue is also a group. And the trim colour to yellow or orange into one group. Incidentally remaining colour of clothes many colours there were 5 students, then they merged into one group.

While the second meeting of the group division is divided Based on the month of birth. It is also beyond their predictions. Because they predict the division would be the same group based on colour, too. Looking at the distribution of the variation of the group they were very excited and happy because there were comfortable stuffs. The activities carried out with action.

b. action

There were several actions in conducting research such as discussion with the steps directly.

c. observation

Based on previous actions, we conducted observations of the process and results of the study of linguistics. In terms of the process has shown the improvement. Students already feel they have the responsibility for the task to understand the material obtained in the first group. Thus, they feel should be able to master the material. Growing sense of responsibility makes them seriously in a discussion on the origin of the group.

C. Cycle III

In this step, the researcher found students' increasing in comprehending the linguistics material, especially about syntax. The students are able to write the diagram tree of the sentences.

The sentences are given by the researcher have different tenses. In here, the students with their group discussed how to draw each diagram tree.

Based on the result of the test, it shows that there is improvement in their score so that we can say this strategy can improve students' comprehension.

CONCLUSION

Of the two ways that the researchers describe, it is reflecting the result significantly to the learning process and the results of student understanding in the course introduction to linguistics. Suggested to other educators, strategy this jigsaw can be applied to all levels of education and the subjects of any particularly subjects that contain solid material.

REFERENCES

Browne, Ann. 2009. *Developing Language and Literacy 3-8.*London: SAGE Publication Ltd

Cohen, Louis, Lawrence Manion and Keith Morrinson. 2005. *Reseach Method in Education.* Canada: Taylor & Francis e-Library.

Crane, L. Ben, Edward Yeager, and Randal L.Whitman. 2003. *An Introduction to Linguistics* Boston: Little Brown and Company

Creswell, John W. 2012. *Educational Research.* Boston: Perason Education

G. Cohen, Elizabeth, Celeste. M Brody and Mara Sapon-Shevin, Ed D. *Teaching Cooperative Learning in: The Challenge for Teacher Education* :Sunny Press

Gould Lundy, Kathleen. *What do I do about the Kid who…?. 50 Ways to Turn Teaching into Learning.* 2004. Ontario: Pembroke Publisher

Glasow, A Neal, Cathy D Hicks. 2003. *What Successful Teacher do.* Corwin A Sage Company.

Ipsen, Guido

Johnson, Andrew P.2005. *A.Short Guide to Action Research.* Indianapolis: The Bobbs-Merril Company, Inc

Refnaldy, dkk. 2007. *Introduction to Linguistics.* Universitas Terbuka. Diknas: Jakarta

Siberman, Mel. 2003. *Active Learning. 101 Strategies to Teach Any Subject.* Massachusetts: A Simon & Schuster Company.

Todd, Loretto. 2004. *An Introduction to Linguistics.* Singapore: Longman Singapore Publisher Pte Ltd