DEVELOPING FORMATIVE ASSESSMENT ON RECOUNT FOR THE TEN GRADERS OF SENIOR HIGH SCHOOL 11 SEMARANG

A THESIS

In partial fulfillment of the requirements

for Master`s Degree in Linguistics

Ignatius Maryoto

1302021 1400006

FACULTY OF HUMANITIES

DIPONEGORO UNIVERSITY

SEMARANG

2014

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

A THESIS

DEVELOPING FORMATIVE ASSESSMENT ON RECOUNT FOR THE TEN GRADERS OF SENIOR HIGH SCHOOL 11 SEMARANG

Submitted by:

Ignatius Maryoto

13020211400006

Appoved by

Advisor

Dr. Suwandi, MPd.

Master Program in Linguistics

Head,

Dr. AgusSubiyanto, M.A

NIP. 19640814199001101

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

A THESIS

DEVELOPING FORMATIVE ASSESSMENT ON RECOUNT FOR THE TEN GRADERS OF SENIOR HIGH SCHOOL 11 SEMARANG

Submitted by

Ignatius Maryoto

13020211400006

VALIDATION

Approved by
Thesis Examination Committee
Master program in Linguistics Postgraduate program at Diponegoro University
Faculty of Humanities Diponegoro University
On … July 2014

Chairperson
Dr. Suwandi, M.Pd. ----------

First Member
Dr. AgusSubiyanto, M.A. ----------

Second Member
Dr. Deli Nirmala,M.Hum. ----------

Third Member
Dr. Suharno, M.Hum. ----------

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

CERTIFICATION OF ORIGINALITY

I hereby declare that this submission is my own work and that, to the best of my knowledge and belief, this study contains no material previously published or written by another person or material which to a substantial extent has been accepted for the award of any other degree or diploma of a university or other institutes of higher learning, except where due acknowledgement is made in the text of this thesis.

Semarang, July 2014

Ignatius Maryoto
13020211400006

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

ACKNOWLEDGEMENTS

First and foremost, my praise to God the Almighty and merciful, who has given strength to me so this thesis entitled `Developing Formative Assessment on Recount for the Ten Graders of Senior High School 11 Semarang` came to a completion.

I also would like to express my deepest gratitude to:

1. **Dr. AgusSubiyanto, M.A, the Head of Master program in Linguistics at Diponegoro University Semarang.**
2. **Dr. Deli Nirmala, M.Hum, the Secretary of Master program in Linguistics at Diponegoro University Semarang.**
3. **Dr. Suwandi, MPd.the advisor. Thanks very much for his precious guidance, advice, suggestion, and continuous motivation until this thesis is completed.**
4. **Hartoyo, M.A, PhD, the expert on assessment who has time for validating this thesis.**
5. **My wife, Lucia Emi Wiharjanti, SE and my beloved son JuventiusWahyuUtama, SE who have prayed and supported me a lot. I also would like to give my special gratitude to my big families in Temanggung, Hj. Immanah and my late father, F.X. Marsoedi. My wise parents – in- law, Sugito and Mardiati.**
6. **My brothers M. Maryanto, J. Marjadi, my sisters: , CH. Maryati, Th. Maryanah and my late sister, M.M. Martini , thank you very much for your support.**
7. **All lecturers of Masterprogram in Linguistics at Diponegoro University Semarang.**

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

8. **The Headmaster of SMA N 11 Semarang and all teachers, administrative staff and cleaning staff. Thank you for the tolerance as long as I studied at Master`s Program in Linguistics Diponegoro University.**
9. **All my friends in Magister Linguistics, Mrs. Raeda, Mr. sulis, Mrsutomo who help me much.**
10. **Much. Ahlis, WahyuSetiaBudi as the administrative staff of the Master program in Linguistics at Diponegoro University.**

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

Motto

Nothing vantuyres, nothing wins
In the Chriust ways

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

DEVELOPING FORMATIVE ASSESSMENT ON RECOUNT FOR THE TEN GRADERS OF SENIOR HIGH SCHOOL 11 SEMARANG

ABSTRACT

This study aims at developing formative assessment on recount for the ten graders of senior high school 11 Semarang in the academic year 2013 – 2014. It also attempts to find out the effectiveness of the developed formative assessment on recount based on syllabus for the ten graders of senior high school. This study was conducted by using research and development (R&D), with the students of senior high school 11 Semarang as the sample.

80 English teachers in 16 state senior high schools in Semarang municipality are chosen as population. Whle forty English teachers from 16 state senior high schools were taken as samples in this research. In preliminary research, the researcher conducted need assessment to the 80 English teachersin 16 senior in Semarang municipality.WhileFortyof them were chosen as sample.They were chosen on purpose. They were selected since they were not only active in their schools, but also participate in English teachers` union. The writer chose two or three English teachers of every school. While for the students, the population was 11 classes of the first grade of senior high school 11 Semarang. The writer selected thirty five students of class X IPA 4 in senior high school 11 Semarang as the sample of this research. The reason is because the writer teaches them once a week.

There were thirteen steps taken for developing formative assessment on recount for the ten graders of senior high school 11 Semarang. The first step was the need assessment for 40 English teachers in 16 state schools of Semarang municipality. The second was analyzing curriculum 2013 as stated by the government regulation number 69, 2013. The third was choosing the core competences from curriculum 2013 whichare related to recount.The fourth was selecting the basic competences from the of syllabus curriculum 2013 related to recount. The fifth was writing test material which covered expressions and grateful, analyzing social function of recount, and linguistics elements of recount. The sixth was writing indicators. The seventh was writing the draft one of formative assessment on recount. The eighth is revision by English teacher. The ninth is writing draft II. The tenth is revision by a senior English teacher. The twelfth is draft III was validated by Hartoyo, M.A., PhD. Thirteenth is formulating final product.

The instrumentation of this research was the questionnaire of need assessment for grasping data whether the English teachers of stateseniorhigh schools in Semarang municipality needed developed formative assessment by English teachers or not. The result was 82.5 % ofEnglish teachers neededdeveloped formative assessment made by English teachers made.

The result on try out I, II, and III were analyzed by using SPSS 20. It showed that the mean score of try out II was greater than the mean score of try out 1(78.0 > 72.0). It also showed that $t_o$was > t.table (2.881 > 2.03 > on the level of significant 0.005. The mean score of try out IIIwas the greatest of all (81.4 > 78.0 > 72.0). It can be concluded that the development of formative test on recountwas proved effective based on the limited try out.

Key words: research and development, formative, assessment.

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

TABLE OF CONTENTS

TITLE i
APPROVAL ii
VALIDATION..... iii
CERTIFICATION OF ORIGINALITY iv
ACKNOWLEDGEMENTS v
Motto vi
ABSTRACT..... vii
TABLE OF CONTENTS viii

CHAPTER 1 INTRODUCTION
1.1 Background of the Study..... 1
1.2 Research Problems..... 2
1.3 Purposes of the Study 2
1.4 Significance of the Study 3
1.5 Terms in this Research 3
1.6 The Organization of Writing 4

CHAPTER II REVIEW OF RELATED LITERATURE
2.1. Previous Studies 5
2.2. Measurement, Evaluation, Assessment, and Test 6
2.3. Formative and Summative Assessment 6
2.4. Reading in a Second Language..... 9
2.5. What Works in Reading Instruction..... 10
2.6. Implication for Assessment 12
2.6.1 Figure Reading Assessment Matching Purpose to Task 13
2.7. Authentic Assessment of Reading..... 14
2.8. Language Competences..... 14
2.8.1 Canale and Swain Model of Communicative Competence 14
2.8.2. Bachman Model of Communicative Language Ability (CLA) 16
2.8.3. Interactional Competence 19
2.9. How to Construct Formative Assessment..... 21

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

2.10. Test Specification and Design 25
2.10.1. Planning in Test Authoring..... 27
2.10.2. Guiding Language Versus Samples..... 28
2.10.3. Congruence (Fit-To-Spec)..... 29
2.10.4. How Do Test Questions Originate?..... 29
2.10.5. Reverse Engineering..... 30
2.10.6. Spec-Driven Test Assembly, Operation and Maintenance 32
2.11. Curricullum 2013..... 33
2.11.1. Characteristics of Curriculum 2013..... 33
2.11.2. Basic Curricullum Framework..... 34
2.11.2.1. Philosophical Basis..... 34
2.11.2.2. Theoretical Basis..... 35
2.11.3. Curricullum Structure..... 36
2.11 3.1. Core Competencies 36
2.11.3.2. Core Competence and Basic Competence Curricullum 2013..... 37
2.12. Indicators of Assessment on Recount..... 37
2.13. Bloom Taxonomies..... 38

CHAPTER III RESEARCH METHODS
3.1. Research Design 39
3.2. Population/Samples..... 40
3.3. Instrumentation..... 40
3.4. Data Collection..... 40
3.5. Data Analysis..... 41
3.6. Revision of Formative Assessment 41

CHAPTER IV FINDINGS AND DISCUSSION
4.1. The Result of the Need Assessment (I)..... 42
4.2. The Result of the Existing Formative Assessment on Recount (II)..... 42
4.3. The Result of Model of Developing Formative Assessment on Recount 44
4.4. Revision of Formative Assessment 48
4.4.1. First Revision by Dian Arini Rosita, SPd 48
4.4.2. First Revision by Dra. TheresiaMelaniaSudarwati, MSi..... 49
4.4.3. The Formative Assessment Items were validated by Hartoyo, MA, PhD 53

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

4.5. The Old Versus the Developing Principles in Constructing Formative Assessment 64
4.6 The result of try out 66
4.7. Findings Interpretation 68

CHAPTER 5CONCLUSION AND SUGGESSTION
5.1 Conclusion 69
5.2. Suggestions..... 71
5.2.1. Suggestions for the English Teachers..... 71
5.2.2. Suggestions for the next researcher..... 71

REFERENCE 73
APPENDICES 77

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

LIST OF TABLES

Name **Page**

1. **Core and Basic Competence 77**
2. **The Result of Try Out 1, Try Out II, and Try Out III 77**
3. **Appendix 1. Final Product on Developing Formative Assessment on Recount 78**

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

CHAPTER I

INTRODUCTION

In this chapter the writer focuses on; (1) the background of the study, (2) the research problems, (3) the purposes of the study, (4) the significance of the study, (5) scope of the study, and (6) the organization of writing.

1.1 Background of the Study

The study aims at developing formative assessment on recount for ten graders of senior high school 11 Semarang in the academic year 2013 – 2014. It also attempts to find out the effectiveness of the developed formative assessment on recount based on curriculum 2013 for the ten graders of senior high school.

There are two main reasons which become the background of writing the thesis. First, formative assessment is a necessary test as it assesses the activities that are being processed.However, not many tests assess the achievement which is being processed. In fact, all English teachers should be able to compose their own formative assessments, so they do not absolutely rely on the students`worksheets made by the publishers.

Second, the use of formative assessment will support the summative assessment which measures the students` achievements at the end of the semester. Formative assessment measures the ongoing capability of the students. conssequently, the students will remember what the lessons are taught. By gaining good marks in their formative assessments, consequently. they will have more and more knowledge. If the students master their all formative assessments well, they will have more understanding on the summative test.Therefore, when the students master their basic competences, they will automatically have competences on summative test. As stated by Brown (2007:4) that summative

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

assessment aims to measure or summarize what a student has grasped and typically happens at the end of a course or unit instruction.

1.2 Research Problems

Based on the background above, there are three proposed questions in the research problems:

1) How are the existing formative assessments constructed by English teachers in state senior high schools in Semarang municipality?
2) What kind of model of formative assessment on recount is developed?
3) To what extent is the effectiveness of formative assessment on recount applied in senior high school 11 Semarang?

1.2 Purposes of the Study

The purposes of the study were:

1) To present the existing formative assessments constructed by English teachers in state senior high schools in Semarang municipality.
2) To present the model of formative assessment on recount.
3) To find out the effectiveness of formative assessment on recount applied in senior high school 11 Semarang

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

1.4 Significance of the Study

The findings of this study give some benefits to be as follows:

1) Theoretically, the result of the research can as the basus of the English teacher due to limited time for teaching.

2) Pedagogically, the result of this research can contribute to the English teachers for the improvement of teaching and learning English.

1.5 Terms in this Research

In order to avoid the wrong understanding, the writer illustrates each term applied in this study. They are as follows:

1) Development is a new product or idea/ the process of developing or being developed *(Concise Oxford English Dictionary Eleventh Edition)*. Developing, of course, is not something new, rather dates back socializing some available material by adapting, modifying, restructuring, and simplifying modern concepts along with students' cultural background.
2) Formative assessments written by Huges Arthur (2003:5) stated that an assessment when teachers use it to check on the progress of their students, to see how far they mastered what they should have learned, and then to use this information to modify their future teaching plan.
3) Recountis a genre in the curriculum 2013. A recount according to Anderson Mark and Anderson Kathy (2003: 48) is a piece of text that retells past events, usually in the order in which they happened.

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

1.6 The Organization of Writing

This thesis contains five chapters. Chapter one presents the background of the study, research problems, purposes of the study, significances of the study, the scope of the study, and the organization of writing.

Chapter two accounts for review of related literature which consists of four main chapters: previous studies, relevant research, curriculum 2013, and the theories of assessment. The previous studies refer to some related research about formative assessment, especially taken from some journals. There is the review of related literature of the study which shows the curriculum in senior high school or junior high school, syllabus, lesson plan, basic competence of recount, the theories of assessments, theories of how to compose good assessment.

Chapter three discusses the research design by applying educational R & D (Gall and Borg. 2003:569), population and samples, instrumentation, data collection, and data analysis. The researcher elaborates the process of corrections by two English teachers and validation by an expert on assessment indetail.

Chapter four provides the research findings. Discussing in this chapter is the result and analysis of the try out of formative assessment on recount and narrative texts based on syllabus, the analysis of the material development.

Chapter five presents conclusion and suggestion. The conclusion is drawn referring to the finding of this study presented in the previous chapter. The suggestion deals with several recommendations for the use of formative assessment.

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

CHAPTER II

REVIEW OF THE RELATED LITERATURE

2.1. Previous Research

Rin (20006) focused her thesis on the correspondence between the final test and English program objectives at SD Muhammadiyah I Surakarta. She did not mention the requirement of writing good assessment. She concluded that the English teacher made some errors in constructing the English tests, consequently the tests are not appropriately used to identify whether or not the English teaching objectives based on the 1994 curriculum of elementary school are achieved. The teachers should be able to choose appropriate English test type since certain types are more appropriate for certain objectives of teaching English. In this case, it is worth knowing, whether it is more appropriate to give oral or written test, direct or indirect test, subjective or objective test, etc.

While Umiyatun (2010) discussed her thesis on the problems of writing recounten countered by the students of the state junior high School 2 Purworejo in the academic year 2009/2010. She then concluded that thegrammatical problem on recount writing of eighth graders of SMP N 2 Purworejo 2009 – 2010 is as past tense: 124 problems or 24,41 %. She did not discuss developing writing assessment.

Both Rini and Umiyatun discussed recount, but Rini focused on the correspondence between the final test and English program objectives at SD Muhammadiyah I Surakarta. Umiyatun concentrated on the grammatical problems on recount. On the other hand, the researcher developed formative assessment on recount for the first graders of senior high school 11 Semarang.

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

2.2. Measurement, Evaluation, Assessment, and Test

Nitko (1983:5) states that measurement refers to quantitative aspects of depicting the attributes of persons.Calmorin (2004:15) formulates measurement as a device to measure individual`s accomplishment, personality, attitudes, intelligence and among others that can be declared quantitatively. The other definition is by Bechman (1997:18) stating that measurement is as the social sciences in the process of quantifying the characteristic of persons according to certain procedures and rules. For another other point of view by Gonzales and Calderon (2007:6) as quoted by Hartoyo (2011:2) defines measurement as the process of deciding the quantity of achievement of learners by means of suitable measuring instruments.

Another term connected to language assessment is evaluation. Bloom et al. (1971) as quoted by Hartoyo (2011:3) definethat evaluation is the systematic gathering of proof to determine whether factcertainchanges are occurring in the learners as well as to determine the amount or degree of change in individual students. Bechman (1997:22) defines evaluation as the series of collecting information for the aim of making decision. Calmorin (2004:18) determines that one of the scope of evaluation is assessment of students. It means that the students should be assessedto decide whether they obtain the aim of learning tasks.

Brown (2004:4– 6) states that tests are administeredadministrative procedures that happen at identifiable times in curriculum when learners master all their capabilities to offer the peak performance, knowing that their response is being assessed.

2.3. Formative and Summative Assessment

Brown (2004:6) states that there are two kinds of assessment: formative and summative assessment. Formative assessment means evaluating students in the procedureof "forming" their competencies and skills with the aim of helping them to continue that growth process.

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

Brown also stresses on the delivery by the teacher and the internalization by student of appropriate feedback on performance, with an eye toward the future continuation or formation of learning. The implication in class is when a teacher declaresa student a comment or a suggestion or call attention to error, that feedback is submitted in order to increase the learner`s language ability. The conclusion is that all practical purposes, practically all kinds of informal assessment are formative.

Meanwhile, summative assessment has a goal to assess or summarize, what a student has obtained, and typically happens at the end of course or unit. A summative of what a student has studied implies looking back and taking stock of how well that pupil has accomplished goals.

In other words, the writer thinks that a well-constructed test is an instrument that supplies a proper measurement of the test-taker`s ability within a particular domain. Both science and art are needed to compose a good test.

Assessmentingeneraliscloselyassociatedwithlanguageassessment,whichcanbeaccomplishedthroughlanguage test. A testis a method of measuring a person's competence or knowledgeon a given domain (Iseni, 2006). Iseni (2006) formulates the differences between traditional assessment and authentic assessment. Traditional assessment is evaluation that belongs to standardized and classroom achievement tests with different types of item. While authentic assessment is defined as fresh method to assess that reflects pupils learning, achievement, motivation, and attitudes on instructionally-relevant classroom activities.

In the learning process, teachers sometimes are confused to differentiate between informal and formal assessment. Brown (2004:5) stresseson that formal assessments are exercises or processes specifically arranged to strikeinto a storehouse of skills and knowledge.

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

Asgher (1999:205-223)states that higher institution policy is dominated by summative assessment regulation, with little emphasis on the role of formative assessment to improve the student lecture. On the other hand, Bennet (2011:5-25) finds that the conceptualization should also allow the substantial time and professional support need of majority of teachers to become proficient users of formative assessment. While Cowie&Bell(1999) conclude that formative assessment is defined as the process used by the teachers and students to recognize and respond to student learning in order to enhance that learning. According to Hodgson (2012:215-225) draws a conclusion that the process of formative assessment in universities has the potential to engage the students in reflection and to take greater ownership. Vickkerman(2001:221-230)concludes that the study found that the whole formative peer assessment was a positive in enhancing students learner and development. Bookhart (2011)concludes that these successful students engaged in self assessment as a regular, on going process and actively tried to fit the new in information about learning into the career as students` assessment. Macdonald (2004)states that paper discusses thepractical implication online pedagogies and illustrates the powerful formative effects, both intended and unintentional assessment on student learning and behavior.

The conclusion is that in teaching process there are ongoing assessments. Brown(2004:1-15) states that assessment is ongoing process that comes forward to a larger domain, while a test part of assessment which is provided administrative procedures that occur at identifiable times in a curriculum when pupils master all their capabilities to indicate peak performance, know that their responses are being assessed and examined. In other words, tests are subsets of assessment.

The writer agrees that some tests measure general ability, while others focus on very specific competencies or objectives. In some cases, a test measures an individual `s ability,

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

knowledge, or performance. The writer thinks that in this test, testers need to recognize who test-takers are.

The writer is in line with Brown that a test measure performance, however, the results denote the test-takers` ability, or to utilize concept common in the field of linguistics, competence. Finally a test measures a given domain. In the case ofproficiency test, the actual performance on the test involves only a sampling of skills, that domain is overall proficiency in a language – general competence in all skills of language.On the other hand, the writer stresses on assessment in class. Brown (2004: 4) proposes that an assessment is a welknown and sometimes misunderstood term in educational practice.

2.4. Reading in a Second Language

Reading processes in a second language are similar to those acquired in the first language in that they call for knowledge of sound/symbol relationships, syntax, grammar, and semantics to predict and confirm meaning (Peregoy and Boyle 1993). As they do in their first language, second language readers use their background knowledge regarding the topic, text structure, their knowledge of the world, and their knowledge of print to interact with the printed page and to make predictions about it.

Two important differences between first and second language reading can be found in the language proficiency and experiences of the students. Students reading in a second language have varied levels of language proficiency in that language. The second language learner may be in the process of acquiring oral language while also developing literacy skills in English. Limited proficiency in a second language may cause a reader literate in the native language to "short circuit" and revert to poor reader strategies (such as reading word by word) (Clarke 1988). Also, students may not have the native language literacy skills to transfer concepts or strategies about reading to the second language. Those who do have native language

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

literacy skills may not know how to transfer their skills to the second language without specific strategy instruction. No empirical evidence exists to show that readers do in fact transfer reading strategies automatically from their first to a second language (Grabe 1988; McLeod and McLaughlin 1986).

Another difference between first and second language reading is that second language readers may have more varied levels of background knowledge and educational experiences (Peregoy and Boyle 1993). Students with a limited range of personal or educational experiences on a reading topic will have little to draw on in constructing meaning from text. In fact, the biggest single challenge to teachers of ELL readers may be the range of educational experiences presented by their students (Chamot and O'Malley 1994b).

2.5. What works in Reading Instruction

In addition to having new knowledge about the reading process, we also know what works in reading instruction. In particular, reading programs having the following four components can lead to student success: (1) extensive amounts of time in class for reading, (2) direct strategy instruction in reading comprehension, (3) opportunities for collaboration, and (4) opportunities for discussions on responses to reading (Fielding and Pearson 1994). We briefly discuss each of these components below and follow with an update on the phonics versus whole language debate.

Spending time reading in class is important because students benefit from the time to apply reading skills and strategies and also because time spent reading results in acquisition of new knowledge (Fielding and Pearson 1994). In turn, knowledge aids comprehension, vocabulary acquisition, and concept formation. Research has shown a consistent positive and

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

mutually supportive relationship between prior knowledge and reading comprehension. However, providing time for sustained silent reading is not enough. To improve reading comprehension, teachers must: (1) provide a choice of reading selections, (2) ensure that students are reading texts of optimal difficulty which challenge but

do not discourage them, (3) encourage rereading of texts, and (4) allow students to discuss what they read with others to encourage social negotiation of meaning.

One of the more important findings to emerge from research on reading instruction over the last fifteen years is that reading comprehension can be increased by teaching comprehension strategies directly (Fielding and Pearson 1994). Many reading strategies can be taught directly, including: using background knowledge to make inferences; finding the main idea; identifying sources of information needed to answer a question; and using story or text structure to aid comprehension. The most promising result of the comprehension strategy research is that instruction is especially effective with "poor comprehenders."

In addition to class time for reading and direct strategy instruction, peer and collaborative learning also contribute to reading acquisition !(Fielding and Pearson 1994). By working collaboratively, students gain access to each other's thinking processes and teach one another effective reading strategies. In particular, cooperative learning and reciprocal teaching, when implemented correctly, appear to promote reading comprehension. (See the discussion below on reciprocal teaching.) These approaches acknowledge the social nature of learning and the role of the reader as a negotiator of meaning.

Traditionally, teachers have led discussions of reading texts by posing a question for student response and then evaluating that response. However, current trends in reading instruction indicate a move away from primarily teacher-directed discussions to student-driven discussions, allowing for acceptance of personal interpretations and reactions to

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

literature (Fielding and Pearson 1994). These discussions are most effective when they incorporate reading strategy instruction. Changing teacher/student interaction patterns is challenging, however, since many teachers feel the need to maintain control while also "covering" the curriculum.

Similar to reading programs for native Speakers of English, reading instruction for English language learners should include at least five important components: a large quantity of reading, time in class for reading; appropriate materials that encourage students to read; direct-teaching of reading strategies; and a teacher skilled in matching materials and reading strategies to the students' level of interest and language proficiency (Devine 1988; Eskey and Grabe 1988). Such programs result in improved reading ability only when approaches to reading are holistic or integrative rather than skills-based, and when teacher feedback is a core element. In addition, reading instruction for English language learners should tap students' prior knowledge and experiences, focus on comprehension of meaning while teaching skills in context, teach text organization/and allow for collaborative discussions of reading.

2.6. Implications for Assessment

A number of implications for assessment can be drawn from the foregoing description of the nature of reading in first and second languages and effective instructional practices for increasing reading comprehension. These include the importance of determining students' prior knowledge, making students accountable for how they use reading time in class, assessing students' progress in acquiring both decoding skills and reading comprehension strategies, observing how students collaborate in groups as well as how they work individually, and reviewing students' personal responses to reading.

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

Garcia (1994) and Routman (1994) suggest that, in tying instruction to assessment, the key questions become: What do I as a teacher need to know about each student's literacy and language development in order to plan instruction? and What instructional activities and tasks can I use to find ; this out and document it? Information resulting from literacy assessment should help teachers identify students' needs and plan for the most suitable instructional activities. Activities discussed in this chapter that correspond to specific reading assessment purposes are described in Figure 5.1.

In order for reading assessment to become useful in student evaluation, teachers should consider the following (Routman 1994):

1. Be thoroughly familiar with developmental learning processes and curriculum.
2. Articulate a philosophy of assessment and evaluation.
3. Know about and have experience collecting, recording, interpreting, and analyzing multiple sources of data.
4. Be flexible and willing to try out multiple assessment procedures.

Figure 2.6.1 Reading Assessment: Matching Purpose to Task

What Do I Want to Know?	How Will I Find Out?
Reading comprehension	• Retellings • Literature response journals • Anecdotal records • Literature discussion groups • Texts with comprehension questions
Reading strategies	• Reading strategies checklists • Reciprocal teaching •Think-alouds • Anecdotal records • Miscue analysis • Running records

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

Reading skills	• Cloze passages • Miscue analysis • Running records
Reading attitudes	• Reading logs • Interviews • Literature discussion groups • Anecdotal records
Self-assessment	• Interviews • Rubrics/rating scales • Portfolio selections

Adapted from Routman (1994).

Be committed to understanding and implementing an approach to evaluation that informs students and directs instruction.

2.7. Authentic Assessment of Reading

Assessment requires planning and organization. The key lies in identifying the purpose of reading assessment and matching instructional activities to that purpose. After identification of assessment purpose, it is important to plan time for assessment, involve students in self- and peer assessment, develop rubrics and/or scoring procedures, set standards, select assessment activities, and record teacher observations. In this section we discuss each of these steps. We follow this with suggestions for bringing all of the information together in reading/writing portfolios and using reading assessment in instruction.

2.8. Language Competences

2.8.1 Canale and Swain Model of Communicative Competence

Canale and Swain (1980:23-30) produced the first and most influential model of what they call `communicative competence`, which is reproduced as leading to `more useful and

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

effective second language teaching, and allowing more valid and reliable measurement of second language communication skills. The writer tries hard to distinguish between `communicative competence and communicative performance.

Canale and Swain inFulcher and Davidson (2007:38-41) say that Canale and Swain attempt to do this firstly by reviewing how a variety of authors had so far defined communicative competence, and argue that for them it refers `to the interaction between grammatical competence, or knowledge of the rule of grammar , and sociolinguistics competence, or knowledge of the rules of language use`. They then firmly distinguish between communicative competence and communicative performance, the latter term referring only to the actual use of language in real communicative situations.

Canale and Swain present a model of knowledge, into which sociolinguistic competence is added. The model includes two components:

1. Communicative competence which is made up of:

1.1.grammatical competence: the knowledge of grammar , lexis, morphology, syntax, semantic and phonology,

1.2.sociolinguistic knowledge: the knowledge of sociocultural rules of language use and rule of discourse, and

1.3. strategies competence: the knowledge of how to overcome problems when faced with difficulties in communication.

2, Actual communication

2.1. the demonstration of knowledge in actual language performance.

Canale and Swain outline Hymes`s notion of a speech event in terms of participants, settings, form, topic, purpose, key, channel, code, norms of interaction, norms of interaction and genre. The speech event is said to be basis for understanding the rules of language use.

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

The writer thinks that this seminal model of communication is relevant to language testing for several reasons below. The first he answer the distinction between communicative competence and actual performance. The second is talking on communicative competence. And the last is about the model.

Firstly the distinction between communicative competence and actual performance means that tests should contain task that require communicative competence as well as tasks or items types that measure knowledge. These tasks type would allow test takers to demonstrate their knowledge in action. This is a theoretical rationale for the view thatpencil and paper test of knowledge alone can`tdirectly indicate whether a language learner can actually speak or write in a communication situation.

Secondly, as communicative competence was viewed as knowledge, discrete point tests were as useful as for some purposes. Discrete point tests - using terms that tested just one isolated item of grammar, for example – had been heavily criticized in the communicative revolution.

Thirdly, the model, especially if it were more `fine grained`, could be used to develop criteria for the evaluation of language performance, at different levels of proficiency. It isclear that the implications of a model of language competence and use have much to say about how we evaluate language performance , award to score to that performance and therefore interpret the score in terms of what we hypothesize the test taker is able to do im non-test situations.

2.8.2. Bachman Model of Communicative Language Ability (CLA)

Bachman`s model of CIA is an expansion 0f what went before, and does two things which make it different from earlier models. Firstly, it clearly distinguishes between what constitutes a `skill`, which was left unclear in the model of Canale; it explicitly `attempt to

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

characterize the process by which the various components interact with each other and with the context in which language use occurs` (Bachman, 1990:81). The three components of CIA for Bachman are language competence (knowledge); strategic competence (the `capacity for implementing the components of language competence in contextualized communicative use.

The two elements of discourse competence, cohesion and coherence, are spilt up. Cohesion occurs explicitly under textual competence, while coherence as a title appears and is subsumed under illocutionary competence. This is because the left-hand branch of the tree concerns the formal aspects of language usage, comprising grammatical competence and textual competence. The latter concerns knowledge of how text (spoken or written) are structured so that they are recognized as convention by hearers or readers.

The right-hand side of the tree is now described by the superordinate term pragmatic competence, which is defined as the acceptability of utterances within specific context of language use, and rules determining the successful use of a language within specified contexts.

It is strategic competence that now drives the model of the ability for language use. Bachman argues that strategic competence is best seen in terms of a psycholinguistic model of speech production, made up of three components:

Assessment component:

1. Identify information needed for realizing a communicative goal in a particular context.
2. Decide which language competences we have to achieve the goal.
3. Decide which abilities and knowledge we share with our interlocutor.
4. Evaluate the extent to which communication is successful.

Planning component:

5. Retrieve information from language competence.

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

6. Select modality or channel.

7. Assemble an utterance.

Execution component:

8. Use psychophysical mechanisms to realize the utterance.

Srategic competence is said to consist of avoidance strategies, such as avoiding a topic of conversation, and achievement strategies, such as circumlocution or the use of delexicalized nouns (such as 'thing'). Also included are stalling strategies, and self-monitoring strategies such as repair or rephrasing. Finally, but crucially, interactional strategies are listed, such as asking for help, seeking clarification or checking that a listener has comprehended what has been said.

Although the model presented is not unduly different from Canale (1980), and steps back from the non-linguistic elements of Bachman and Palmer (1996), it is nevertheless more specific about what each competence contains, and argues that the interaction of competences is the realm of strategic competence. It therefore contains a knowledge component and an ability for use component, following Hymes. This model appears to have brought us full circle. The authors are also explicit in stating that the model is not directly relevant as a whole to all teaching contexts. Celce-Murcia et al. (1995:30) state that:

> As McGroarty points out, 'communicative competence' can have different meanings depending on the learners and learning objectives inherent in a given context. Some components (or sub-components) may be more heavily weighted in some teaching-learning situations than in others. Therefore, during the course of a thorough needs analysis, a model such as ours may be adapted and/or reinterpreted according to the communicative needs of the specific learner group to which it is being applied.

The researcher agrees with this perspective. Ours is a book on language testing, and so the particular relevance of Celce-Murcia *et al.'s* work is to the design and validation of language tests, which would immediately limit its interpretation to other contexts of application.

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

2.8.3. Interactional Competence

Writers with a particular interest in the social context of speech and how communication is understood and constructed in a specific context have concentrated on developing the concept of interactional competence. With reference to the Celce-Murcia *et al.* model, Markee (2000: 64) argues that:

> The notion of interactional competence minimally subsumes the following parts of the model: the conversational structure component of discourse competence, the non-verbal communicative factors component of sociocultural competence, and all of the components of strategic competence (avoidance and reduction strategies, achievement and compensatory strategies, stalling and time-gaining strategies, self-monitoring strategies and interactional strategies).

The conversational structure component, as we have seen, would include sequential organization, turn-taking organization and the ability to repair speech. This approach draws together aspects of models that we have already considered into a new competence that focuses on how individuals interact as speakers and listeners to construct meaning in what has been called 'talk-in-interaction'.

The origin of interactional competence can be traced to Kramsch (1986), who argued that talk is co-constructed by the participants in communication, so responsibility for talk cannot be assigned to a single individual. It is this that makes testing interactional competence challenging for language testing, for as He and Young (1998: 7) argue, interactional competence is not a trait that resides in an individual, nor a competence that 'is independent of the interactive practice in which it is (or is not) constituted'.

The chief insight is that in communication, most clearly in speaking, meaning is created by individuals in joint constructions (McNamara, 1997). This is part of the theoretical rationale for the use of pair or group modes in the testing of speaking (Fulcher, 2003a: 186-190), as these modes have the potential to enrich our construct definition of the test.

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

Opening up performance in this way has interesting consequences for how we understand the design of tasks and how we treat the assessment of test takers in situations where they interact with an interlocutor (either a partner or a tester or rater). In terms of tasks we need to ask what kinds of activities are likely to generate the type of evidence we need to make inferences to the new constructs. In interaction, we need to investigate what constitutes construct-irrelevant variance in the score, or meaning that cannot be attributed to the individual receiving the score, and what part of the score represents an individual's interactional competence.

We therefore need to ask what aspects of performance might constitute realizations of interactional competence that can be attributed not directly to an individual but only to the context-bound joint construction that occurs in interactions - including an oral test. Such aspects of performance would be those that arise directly out of the *adaptivity* of one speaker to another.

This definition of adaptivity is not to be confused with an oral 'adaptive test', which is a test where the rater adjusts and refines scores for a test-taker, live and in real time, by selecting tasks that optimize themselves to the test-taker's actual ability range. Here, we are speaking of the natural adaptivity that happens in all oral discourse, as human beings engage in complex conversational mechanisms to make themselves understood to one another.

The simplest example of the principle of adaptivity in second-language communication is that of accommodation (Berwick and Ross, 1996), in which a more proficient speaker adapts their speech to the perceived proficiency level of the interlocutor, thus making communication easier for the other. One such example is lexical simplification, perhaps associated with slower delivery. The speaker makes an assessment of the abilities of the interlocutor, brings competences to bear in the adjustment of contributions to speech in real-

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

time processing, and uses contributions that enable the interlocutor to make further contributions by drawing on their own current competences more effectively.

2.9. How to Construct Formative Assessment

The writer reviews some assessments requirements in constructing well-organized assessments, they are Brown, Hughes, ESOL, and Harris. The general principles of assessment and guidance carry out equally to the context of mainstream delivery. However, the setting of language learning vocational context produces some supplementary considerations.

Philida (2011:191) guides some thoughts on what assessment duties should appear about the learner:prior occupational skills, educational accomplishments, and work experience, theycan have a significant impact on capacity to gain; language skills and needs;study skills, which will rely on the course, but most generally required are the ability to incorporate information, take notes, read materials and write assignment.

Brown (2004:56-58) states that there are four guidelines in designing multiple-choices test items. The first is designing each item to measure a specific objective. The second is stating both stem and options as simply and directly as possible. The third is making certain that the intended answer is clearly the only correct one. The fourth is using item indices to accept, discard, or revised items. Brown (2004:59-81) also adds that there are six steps in developing a standard test. The first is determining the purpose and objective of the test. The second is designing test specifications. The third is designing, selecting, and arranging testtasks/items. The fourth is making appropriate evaluations of different kinds of items,the fifth is specify scoring procedures and reporting formats. The sixth is performing ongoing construct validation studies. Brown (2004:206) elaborates the reading comprehension features which cover: main idea (topic); expressions/idiom/phrases in context; inference

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

(implied detail); grammatical features; detail (scanning for a specially stated detail); excluding facts not written (unstated details); supporting idea(s); vocabulary in context.

Hughes (2003:76-78) cautions against a four number of weaknesses of multiple-choice items:the technique tests merely recognition knowledge;guessing may have a considerable consequence on test scores; 1) the technique severely limits what can be tested; 2) it is very difficult to write prosperous items; 3) washback may be injurious;4) swindling maybe made easy.

Hughes (2005:58) recommends ten stages of test development. The procedures are: making a full and clear statement of the testing 'problem'; writing specific specifications for the test;writingand moderating items; trying the items informally on native speakers and rejecting or modifying problematic ones as necessary;trying the test on group of non –native speakers similar to those for whom the test is intended;analyzing the result of the trial and make any necessary changes;calibratingscales;the validating;writinghandbook for test takers; and training any necessary staff (interviewers, raters, etc).

There are seven basic rules in reproducing the test according to Harris (1969:108-110):1) it is necessary that test materials be reproduced as obviously as possible;2)test materials should be spaced so as to equip maximum readability;3)no multiple-choice item should be started on one page and continued on the next;4) when blanks are left for the accomplishment of short answer items, a guideline should be supplied on which the examinee may write his answer;5) it is advisable to denote at the base of each page if the examinee is to proceed on the next page or halt his work;6)if each part of the test is segregated timed, the directions for each part should fill a right-hand page of the book;7) the use of a separate cover sheet will prevent examinees from looking at a test material before the actual material administration begins.

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

The writer thinksO`Mallaey and Valdez (1996:11-14) play important roles in the process of teaching-studying.Feuer and Fulton in Mallaey (1996:11) state that there are numerous kinds of authentic assessment used in classrooms today. The range of possibilities is sufficiently broad that teachers can choose from a number of options to meet specific aims or adapt approaches to meet instructional and pupil needs. Teachers already use that strategy.

Philida (2008:128) writes that it is difficult to write obvious task instructions which are in the same level of language difficulty as the task itself. Piloting will give real information on how effective the task instruction are and the learner participating in the pilot give useful feedback.

Fulcher and Davidson (2007:28) state that language tests are designed by teachers with a particular skill and training in test design, or by people who specialize in test design. The researcher supports that the teacher constructs an assessment or a test should have a training or a course which is suitable to construct atest.

Nunan (1992:185) states that assessment refers to the process and procedures whereby we determine what learners are able to do in the target language. They express their ideasin ESL reading level accomplishment way out criteria. Level 4: The individual is able to read simple descriptions and narratives on familiar subjects or from which fresh vocabulary can be decided by context; can make some minimal inferences and contrast information form such texts but not consistency.

There are eight reasons tests and measurements are applied in the eveluation according to Tuckman (1975:7- 8): 1) to grant objectivity to our attentions;2) to elicit behavior under relativity controlled conditions;3)to sample performances of which the person is capable;4)to gain performances and measure gains relevant to goals or standard;5)to apprehend or to catch the materials and catch them or the unseen; 6) to detect the distinctions

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

and components of behavior;7) to forecast future behavior;8) to make data available for continuous feedback and decision making.

Tuckman (1975:77) also states that short-answer items typically ask students to identify, diffrentiate, state, or name something. In the free choice format, the measurement basically include asking students a question that requires that they state, or name the specific information or knowledge called for (recall it) indicating acquisition of that knowledge.

Nitko (1983:322) has argumentation on possible responses to a multiple choice has a correct choice, incorrect choice, and omit the item. Correct choice means possible interpretations: the person surelyknows the answer, makes a lucky random guess, answer uses partial knowledge, answer uses testwiseness. Incorrect choice means the person makes unlucky random guess, learns an errorneous response, learns an incomplete, learns an incomplete response, he is locked by a clever item writer, truly knows the answer but inadvertently makes the wrong mark. Because of this he is sacred to respond, he did not have adequate time to response. Therefore, he is unable to differentiate partial, incomplete, and lack of knowledge.Nitko (1983:193-194) states five advantages of multiple-choice items are frequentlylisted:1) among the various types of response-choice items, the multiple-choice item can be utilized to test a larger variety of instructional goals; 2) multiple-choice test do not require the examinee to write out and develop their answers, minimizing the chance for less knowledge examinees to “bluff” or “dress up” their answers (Wood, 1977) ;3)multiple-choice tests focus on reading and thinking and thus do not need the writing process to occur under examination circumstances;4) there is less chance for the examinee to surmise the correct answer to a multiple-choice item to a true-false item or to poorly constructed matching exercise;5)and if the distracters on multiple-choice items are relied on common pupil errors or mis-conception, then the items may give “diagnostic insight” into difficulties an individual pupil may be encountering.

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

The old criticism of the multiple-choice item as being something that we do not do `in real world` (Underhill, 1992) is therefore one that we can no longer recognize as meaningful. The researcher tries to minimize criticism on composing the multiple-choice item in this research. He collected the existing formative assessment made by 15 English teachers teaching in state senior high schools in Semarang municipality then analyzed them.

O`Malley and Valdez (1996:17-19) propose eight steps in designing authentic assessment: 1) creating an assessment team of teachers, parents, and an administrators to begin discussion; 2) determining the purposes of authentic assessment; 3) specifying objectives; 4) conducting professional development on authentic assessment; 5) collecting of authentic assessment; 6) adapting existing assessment or develop new ones; 7) trying out the assessment; 8) reviewing the assessment.

McNamara (1997:48) argues that all models of language ability have three main dimensions, constituted by statements about: 1) what it means (a model of knowledge); 2) underlying factors relating to the ability to use language (a model of performance; 3) how we understand specific instances of language use (actual language use). Ruch (1924:95-96) explains that detailed rules of procedures in the construction of an objective examination which would possess general utility can hardly be formulated. He adds that the type of questions must be decided on the basis of such facts as the school subjects concern.

2.10. Test Specification and Design

The specifications – usually called `spec` - are generative explanatory documents for the creation of test tasks. Specs tell us the nut and bolts of how to phrase the test items, how to structure the test lay out, how to locate the passages, and how to make a host of difficult choices as we prepare test materials. More importantly, they tell us the rationale behind the

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

various choices that we make. The idea of spec is rather old – Ruch (1924) may be the earliest proponent of something like the test specs.

Specs are often called blueprints, and this is an apt apology. Blueprints are used to build structure can be erected. For example, we may find ourselves looking at the row of new home, and, while each has unique design elements, we realize that all of the homes share many common features. Without going inside, we guess that the rooms are probably in the same spot, and so forth. We surmise that all homes on this row were built from common blueprint,

Alternatively, perhaps we are in the housing development and we note that there are about five or six models of homes, We see that another seem like one we saw on an adjacent street. In the second house in development, we surmise that the homes were built from a small set of some five or six blueprints. This second analogy is very like a test: there are a number of items or tasks drawn from a smaller number of specs, because we wish to re-sample various testing in order to improve our reliability and validity.

The classic utility of specs lies in the test equivalence. Suppose we had a particular test task and we wanted an equivalent task – same difficulty level, same testing objective, but different content. We want this to vary our test without varying the results. Perhaps we are concerned about test security, and we simply a new version of the same test with the same assurance of reliability and validity. This was the original purpose of specs. Following Davidson and Lynch (2002,Chap.7) we believe that test specs have a boarder and more profound impact on test development. They serve as a focus of critical review by test developers and users.

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

2.10.1. Planning in Test Authoring

A first and logical question might be; how much can we actually plan in any test? Ruch (1924:95-96) answers that detailed rules of procedures in the construction of an objective examination which would possess general utility can hardly be formulated. The type of questions must babe decided on the basis of such facts as the school subject concerned , the purposes of the examination, the length and reliability of the proposed examination, preferences of teachers and pupils, the time available for examination, whether factual knowledge or thinking is to be tested.

Kehoe (1995) presents a series of guidelines for creating multiple-choice test items. These guidelines are specs like in their advice. Here are the first two, which concern the stem of multiple-choice (a stem is the top part of multiple-choice item, usually a statement or question)

1. Before writing the stem, identify the one point to be tested by that item. In general, the stem should not pose more than one problem, although the solution to that problem may require more than one step.
2. Construct the stem to be either an in complete statement or a direct question, av avoiding stereotyped phraseology as rote responses are usually based on verbal stereotypes.

There are two elements to his – or any – multiple-choice item, each functioning somewhat differently. We first see a statement and question, known as the `stem`. We then see four `choices`. Most likely, the test taker was told that each item has one correct choice (something that testers call the `key`) and three incorrect choices (known as distracters`). To answer the item correctly, the students must read carefully each word of the stem and each word of choice; furthermore, the student probably knows well that three of four choices are intended to be incorrect, and so this close reading becomes a process of elimination. Suppose that a particular testing or teaching situation routinely uses such close-reading items, and in that settinglike this is actually very familiar to the students. In such a case, perhaps the students do not (really) read the item closely or analyze its component parts. They see it,

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

recognize it as familiar type of task, and engage the relevant cognitive and language processing skills – from their training – to attempt the item.

2.10.2. Guiding Language Versus samples

There are many styles and layouts for test specs. All test specifications have two components: sample(s) of the items or tasks we intend to produce and 'guiding language' about the sample(s). Guiding language comprises all parts of the test spec other than the sample itself. For the above seat-belt sample, guiding language might include some of these key points:

[1] This is a four-option multiple-choice test question.

[2] The stem shall be a statement followed by a question about the statement.

[3] Each choice shall be plausible against real-world knowledge, and each choice shall be internally grammatical.

[4] The key shall be the only inference that is feasible from the statement in the stem.

[5] Each distracter shall be a slight variation from the feasible inference from the stem; that is to say, close reading of all tour choices is necessary in order to get the correct answer.

Taking a cue from Ruch, we can assume one more important bit of highly contextualized guiding language:

[6a] It is assumed that test takers are intimately familiar with this item type, so that they see instantly what kind of task they are being asked to perform; that is to say, the method of the item is transparent to the skill(s) it seeks to measure.

An alternative would be to assume that our test takers are not familiar with this item type:

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

[6b] Test takers may or may not be familiar with this item type. The level of familiarity is not of any importance. The focus of this item is close reading, because the item must function as an assessment of proficiency, preferably of high-level proficiency.

2.10.3. Congruence (Fit-To-Spec)

The seat-belt item was first presented, above, without any guiding language. We then encouraged a critical reflection about the item. This is rather typical in test settings. We have test tasks in front of us, and somehow we are not happy with what we see.

Our next step above was to induce some guiding language such that items equivalent to the seat-belt question could be written. For example, something like this should be equivalent – it appears to follow our evolving spec here:

The vast majority of parents (in a recent survey) favourstricter attendance regulations at their children's schools. Which of the following *could* be true?

(a) Most parents want stricter attendance rules.

(b) Many parents want stricter attendance rules.

(c) Only a few parents think current attendance rules are acceptable.

(d) Some parents think current attendance rules are acceptable.

2.10.4. How Do Test Questions Originate?

We have seen a test question (the seat-belt item), and from it we induced some guiding language. We stressed our evolving guiding language by trying out a new item - on school attendance. Certain problems began to emerge.

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

Firstly, it is clear that our item requires very close reading-perhaps we want that; perhaps we do not. Perhaps we presume our students are accustomed to it boom prior instruction; perhaps we can make no such presumption. When the attendance item is compared to the seat-belt item, there are some new discoveries emerging. It seems that the four choices have a kind of parallel structure. Both items have two choices ((a) and (b)) that follow similarly structured assertions, and then two subsequent choices ((c) and (d)) that follow a different assertion structure. Guiding language like this might be relevant, and note that this guiding language refers directly to the two sample items - a good idea in test specs:

The purpose of this item is to test close inferential reading of assertions about scientific surveys. Items can contain precise counts and percentages (e.g. 'eighty per cent' in the first sample item) or generalities of a survey nature (e.g. 'the vast majority of parents' in the second).

Furthermore, we see in the second sample item that inference is broadly defined. Test takers may be asked to perform a mathematical calculation ('eighty of one hundred' is eighty per cent) or a double inference ('only a few' is the opposite of 'the vast majority').

Spec writing is an organic process. Time, debate, consensus, pilot testing and iterative re-writes cause the spec to grow and to evolve and to better represent what its development team wishes to accomplish.

2.10.5. Reverse Engineering

Reverse Engineering (RE) is an idea of ancient origin; the name was coined by Davidson and lynch (2002), but they are the first to admit that all they created was the name. RE is an analytical process of test creation that begins with an actual test question and infers the guiding language that drives it, such that equivalent items can be generated. As we did the

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

process here, it is a very good idea to stress an evolving spec by trying to write a new item during our reverse engineering. This helps us understand better what we are after. There are five types of RE, and the types overlap:

1. Straight RE: this is when you infer guiding language about existing items without changing the existing items at all. The purpose is solely to produce equivalent test questions.
2. Historical RE: this is straight RE across several existing versions of a test. If the archives at your teaching institution contain tests that have changed and evolved, you can do RE on each version to try to understand how and why the tests changed.
3. Critical RE: perhaps the most common form of RE, this is precisely what is under way here in this chapter - as we analyze an item, we think critically: are we testing what we want? Do we wish to make changes in our test design?
4. Test deconstruction RE: - whether critical or straight, whether historical or not, provides insight beyond the test setting. We may discover larger realities - why, for instance, would our particular test setting so value close reading for students in the seat-belt and attendance items? What role does close inferential reading have to the school setting? Are these educators using it simply to produce difficult items and thus help spread out student ability - perhaps in a bell-shaped curve? The term 'test deconstruction' was coined by Elatia (2003) in her analysis of the history of a major national language test.
5. Parallel RE: In some cases, teachers are asked to produce tests according to external influences, what Davidson and Lynch (2002, chapter 5) call the 'mandate'. There may be a set of external standards outside the classroom - as, for example, the Common European Framework. Teachers may feel compelled to design tests that adhere to these external standards, and, at the same time, the teachers may not consult fully with

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

one another. If we obtain sample test questions from several teachers which (the teachers tell us) measure the same thing, and then perform straight RE on the samples, and then compare the resulting specs, we are using RE as a tool to determine parallelism (Nawal Ali, personal communication).

2.10.6. Spec-Driven Test Assembly, Operation and Maintenance

Spec-driven testing lends itself very well to higher-order organizational tools. Specs are, by their very nature, a form of database. We might develop a test of one hundred items, which in turn is being driven by some ten to twelve specs. Each spec yields several equivalent items in the test booklet. When the time comes to produce a new version of the test, the specs serve their classic generative role to help item writers write new questions. Over time, each spec may generate many tens of items, and from that generative capacity an item bank can be built. The items are the records of the bank, and in turn, each item record can be linked to its spec. Over time, also, shared characteristics from one spec to another can be identified and cross-linked; for example, the spec under production in this chapter uses inferential reading.

Over time, the assembly and operation of a lest represents an investment by its organization. Effort is expended, and there is a risk: that the value of the effort will outweigh the need to change the test. A kind of stasis becomes established - the stable, productive harmony of an in-place test (Davidson, 2004). Stasis yields archetypes, and that is not necessarily a bad thing. Well-established specifications which produce archetypes seem to produce trust in the entire testing system, and from such trust many difficult decisions can be made. In placement testing (for example), stasis helps to predict how many teachers to hire, how many classrooms will be needed, how many photocopies to budget and a host of other logistical details - because the educators who run the test are so familiar with its typical results.

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

There may come a time at which members of the organization find themselves frustrated. 'This test has been in place for years! Why do we have this test? Why can't we change it?' Grumbling and hallway complaint may reach a crescendo, and if funding is available (often it is not) then 'finally, we are changing this test!'.

2.11.Curricullum 2013

Law No. 20 Year 2003 on National Education System states that the curriculum is a set of plans and arrangements regarding the purpose, content, and teaching materials and methods used to guide the organization of learning activities to achieve specific educational goal. Based on this definition, there are two dimensions of the curriculum, the first is a plan and setting the objectives, content, and material, while the second is the means used for learning,activities.

2.11.1. CharacteristicsofCurriculum2013

Curriculum 2013 is designed with the following characteristics: develop a balance between the spiritual and social development of attitudes, curiosity, creativity, cooperation with intellectual and psychomotor abilities; schools are part of that provides a planned learning experience; develop the attitudes, knowledge and skills and apply them in various situation in school and community.

The curriculum aims to prepare the 2013 Indonesian people that have the ability to live as individuals and citizens who believe, productive, creative, innovative, and affective and able to contribute to society, nation, state, and world civilization.

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

2.11.2. Basic Curricullum Framework

2.11.2.1. Philosophical Basis

Philosophical foundation for curriculum development to determine the quality of students who will achieve the curriculum, the source and content of the curriculum, the learning process, the position of learners, learning outcomes assessment, student relationships with the community and the surrounding natural environment.The curriculum was developed in 2013 with the philosophical foundation that provides the basis for the development of all human potential learners become qualified Indonesian listed in the national education goals. Basically, none of the philosophy of education that can be used specifically for the development of curriculum that can produce quality human. Accordingly, the 2013 curriculum was developed using the philosophy as follows:

1 . Education rooted in the cultural life of the nation to build the nation's present and future. This view makes the 2013 curriculum was developed based on a variety of Indonesian culture, geared to build the present life, and to build the foundation for a better life in the nation's future. It implies that the education curriculum is designed to prepare the young generation of the nation 's life. Thus, the task of preparing the young generation becomes the main task of a curriculum. To prepare for the life of the present and future learners, the curriculum in 2013 to develop learning experiences that provide opportunities for learners to master the competencies necessary for life in the present and the future, and at the same time continuing to develop their capacity as their nation's culture and people who care about the problems of contemporarysociety and the nation.

2 . Learners are developing their nations into creative culture. In the view of this philosophy, the nation's achievements in various fields of life in the past is something that should be included in the curriculum for students to learn. The education process is a process that provides an opportunity for learners to develop her potential into ability to think rationally

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

and academic excellence by giving meaning to what is seen, heard, read, learned of the cultural heritage based meaning that the lens is determined by culture and in accordance with the level of maturity psychological as well as physical maturity learners. In addition to developing the ability to think rationally and brightest in the academic curriculum in 2013 to position the cultural superiority studied to create a sense of pride, applied and manifested in personal life, in social interactions in the surrounding communities, and the nation 's life today.

3 . Education aimed at developing the intellectual and academic excellence through education disciplines. This philosophy determines that the content of the curriculum is the discipline and learning is learning disciplines (essentialism). This philosophy requires that the curriculum has the same name with the name the lessonsdisciplines, always aiming to develop intellectual abilities and academic excellence.

4 . Education to build the life of the present and future better than the past with different intellectual abilities, communication skills, social attitudes, concerns, and participate in community life and build a better nation (experimentalism and social reconstructivism). With this philosophy, curriculum 2013 intends to develop the potential of learners into reflective thinking skills in the settlement of social problems in the community, and to build a democratic people's lives better. Thus, using philosophy ascurriculum 2013 over the life of the individual in developing learners in religion, art, creativity, communication, and the various dimensions of intelligence value that corresponds to a self- learner and neededcommunity , nation and mankind.

2.11.2.2.Theoretical Basis

Curriculum 2013 was developed on the theory of"standards-based education" (standards - based education), and the theory of competency-based curriculum (competency- based curriculum). Education basedstandard provides for national standards as a minimum quality

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

standard of citizens broken down into content, process standards, competency standards, teachers and standards, infrastructure standards, management standards, standard finance, education and assessment standards. Competency -based curriculum is designed to provide the widest possible learning experience for students in developing the ability to act, knowledgeable, skilled, and act.

Curriculum 2013 adheres to: (1) the teachers teach curriculum in the form of a process developed in the form of learning activities in the school , classroom, and community, and (2) direct the learning experience of students (learned - curriculum) in accordance with the background, characteristics, and the ability of early learners. Learning experience directly into individual learners learning outcomes for themselves, while learning outcomes across the curriculum learners into results.

2.11.3. Curricullum Structure

2.11 3.1. Core Competencies

Core competencies are designed in line with the increasing age of students in a particular class. Through its core competencies, vertical integration of various basic competence in different classes can be maintained .The formulation of the core competencies usingthefollowingnotation:

1. Core Competencies - 1 (KI - 1) for the core competencies of spiritual attitude;
2. Core Competency - 2 (KI - 2) for the core competencies of social attitudes;
3. Core Competency - 3 (KI - 3) for the core competencies of knowledge, and
4. Core Competency - 4 (KI - 4) for the core competency skills.

The description of the Core Competencies for High School level up / Madrasah Aliyah can be seen in the following table.

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

2.11.3.2. Core Competences and Basic Competences of Curriculum 2013

See Table 1

2.12. Indicators of Assessment on Recount

Derewianka (1995:11-15) states that there are two kinds of recount: factual recount and imaginative recount. The characteristics of factual recount are using third person pronoun; mention of personal feelings is probably not appropriate; the passive voice may be used; details are usually selected to help the reader reconstruct the activity or incident accurately; detail of time, place and manner may need to be precisely stated; descriptive details may also be required to provide precise information.While, the characteristics of imaginative recount: it is usually written in the first person, and it may be appropriate to include personal reactions.

The researcher studies curricullum 2013, especially on core competences and basic competences which related to recount. He makes sentences as indicators. The indicators of curricullum 2013 which relate to recount : expressings: grateful for the opportunity to learn English as a language of international communication embodied in the spirit of learning; demonstrate courteous behavior and care in performing interpersonal communication with teachers and friends; demonstrate courteous behavior and care in performing interpersonal communication with teachers and friends; demonstrate responsible behavior, caring, cooperation, and peace-loving;analyze social functions, text structure and linguistic elements of the text presentation of identity, according to the context of its use. analyze social functions, text structure and linguistic elements in the expression of praise winged (extended), as well as the response, according to the context of its use; analyze social functions, text structure and linguistic elements of the text recount about experiences, events, and events, simple, according to the context of its use; Analyze social functions, text structure

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

and linguistic elements of a simple descriptive text about the people, sights, and famous historical buildings, according to the context of its use.

2.13. Bloom Taxonomies

The researcher thinks that in composing formative assessment, one needs to decide that the item is easy, medium, ordifficult. Therefore the researcher has to understand the Bloom taxonomies in detail. Nitko (1983:99) states that the teacherswho develop formative assessment, a teacher will obtain the taxonomy of educational objectives. The taxonomy classifies cognitive performances into six major headings. First, in Indonesia commonly called C1, is knowledge. It includes recall. Second, (C2) is comprehension that represent the lowest level of understanding. Third, (C3) is Application that utilizes of abstraction in particular and concrete circumstance. Fourth, Analysis (C4) means obscuring the communication. Fifth, synthesis (C5) means putting together of elements and parts. Sixth, evaluation (C6) means judgment about the value.

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

CHAPTER III

RESEARCH METHOD

3.1. Research Design

According to Borg and Gall (2003:567) evaluation plays a key role in educational research and development (R & D). They think that educational R & D is an industry based development model in which findings of research are used to design fresh products and procedures, which then are systematically field-tested, evaluated, and refined until they meet specified criteria of effectiveness, quality, or similar standard. Borg and Gall also stress that educational R & D, accounts for less than one percent of each education dollar. Nonetheless, it has promised for improving education because it involves a close connection between systematic program evaluation and program development.

The writer uses R &D because this research develops the formative assessment on recount for the ten graders of senior high school 11 Semarang. He divides this R & D into two phases: the first, the need assessment for English teachers to know the formative assessments which are needed or not. The second is developing the formative assessment.

The researcher carried out his research on developing formative assessment on recount by doing the thirteen steps in the model of writing formative assessment on recount: (1) need assessment to identify the need of formative assessment; (2) Analysing curricullum 2013; (3) studying core competences; (4) analyzing basic competences related to recount; (5) determining materials related to recount, expression grateful, analyze social function, linguistic elements of recount; (6) writing indicators; (7) writing draft I formative assessment on recount; (8) revision by English teacher; (9) writing draft II formative assessment on recount; (10) revision by senior

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

English teacher; (11) writing draft III formative assessment on recount; (12) validated by Hartoyo, MA, PhD.; (13) formulating final product of formativer assessment on recount.

3.2. Population/Samples

In the preliminary research, the researcher conducted need assessment for English teachers who teach in state senior high schools in Semarang municipality. Forty English teachers who teach in senior high schools were taken as the sample.

Meanwhile for the try out, the population of the research is the students of the ten grade of senior high school 11 Semarang, and class X IPA 4 consisting thirty-five students are taken as the sample.

3.3. Instrumentation

The researcher distributes questionnaires to 40 English teachers as sample in the need of assessment. While he applied test in the R & D of developing formative assessment on recount for the ten graders of senior high school 11 Semarang; he tried out for three times: try out 1, try out II, and try out III.

3.4. Data Collection

First, the researcher collects the data by using questionnaire on the need assessment whether they need formative assessment or not. He collects from the English teachers who teach in state senior high schools. There are 40 English teachers from Senior High School (SHS) 1 up to SHS 16 in Semarang municipality. The procedures of collecting data: collecting the result of

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

questionnaires, Making table for recording the result of the English teachers, gving the mark or thick of the English teachers` questionnaires results.

Second, he collects the existing formative assessments from fifteen English teachers who teach in 16 state senior high schools in Semarang municipality. Then he reviewed them.

Third, he collected data from X IPA 4 or thirty five students tests the formative assessment in try out I, try out II, and try out III. The reseacher then analysed the result of the draft formative on recount I, the draft formative on recount II, and the draft formative on recount III. Draft I was revised by the English teacher in senior high school 11 Semarang. Draft II was revised by senior English teacher in the same school, while draft III is validated by an expert on assessment (Hartoyo, MA, PhD).

3.5. Data Analysis

Data from the questionnaires of the need of assessessment were analysed properly and carefully tabulated. The data of the result of try out I, try out II, and try out III then are analyzed to find means, and to differentiate in order to identify whether there is significant difference among the try out I, try out II, and try out III. The function of t is to know the improvement of every try out.

3.6. Revision of Formative Assessment

There are two revisions by English teachers: Dian Arini Rosita, SPd, and Dra. Theresia Melania Sudarwati, MSi. After those revisions, The drafts are corrected and then they are validated by an expert from on assessment Hartoyo, MA, Phd.

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

CHAPTER IV

FINDINGS AND DISCUSSION

In this chapter the conclusions of phenomenon discussed in previous analysis as the answers t the research questions and some possible suggestions for the teacher and the next researcher.

4.1. The I Result of the Need Assessment

The need assessment was carried out on December, 8 until 18, 2012.The population was 40 English teachers from senior high school 1 Semarang until Senior High School 16 Semarang municipality. The conclusion from part A is that 82.5 % of English teachers need formative assessment.

The conclusion of B or Recount text is 33 or 82.5 % choose "very important", 7 or 17.5% answer "important", 0 or zero % answer "not important".The conclusion of part C is that the samples say that four skills in English language are important, but the rank of the importance according to the samples is writing skill (92.5%), reading skill (85%), listening skill (72.5%), and speaking skill (65%).

4.2. The II Result of the Existing Formative Assessment on Recount

The researcher collected 15 existing formative assessment from English teachers which can be categorized: 1) `open answer`; 2) `multiple-choice`; 3) `free composition`; 4) summary. The number of English teachers who wrote formative assessment : `open answer` was 3; `multiple-choice` with distracters were 4; `multiple-choice with five distracters were 5; and `summary` is 3.

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

The researcher reviews those formative assessment made by the English teachers by using criteria: the instrument is grammar and form of formative assessment (number of items). The scores are based on these requirements:

1. Grammar:no mistake (0%) 100,very rare mistakes (1%-10%) 90, rare mistakes (11% - 15%) 80, more mistakes (16% - 20%) 70;

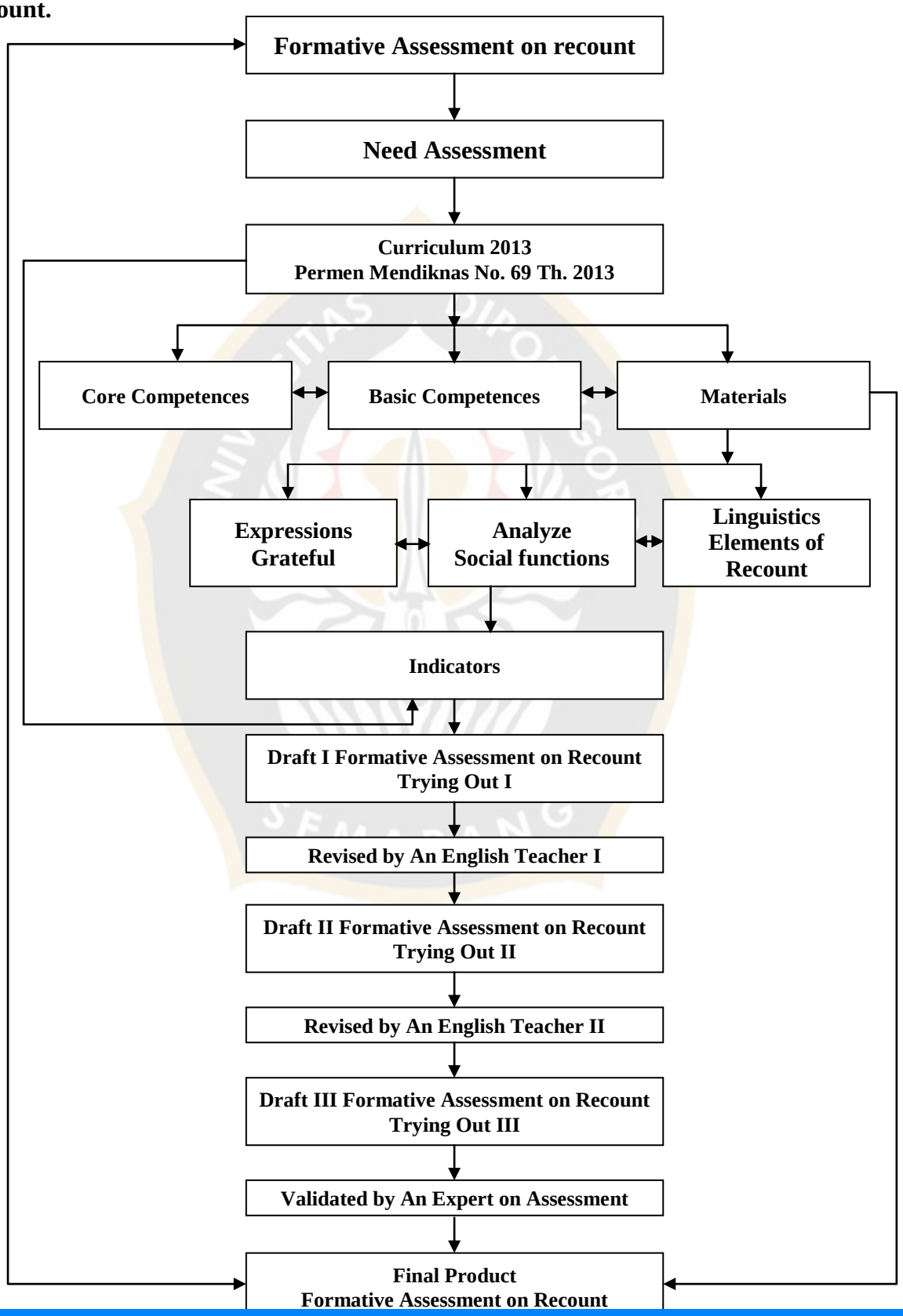
2. Form of Formative Assessment/Number of items:multiple – choice 5 distractors: 21 - 30 items 90, 16 – 20 items 80, 11- 15 items 70, 6 – 10 items 60; multiple – choice 4 distractors:31 – 40 items 90, 21 -30 items 80, 16 – 20 items 70, 11- 15 items 60; open answer: 25 - 30 items 90, 16 – 20 items 80, 11- 15 items 70, 6 – 10 items 60, 1 – 5 items 50; summary: 1- 5 items 90, 6 – 10 Items 80, 11- 15 Items 70, 16 – 20 Items 60, 20 – 30 Items 50

From 15 English teachers: 3 English teachers got A, 6 English teachers got B+, 6 English teachers got B. The researcher took mark 80 or B as the standard of the developing formative assessment.The result was: three English teachers got A, They were from state senior high school 4, state senior high school 6, and , state senior high school 7, while six English teachers got B+. They were from state senior high school 2, state senior high school 3, state senior high school 10, state senior high school 11, and state senior high school 12. Six other English teachers got B. They were from state senior high school 1, state senior high school 5, state senior high school 9, state senior high school 13, state senior high school 15, and state senior high school 16. See Table III in the appendices

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

4.3. The Result of Model of Developing Formative Assessment on Recount

This flowchart is the model used by the researcher to develop formative assessment on recount.



© Master Program in Linguistics, Diponegoro University

The flowchart above is used for developing formative assessment on recount. This flowchart is modified from the steps of the systems approach model of educational research and development (R&D) which is illustrated in Borg and Gall (2003: 571). Some steps are the same such as assing needs to identify goal (s), writing performance objectives, developing assessment instruments, and revising.The researcher priorities on analyzing curriculum 2013 to find out the indicators for developing the formative assessment on recount. He starts from analyzing core competentes, basic competences, and materials of recount: expressing of grateful, analyzing social functions, and linguistic element of recount. After the researcher is able to decide indicators, he starts to write items in multiple-choice with five options, The reason is to adjust the national examination in order the students recognize the characteristic of national examination.

The important steps are revising and validating. Draft I is revised by his friend, an English teacher who teach in senior high school 11 Semarang. Draft I is then tried out in X IPA 4 which is chosen as samples. Draft I and draft II are revised by different English teachers. Draft III is validated by an expert on assessment, Hartoyo, M.A, PhD,

There are thirteen steps in the model of writing formative assessment on recount:

1. Need of assessment. The base of developing formative assessment is the need of assessment as prelimanary research. The need assessment was carried out on December, 8 until 18, 2012.The population was 40 English teachers from Senior High School 1 Semarang until Senior High School 16 Semarang. The result of the need assessment is below.

2. Analyzing curricullum 2013. The researcher makes a library research to analyze curricullum 2013.Curriculum 2013PermenMendiknas No. 69 , 2013.The curriculum aims to

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

prepare the 2013 Indonesian people that have the ability to live as individuals and citizens who believe , productive , creative , innovative , and affective and able to contribute to society, nation , state , and world civilization. It has core competences and basic competences for each class. Theoretical Basis of curriculum 2013 was developed on the theory of " standards-based education " (standards - based education) , and the theory of competency-based curriculum (competency- based curriculum) . Education basedstandard provides for national standards as a minimum quality standard of citizens broken down into content, process standards , competency standards , teachers and standards , infrastructure standards , management standards , standard finance , education and assessment standards . Competency -based curriculum is designed to provide the widest possible learning experience for students in developing the ability to act , knowledgeable , skilled , and act.

3. Selecting core competentences which are related to recount.Core competencies are designed in line with the increasing age of students in a particular class . Through its core competencies , vertical integration of various basic competence in different classes can be maintained. .

The formulation of the core competencies using the following notation :

1) Core Competencies - 1 (KI - 1) for the core competencies of spiritual attitude ;2) Core Competency –

2 (KI - 2) for the core competencies of social attitudes ;3) Core Competency - 3 (KI - 3) for the core competencies of knowledge , and4) Core Competency - 4 (KI - 4) for the core competency skills .

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

4, Choosing basic competences which related to recount. Actually basic competences are the elaborating of core competences. One core competences can be elaborated to one until fifteen basic competences. Basic competences are operational in verbs. They can be assessed. They can also guide the teachers to achieve the basic competences.

5. Elaborating materials which related to recount. They consist of expression of grateful, analyze social function, and linguistics elements of recount.

6. Writing indicators. Indicators must be very specific. They can be measured and assessed. They always related to basic competence and core competence.

7. Writing items of formative assessment on recount draft I, revised by senior English teacher, carrying try out I. The reseacher analyzes the result of try out I by using SPSS 20. The sample is class X Science 4 of senior high school 11 Semarang.

8. Being revised by English teacher in high school 11 Semarang

9. Writing items of formative assessment on recount draft II, carrying try out II. The reseacher analyzes the result of try out II by using SPSS 20.

10. Being revised by senior English teacher in high school 11 Semarang

11. Writing items of formative assessment on recount draft III, revised by senior English teacher, carrying try out III. The reseacher analyzes the result of try out III by using SPSS 20. The sample is class X Science 4 of senior high school 11 Semarang.

12. Being validated by an expert on assessment, Hartoyo, MA, PhD.

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

13. Composing final product: developing formative assessment on recount for the ten graders of senior high school 11 Semarang.

Based on the above model, the researcher offers to develop a formative assessment on recount for the ten graders of senior high school 11 Semarang. See appendix.

The model above has been done for developing formative assessment on recount in the first grade of senior high school 11 Semarang.

4.4. Revision of Formative Assessment

4.4.1. First Revision by Dian Arini Rosita, SPd.

The first, number one, Dian Arini Rosita reminds that the requirements of constructing answers are in order from the longest to the shortest or the opposite. The revision is 14 characters, 13 characters, 9 characters,9 characters, and 7 characters. The number of characters are included the spaces. The second, she reminds the researcher that the item number two is the answer should be "Good bye or good night as a culture in western countries.

No. Item	**Assessment Item is Corrected**	**The result of Revision**
1	Exercise 1: Listening For each question from 1 to 5, five options are given. One of them is the correct answer. Make your right choice (A, B, C, D, or E) in the brackets provided. Man: Hi my name is Amir. I am 15 years old. How do you do? Woman: ….. **A.** How do you do? (**14 characters)** B. I am ok **(7 characters)** C. I am very well**(13 characters)**	Exercise 1: Listening For each question from 1 to 5, five options are given. One of them is the correct answer. Make your right choice (A, B, C, D, or E) in the brackets provided. Man: Hi my name is Amir. I am 15 years old. How do you do? Woman: …. What is the best response for the woman to reply? A. How do you do? (**14 characters)** B. I am very well (**13 characters)** C. I am Tina**(9 characters)**

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

	D. I am fine**(9 characters)** **E.** I am Tina**(9 characters_** **The requirements of constructing answers are in order from the longest to the shortest or the opposite**.	D. I am fine **(9 characters)** E. I am O.K. (**7 characters)**
2	Man : It is 9.00 pm. It is time for me to go home. Good bye all. Woman …. A. Good evening B. Good morning C. Good afternoon D. Good noon E. Good bye **The answer should be "Good bye or good night as a culture in western countries.**	Man : It is 9.00 pm. It is time for me to go home. Good bye all. Woman …. What is the best response to give? A. Good afternoon B. Good morning C. Good evening D. Good noon E. Good night

4.4.2. First Revision by Dra. TheresiaMelaniaSudharwati, MSi.

TheresiaMelaniaSudharwati states that problem item number three, there is not important "great dad" and "only" and she recommends to add "What is the best response for the man to reply".Shealso revises "I only …." Is omitted because to avoid with "if only" which means conditional.

Problem item number five must be " Subject + please + to + infinitive" to express willingness. To ask for a conclusion uses "What can we infer from the dialogue?" . The options should be A.The woman shows her attention, B. The woman refuses his offer, C. The man shows his attention, D. The man refuses her offer, E. The woman greets him.

Number item number eight must be all options are correct, but option B, is not correct "Does you see?". That option must be "Are you seeing?". The question is "Amin: …your grandma last year John? I missed her so much.

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

Problem item number nine is , option does not trap, grammatically must be correct. "Juvent: What did you send last week?Ivan: I … some papers to Semarang last week. The options are A. were sending/ was sending, B. are sending/am sending, C. have sent, D.sent/ had sent, E.send/sent.

Problem item number ten is check to be +ving form. "Toto: Where did you hide yourself thirty years ago? Yanti: I …. In Indonesian policewoman, how about you? The options are

A. Hiding / was hiding, B. hidden/have hidden, C.hide/ had hidden, D.hides/hide, E.hid

The problem item number twelve is pronoun. "I" is changed into "the writer" so the options become

A. The writer found a spider in the bathroom yesterday, B. The writer went to the kitchen and got a glass, C. After that the writer went to the window, D. The writer had to take care of himself, E. The writer needed to take a bath.

Problem number twenty six is the options must be a sequenced whether from the biggest characters or the smallest characters. The best construction must be "The first paragraph of a recount text belongs to …. A. events, B. orientation, C. conclusion, D. Complication, E. Reorientation"

Problem number twenty nine is "the question should be clearly stated". The correction isWhat can you find in the first paragraph of a recount text?" It is better than "There are three kinds of orientation. What are they?"

"What can you find in the first paragraph of a recount text?" It is better than "There are three kinds of orientation. What are they?"

" Problem number thirty is reference. The person who refers needing an adjective clause. Once upon a time there lived a beautiful princess, named princess who lived with an old woman". She was very kind.**She refers to ….**A. A beautiful princess,B. A beautiful woman, C.An old princess, D.An old woman, E. A woman.

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

No. Item	Assessment Item is Corrected	The result of Revision
3	Man: What was your result of your English test? (son dihilangkan) Woman; Great dad. I got nine and a half. I onlymissed one. (great dad hilang) Man: …. Narrator : What is the best response for the man to reply? A. Oh, I`m happy to hear it /**Great, I'm happy to hear it.** B. Oh, I`m sad to hear it C. I`m sorry to hear it D. You are so foolish E. You are joking	Man: What was your result of your English test? Woman: I got nine and a half. I missed one. Man: …. What is the best response for the man to reply? A. Great, I'm happy to hear it. B. Oh no, I`m sad to hear it C. I`m sorry to hear it D. You are so foolish E. You are joking
5	Listening, tidak bisa dengan model soal ini Woman: Tomorrow is my sister`s birthday. Would **you please to come** to my house to celebrate my Sister`s birthday?.....diganti you please come Man :I am so sorry, my mother has asked me to accompany her to go shopping. **The underlined expression belongs to** ….diganti Narrator : What can we infer from the dialogue? A. Showing attention /The woman shows her attention B. Refusing an offer/The woman refuses his offer C. Leave taking/The man shows his attention D. Introduction/The man refuses her offer E. Greeting/ The woman greets the man	Woman: Tomorrow is my sister`s birthday. Would **you please to come** to my house to celebrate my sister`s birthday, please come. Man :I am so sorry, my mother has asked me to accompany her to go shopping. What can we infer from the dialogue? A. The woman shows her attention B. The woman refuses his offer C. The man shows his attention D. The man refuses her offer E. The woman greets him
8	Amin: … you see your grandma last year John? I missed her so much. John: Yes, I saw my grandma last year. A.Do/ do you see B. Does/ are you seeing C. Have/ have you seen D. Did/ did you see E. Are/ are you going to see Clue tetaplah dalam bentuk grammar yang	Amin: …your grandma last year John? I missed her so much. John: Yes, I saw my grandma last year. A.are you going to see B. are you seeing C. have you seen D. did you see E. do you see

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

	benar…..	
9	Juvent: What did you send last week? Ivan: I … some papers to Semarang last week. A. were sending/ was sending B. are sending/am sending c. have sent C.sent/ had sent D.send/sent *Option jangan menjebak, grammar tetap* *harus benar*	Juvent: What did you send last week? Ivan: I … some papers to Semarang last week. A. was sending B. am sending C. have sent C had sent E. sent
10	Toto: Where did you hide yourself thirty years ago? Yanti: I …. In Indonesian policewoman, how about you? A.hidding/ was hidding B. hidden/have hidden C.hide/ had hidden D.hides/hide E.hid	Toto: Where did you hide yourself thirty years ago? Yanti: I …. In Indonesian policewoman, how about you? A. was hidding B. have hidden C. had hidden D.hide E.hid
12	What's the first event of the story? A. I found a spider in the bathroom yesterday**. / I diganti the writer** B. I went to the kitchen and got a glass. Then,/then hilang, **I diganti the writer** C. I'll have to take care of myself!/ The writer had to take care of himself D. After that I went to the window/I diganti the writer E. I needed to take a bath/ I diganti the writer	What's the first event of the story? B. The writer found a spider in the bathroom yesterday. B. The writer went to the kitchen and got a glass. C. . After that the writerwent to the window D. The writer had to take care of himself E. The writer needed to take a bath
26	What are the generic structures of Recount text?/ The first paragraph of a recount text belongs to A. orientation B. events C. conclusion D. Reorientation E. Complication *Answer:* ***A. Orientation*** } ***dihilangkan*** *A. Event* *B. Reorientation*	The first paragraph of a recount text belongs to A. events B. orientation C. conclusion D.Complication E. Reorientation

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

27	What is the tense mostly used in a recount text? of Recount? a. Past tense b. Present tense c. Present continuoes d. Past perfect e. Past continuous **Answer; Past Tense**	What is the tense mostly used in a recount text? a. Past tense b. Past perfect c. Present tense d. Past continuous e. Present continuous
29	There are three kinds orientation. What are they?/ What can you find in the first paragraph of a recount text? a. a sequence of events b. who, where,when c. an explanation d. pro and cons e. Description **Key: Characteristic** **Time** **Place**	What can you find in the first paragraph of a recount text? a. a sequence of events b. who, where,when c. an explanation d. pro and cons e. Description
30	Once upon a time there lived a beutiful princess, named princess who lived with an old woman. She was very kind. **She refers to ….** a. A beautiful princess b. A beautiful woman c. An old princess d. An old woman e. A woman **Key A beautiful princess.**	Once upon a time there lived a beautiful princess, named princess who lived with an old woman. She was very kind. **She refers to ….** a. A beautiful princess b. A beautiful woman c. An old princess d. An old woman A woman

4.4.3. The Assessment Items were validated by Hartoyo, PhD, MA.

Validation Instrument

(Expert Judgement)

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

Subject : Developing Formative Assessment on Recount for the First Grade of State Senior High School 11 Semarang on the Academic Year 2013/2014

Topic : Developing Formative Assessment on Recount

Target : The First Grade Students

Signature :

Day/Date : March 21st 2014

A. Directions

- This evaluation instrument is filled out by an expert who specializes in teaching academic writing or language assessment.
- The evaluation embraces several aspects: goals or objectives, the content or aspects of the model, self-assessment instruments, and the implementation procedure (manual guide).
- The answers are given by putting a cross (X) stated in numbers 1, 2, 3, 4 and 5 On the provided columns.
- Score criteria:

  1 : very good: not necessarily changed

  2 : good: a little changes may be made

  3 : fair: partially removed

  4 : poor: greater revision should be made

  5 : not at all: removed.

- The model is attached. Some necessary changes are required for its improvement.

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

B. Instrument

The indicators are clearly described

No		1	2	3	4	5
Aspect :The core instructional/The basic instructional						
1.	The core instructional are adequately stated	**V**	**V**			
2.	The core instructional are clearly taken from curriculum 2013				**V**	
3.	The core instructional are accurately developed in the basic competences					**V**
4.	The core instructional are are accurately and the same principals with the basic competences					**V**
Aspect : The indicators						
5.	The indicator of formative assessment can Be assessed			**V**		
6.	The indicator of formative assessment is Stated depending on the core/basic Instructional			**V**		
7.	The indicator of formative assessment refers to clear item			**V**		
8.	The indicator of formative assessment is Clearly written in accordance with core or basic instructional			**V**		

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

Aspect: Construction						
9.	The item of formative assessment is developed from indicator			**V**		
10.	The item of formative assessment and alternative answers must be short					**V**
11.	The item of formative assessment must be in good structure					**V**
12.	The item of formative assessment has only correct answer					**V**
13.	The item of formative assessment should be started on one page and continued on the next; when blanks are left for the accomplishment of short item					**V**

C. Comments and Suggestions

Great job. Please be careful with formative assessment because it will make our young generations better.

Hartoyo,MA,PhD

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

The problem of item number one is spelling of O.K. It is an important part of writing test item because there will not be understood if the writer does not follow the rules. O.K. means all right for using to attract or to introduce a comment, used to check that somebody agrees with you or understands you.Problem of item number two is Hartoyo, M.A., PhD suggests that the answer is "Good night" he focuses on "It us 9.00 pm. It is time for me to go home. Good bye"The problem item number the term of '95' is "I got ninety-five.The problem item number four is using " Would you like to + infinitive", while the problem item number five is cross cultural understanding that woman and man usually do not attend the birthday party. The man is changed into Dani and woman is changed into Dani.

The problem item number seven is an adverb of time, "Did you visit Kintamani lake last year?. Last year is changed into "during last vacation". The answer will be "No, I did not; as it was raining hard when I was there". And, the problem item number eight is we (both John and I recognize grandma well. While the problem item number nine is it needs an adverb of place 'Semarang".

The problem item number eleven is Hartoyo adds re-orientation on the text,finally, I had a bath. I did it myself! I was proud of myself (What`s for), then adds preposition of on option C.The problem item number thirteen is options C, D, and E are not capital letter because in the middle of a sentence. They must be C. in front of, D. at the end, and E. firstly.

The problem item number sixteen is parallelism, for example the smartest and the funny (must be the funniest). While the problem item number eighteen is competition must be competitions. And the problem item number twenty – one is added by "that is told" after there is a famous story. Options B & C are similar in meaning: B. Washington`s dishonest C. Washington`s dishonesty.

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

The problem item number of number twenty –three is to sharpen the question: Which one of the following words has the same meaning as the underlined word? Next is the problem item number twenty – four is parallelism. Option C is not a noun, while other options are nouns. The problem item number twenty – five is capital letters. In the position in the middle of a sentence must not be a capital letter.

The problem item number twenty –six is capital letter. After the statement "The underlined word refers to …. The options should not be capital letters. While the problem number twenty – nine is option D must be "Pros and cons". The last item number thirty is adding an adjective clause "Once upon a time there lived a beautifulprincess, named princess who lived with a poor family She was very kind.

No. Item	Assessment Item is Corrected	The result of Revision
1	Exercise 1: Listening For each question from 1 to 5, five options are given. One of them is the correct answer. Make your right choice (A, B, C, D, or E) in the brackets provided. Man: Hi my name is Amir. I am 15 years old. How do you do? Woman: ….. A. How do you do? B. I am very well C. I am Tina D. I am fine E. I am ok **Ok should be O.K.**	Exercise 1: Listening For each question from 1 to 5, five options are given. One of them is the correct answer. Make your right choice (A, B, C, D, or E) in the brackets provided. Man: Hi my name is Amir. I am 15 years old. How do you do? Woman: …. What is the best response for the woman to reply? A. How do you do? B. I am very well C. I am Tina D. I am fine E. I am O.K.
2	Man : It is 9.00 pm. It is time for me to go home. Good bye all. Woman …. A. Good evening B. Good morning C. Good afternoon	Man : It is 9.00 pm. It is time for me to go home. Good bye all. Woman …. What is the best response to give? A. Good afternoon B. Good evening

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

	D. Good noon E. Good bye **Consider E. "Good night"**	C. Good morning D. Good noon E. Good night
3	Man: What was your result of your English test? Woman: I got nine and a half.(**I got the ninety-five)** Man: …. What is the best response for the man to reply? A. Oh, I`m happy to hear it B. Oh, I`m sad to hear it C. I`m sorry to hear it D. You are so foolish E. You are joking	Man: What was your result of your English test? Woman: I got the ninety-five. Man: …. What is the best response for the man to reply? A. Great, I'm happy to hear it. B. Oh no, I`m sad to hear it C. I`m sorry to hear it D. You are so foolish E. You are joking
4	Man : What's the film about? **Woman : Well, it's about** the fighting of young heroes in the beginning of independence. Would like to come to the movie with me tonight?**(Should be Would you like…..)** Man: …. Narrator : What is the best response for the man to give? A. It is my business B. It is not funny C. Of course D. Not at all E. Not now	**Man : What's the film about?** **Woman : Well, it's about** the fighting of young heroes in the beginning of independence. Would you like to come to the movie with me tonight? Man: …. What is the best response for the man to give? A. It is my business B. It is not funny C. Of course D. Not at all E. Not now
5	Woman: Tomorrow is my sister`s birthday. Wouldplease to come my house to celebrate mySister`s birthday?(Change woman to Diana and man toDani because both man and woman are not common to celebrate birthday, but teenagers in England. Would you please + verb without to Man :I am so sorry, my mother has asked me to accompany her to go shopping. What can we infer from the dialogue? A. The woman shows her attention B. The woman refuses his offer C. The man shows his attention D. The man refuses her offer E. The woman greets him	Diana: Tomorrow is my sister`s birthday. Would **you please come** to my house to celebrate my sister`s birthday, please come. Dani :I am so sorry, my mother has asked me to accompany her to go shopping. What can we infer from the dialogue? The woman shows her attention B. The woman refuses his offer C. The man shows his attention D. The man refuses her offer E. The woman greets him
7	Teacher: Did you visit Kintamani lake last year?**(last year should be changed by "during your last vacation)** **No, I did not; as it was raining hard when I**	Teacher: Did you visit Kintamanilake during your last vacation? Bagas: No, I did not; as it was raining hard when I was there.

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

	was there. Bagas: Sorry, It was raining.**(No, I did not; as it was raining hard when I was there.)** **A. hadn`t visited** **B. didn't visit** **C. had visited** **D. was visiting** **E. visited**	A. hadn't visited B. didn't visit C. had visited D. was visiting E. visited
8	Amin: … you see your grandma last year John? I missed her so much.**(we both John and I recognize grandma?)** John: Yes, I saw my grandma last year. A. Are you going to see B. Are you seeing C. Have you seen D. Did you see E. Do you see	Amin: …your grandma last year John? I missed grandma so much. John: Yes, I saw my grandma last year. A.are you going to see B. are you seeing C. have you seen D. did you see E. do you see
9	Juvent: What did you send to Semarang last week? Ivan: I … some papers to Semarang last week. A. was sending B. am sending C. have sent D. had sent E. sent	Juvent: What did you send to Semarang last week? Ivan: I … some papers to last week. A. was sending B. am sending C. have sent C had sent E. sent
10	Toto: Where did you hide yourself thirty years ago? Yanti: I …. in Indonesian policewoman, how about you?**(In Indonesia)** A. was hidding B. have hidden C had hidden D/hide E.hid	Toto: Where did you hide yourself thirty years ago? Yanti: I …. In Indonesian policewoman, how about you? A. was hidding B. /have hidden C. had hidden D.hide E.hid
11	Questions 11 through 15 are based on the following passage I found a spider in the bathroom yesterday. I shouted for help, but no one came. There was no one else in the house. I needed to take a bath, but I couldn't because of the spider. So, I thought I was very brave! I'll have to take care of myself! I went to the kitchen and got a glass. Then, I went to the study room and got a paper. Next I went back to the bathroom. Then, I caught the spider in the glass. After that I	Exercise 3. Reading section Questions 11 through 15 are based on the following passage I found a spider in the bathroom yesterday. I shouted for help, but no one came. There was no one else in the house. I needed to take a bath, but I couldn't because of the spider. So, I thought I was very brave! I'll have to take care of myself! I went to the kitchen and got a glass. Then, I went to the study room and got a paper. Next I went back to the bathroom. Then, I caught the spider in the glass. After that I went to the window and threw the spider outside.

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

	went to the window and threw the spider outside. Finally, I had a bath. I did it myself! I was proud of myself (**What`s for)** What is the passage about? A. The writer caught and killed a spider in the house B. The writer could take a bath by himself C. The writer could not take care of himself D. The spider was in the glass E. Finally the writer was sad	What is the passage about? A. The writer caught and killed a spider in the house B. The writer could take a bath by himself C. The writer could not take care of himself D. The spider was in the glass E. Finally the writer was sad
12	What's the first event of the story? A. I found a spider in the bathroom yesterday**. / I diganti the writer** B. I went to the kitchen and got a glass. Then,/then hilang, **I diganti the writer** C. I'll have to take care of myself!/ The writer had to take care of himself D. After that I went to the window/I diganti the writer E. I needed to take a bath/ I diganti the writer	What's the first event of the story? A. The writer found a spider in the bathroom yesterday. B. The writer went to the kitchen and got a glass. C. After that the writerwent to the window D. The writerhad to take care of himself E. The writer needed to take a bat
13	**Finally**, I had a bath. I did it myself! I was proud of myself.The synonym of the underlined word is … A. in the beginning**(I must be changed I)** B. in the middle C.In front of D.At the end E. Firstly	**Finally**, I had a bath. I did it myself! I was proud of myself.The synonym of the underlined word is … A. In the beginning B. In the middle C.In front of D.At the end E. Firstly
14	Where did the writer throw the spider? A. In the house B. In the glass C. In a paper D. Outside**of the window** E. Inside	Where did the writer throw the spider? A. In the house B. In the glass C. In a paper D. Outside **of the window** E. Inside
16	**The followingtextquestions for number 16 -20 Fill in the blanks by choosing A, b, C, D, or** E. **Announcement** To .. 16.. the 23rd …17..of "SMA Indonesia Merdeka", Students **Association** will make some …18 .., include: - Sport : Basketball (M&F) and Soccer (M)	**The followingtextquestions for number 16 -20 Fill in the blanks by choosing A, b, C, D, or** E. **Announcement** To .. 16.. the 23rd …17..of "SMA Indonesia Merdeka", Students **Association** will make some …18 .., include: - Sport : Basketball (M&F) and Soccer (M)

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

	- Art : Singing and English Drama (M&F) - Class : The smartest student and **funniest** (M&F) All competitions will be …19..out on October 23 – 25. All member of Students Intra organization must come to the meeting on September the first. Further ..20.., please contact: 08882522529 (Emmi) 081326304278 (Juvent), 085643317687 (Vice), 0888234234 1 (OC) and 085577333123 (Registrar). Anita – The Chief	- Art : Singing and English Drama (M&F) - Class : The smartest student and **funniest** (M&F) All competitions will be …19..out on October 23 – 25. All member of Students Intra organization must come to the meeting on September the first. Further ..20.., please contact: 08882522529 (Emmi)
18	A union. B. addition C. simulation D. registration E.competitions	A. union. B. addition C. simulation D. registration E.competitions
21	Questions 21 through 30 are based on the following passage. There is a famous story **that is told** about George Washington, the first President of the United States, and a cherry tree. It seems that when George Washington was a boy, he had a little axe. One day, he used it to cut down a cherry tree that his father was very proud of. When his father saw what had happened to his cherry tree, he was very angry. He asked who had cut it down. George said, "I cannot tell a lie." He admitted that he had cut down the tree. His father was proud of George's honesty. This story was not actually true. However, it did demonstrate American's belief about Washington's honesty. What does the text tell about? A. Cherry tree B. Washington's dishonest**(B & C similar)** C. George Washington's dishonesty D. George Washington and the cherry tree E. The first president of the United States	Questions 21through 30 are based on the following passage. There is a famous story about George Washington, the first President of the United States, and a cherry tree. It seems that when George Washington was a boy, he had a little axe. One day, he used it to cut down a cherry tree that his father was very proud of. When his father saw what had happened to his cherry tree, he was very angry. He asked who had cut it down. George said, "I cannot tell a lie." He admitted that he had cut down the tree. His father was proud of George's honesty. This story was not actually true. However, it did demonstrate American's belief about Washington's honesty. What does the text tell about? A. Cherry tree B. George England's dishonesty C. George Washington and the cherry tree D. The first president of the United States E. He used it to cut down a cherry tree that his father was very proud of.
22	What is the main idea of paragraph 2? A. George was liar	What is the main idea of paragraph 2? A. George was liar

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

	B. His father was very angry/ George's father C. George cut down his cherry tree **(his father`s cherry tree)** D. George Washington, the first President of the United States, E. It seems that when George Washington was a boy, he had a little axe.	B. His father was very angry C. George cut down his cherry tree D. George Washington, the first President of the United States, E. It seems that when George Washington was a boy, he had a little axe.
23	"When his father saw what had **happened** to his cherry tree, he was very angry." What does the underline word have similar meaning? **Which one of the following words has the same meaning as the underlined word?** A. Occurred B. Handed C. Seemed D. Named E. Threw	"When his father saw what had **happened** to his cherry tree, he was very angry." What does the underline word have similar meaning? **Which one of the following words has the same meaning as the underlined word?** A. Occurred B. Handed C. Seemed D. Named E. Threw
24	Who is George Washington? **A.** First President of The United B. Owner of a cherry tree C. Demonstrate American**(The others are noun but C is not= not parallel)** D. A dishonest boy E. A proud father	Who is George Washington? **A.** First President of The United B. Owner of a cherry tree B. A demonstrate American C. A dishonest boy D. A proud father
25	"….one day, he used **it** to cut down a cherry tree that his father was very proud of." (Par. 2) The underlined word refers to …. A. Hoe **Because the answers are not in the first so the words start from small letters.** B. Saw C. Nail D. Hammer E. Little axe	"….one day, he used **it** to cut down a cherry tree that his father was very proud of." (Par. 2) The underlined word refers to …. A. hoe **B.** saw C. nail D. hammer E. little axe
26	The first paragraph of a recount text belongs to A. events B. orientation C. conclusion	The first paragraph of a recount text belongs to A. events B. orientation C. conclusion D. reorientation

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

	D. Reorientation**Be consistent. D & E start with small letters!** E. Complication	E. complication
29	What can you find in the first paragraph of a recount text? A. a sequence of events B. who, where,when C. an explanation D. pro and cons**(Pros and cons)** E. Description	What can you find in the first paragraph of a recount text? A. a sequence of events B. who, where,when C. an explanation D. pros and cons E. Description
30	Once upon a time there lived a beutifulprincess, named princess who lived with a poor family She was very kind. **She refers to ….** A. A beautiful princess B. A beautiful woman C. An old princess D. An old woman E. A woman	Once upon a time there lived a beutifulprincess, named princess who lived with a poor family She was very kind. **She refers to ….** A. A beautiful princess B. A beautiful woman C. An old princess D. An old woman E. A woman

4.5. The Old Versus the Developing Principles in Constructing Formative Assessment

No.	Old Principles	Developing Principles
1.	1. Designing each item to measure a specific objective, and Eliciting behavior under relativity controlled conditions	1.Test item must comply with indicator
2.	**2.** Stating both stem and options as simply and directly as possible	2.The arrangement of option should be constructed from the largest characters or from the least characters
3.	**3.** Capital letters in the beginning of sentences	3.Capital letter must be used in the

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

		initial position, names, or after full stops
4.	4. Distracters must be four options	4.Distracters must be grammatically correct
5.	Multiple-choice test do not require the examinee to write out and develop their answers, minimizing the chance for less knowledge examinees to “bluff” or “dress up” their answers	5.The stem of questions should be formulated clearly and firmly
6.	The questions do not answer itself	6.The stem of questions do not give clues to the correct answer
7.	Options should be carefully designed	7.Options should be parallel and logical in terms of material
8.	Test materials should be spaced so as to equip maximum readability;	8.The length of response options should be relatively the same
9.	All options must not show the correct answer	9.The option do not contain the statement "All options above answer is wrong / right"
10.	The position of numbers in options must be from the small ones	10.The options in the form of numbers or time must be arranged in order of the size of the numeric value or chronological time
11.	All figures must be arranged in sequent	11.Figures, graphs, tables,

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

	numbers	diagrams, and the like contained in the questions should be clear and functional
12.	The questions is constructed must be clear	12.The formulation of stem does not use the phrase or words must not be ambiguity
13.	Question must be right and logical to be interpreted by students	13.The stems do not depend on the answer to the previous question
14.	Correct choice means possible interpretations: the person surelyknows the answer, makes a lucky random guess, answer uses partial knowledge, answer uses testwiseness	Each assessment item has only one correct option

4.6. The Result of Try Out

Table 3: The Result of Statistical of Try Out I, Try Out II, and Try Out III by SPSS 20

Activity	N	Mean score	Std.deviation	T	df	Sig (2-tailed)	t-table
Try out I Try out II	35 35	72.0 78.0	8.84507 6.69943	-2.881	34	.007	2.03
Try out II Try outIII	35 35	78.0 81.4	6.69943 4.88868	-7.657	34	.000	2.03
Try out I Try outIII	35 35	72.0 81.4	8.84507 4.88868	-5.006	34	.000	2.03

t-table (2-tailed): 2.03

Based on Table 4: The result of Try out I, Try out II, and Try out III in the appendix shows that the interaction between students and their teacher is effective because the reasons below:.

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

1) The table above shows that the mean score of try out 2 is greater than the mean score of try out 1(78.0 > 72.0). It also shows that the value of $\mathbf{t_o}$is 2.881 with the level of significant is 0.007 and *df*is on 34. The value of t_t on *df* 34 at the level of significant 0.005 is 2.03. The value of t_o is greater than the value of t_t (2.881 > 2.03). It means that there is significant difference score of try out 1 and try out 2. Therefore it can be concluded that try out 2 is effective. It means that formative test Recount try out 2 is more effective than try out test Recount 1. The mean score of try out 3 is the greatest of all (81.4 > 78.0 > 72.0). It can be concluded that the development of formative test on Recount is proved effective based on the limited try out.

The researcher applied statistical calculation by using SPSS 20 in order to compare among the result of try out I, try out II, and try out III. The result was the table above showed that the mean score of try out II is greater than the mean score of try out 1(78.0 > 72.0). it is also showed that the value of $\mathbf{t_o}$is2.881 with the level of significant is 0.007 (0.007 < 0.005) and *df* is on 35. The value of t_t on *df* 35 at the level of significant 0.005 is 2.03. The value of t_o is greater than the value of t_t (2.881 > 2.03). It means that there is significant difference score of try out 1 and try out II. Therefore it can be concluded that try out II is effective. It means that formative assessment on recount try out II is more effective than try out test Recount 1. The mean score of try out III is the greatest of all (81.4 > 78.0 > 72.0). It can be concluded that the development of formative test on recount is proved effective based on the limited try out.

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

4.7. Findings Interpretation

1. From the result of need assessment, A.2 shows that composing formative assessment based on syllabus "is needed much" is 33 English teachers or 82.5%, and 7 English teachers or 17.5% answered "is needed". No teachers answered "isn`t needed" in both A.1 and A.2 questions in questionnaire. The conclusion from part A is that 82.5 % of English teachers need formative assessment.
2. The result the existing formative assessment constructed by the 15 English teachers in Semarang municipality was: three English teachers got A, They were from state senior high school 4, state senior high school 6, and , state senior high school 7, while six English teachers got B+. They were from state senior high school 2, state senior high school 3, state senior high school 10, state senior high school 11, and state senior high school 12. Six other English teachers got B. They were from state senior high school 1, state senior high school 5, state senior high school 9, state senior high school 13, state senior high school 15, and state senior high school 16 for complete data see appendices.
3. Based on the researcher`s finding that the model with 13 steps in 4..3. The Result of Model of Developing Formative Assessment on Recount, this flowchart is the model used by the researcher to develop formative assessment on recount. This model is the answer of research problem number 2.
4. The result of table 4 and table 5 of try out 1, try out II, and try out III show clearly that the interaction or learning process between students and teacher is effective.

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

CHAPTER 5

CONCLUSION AND SUGGESSTION

5.1 Conclusion

In a way to develop the formative assessment on recount, the researcher used three main stages: the need assessment, draft of assessment on recount, revision by the teachers in the same school, try out 1, judgment by the expert on assessment, try out II, final assessment on recount, try out III.

Based on the quantitative data and qualitative data above, the researcher draws a conclusion:

1. On December 10, 2012 the researcher collected 15 existing formative assessments constructed by English teachers in state senior high schools in Semarang municipality. He then analyzed them by using the tools: validity, number of items, and grammar. He found out that the result of 15 existing formative assessments are 3 English teachers got A, 4 English teachers got B+, 8 English teachers got B. The researcher took mark 80 or B as the standard of the developing formative assessment.
2. Based on the result of the need assessment;the writer formulates a model of flochart on model of developing formative assessment on recount for the ten graders of senior high school 11 Semarang as follows elaborating in flowchart 2.
3. The tryout was done in SMA 11 Semarang three times. It was done in class X IPA 4. The first try out was done on September 12, 2013, the second try out was done on September 26, 2013, and the third try out was done on October 10, 2013 in classroom of X IPA 4. The data were analyzed by using SPSS vol.16 to find out the significant difference among the try out test of recount text. the result of trying out will be presented in the

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

following discussion in details. The table 3 in chapter four shows that the mean score of try out 2 is greater than the mean score of try out 1(78.0 > 72.0). It also shows that the value of $\mathbf{t_o}$is2.881 with the level of significant is 0.007 and *df*is on 34. The value of t_t on *df* 34 at the level of significant 0.005 is 2.03. The value of t_o is greater than the value of t_t (2.881 > 2.03). It means that there is significant difference score of try out 1 and try out 2. Therefore it can be concluded that try out 2 is effective. It means that formative on recount try out 2 is more effective than try out test Recount 1. The mean score of try out 3 is the greatest of all (81.4 > 78.0 > 72.0). It can be concluded that the development of formative test on Recount is proved effective based on the limited try out.

4. In constructing formative test there are somerules for our guidelines, the main important rules are: test item must comply with indicator, the arrangement of option should be constructed from the largest characters or from the least characters, capital letter must be used in the initial position, names, or after full stops, distracters must be grammatically correct, the stem of questions should be formulated clearly and firmly, the stem of questions do not give clues to the correct answer, options should be parallel and logical in terms of material, the length of response options should be relatively the same, the option do not contain the statement "All options above answer is wrong / right", the options in the form of numbers or time must be arranged in order of the size of the numeric value or chronological time, figures, graphs, tables, diagrams, and the like contained in the questions should be clear and functional, the formulation of stem does not use the phrase or words must not be ambiguity, the stems do not depend on the answer to the previous question.

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

5.2. Suggestions

The researcher may suggest first for English teachers, second for the next researcher who are interested in assessment.

5.2.1. Suggestions for the English Teachers

1. The English teacher should try to start writing her/his own formative assessment by applying the guideline in constructing formative assessment
2. The researcher understands well that all English teachers are busy in developing their competence by reading the reference, writing interesting topics, on the other hand, the English teachers are better to write an assessment an item at least a day;
3. After writing an item of assessment then the English teachers should consult to their friends to discuss the item;
4. The English teachers should follow the organization of English teachers in every district, town, or city so that they are able to change their ideas of English development;
5. The English teacher should empower himself or herself to write the formative assessment;
6. The senior English teachers should revise the product of the English teachers on the formative assessment.

5.2.2. Suggestions for the next researcher

1. The study may focus on summative by analyzing the existing summative test in Semarang municipality;

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

2. Second, further study can focus on developing summative assessment in Central Java province.

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

REFERENCE

Asgher. 1999. `The Lived Experience of Formative Assessment Practice in a British Universities`, *Volume 6 Number 1, 1999; page 205 – 223, Journal of Further Higher Education*

Anderson Mark and Anderson Kethy. 2003. Text Type in English 1-2.South Yara: MacMillan Education Australia

BalaiPengembangandanPembinaanBahasaKementrianPendidikandanKebudayaan. 2011. Undang-UndangRepublik Indonesia Nomer 24 Tahun 2009 TentangBendera, danLambang Negara, sertaLaguKebangsaan.Jakarta: BalaiPengembangandanPembinaanBahasaKementrianPendidikandanKebudayaan

Bechman.(1990). Fundamental Considerations in Language Testing. Oxford: Oxford University Press.

Bechman and Palmer,1996.Language Testing in Practice.Oxford: Oxford University Press

Bennet. 2011. `Formative Assessment: a Critical Review`, Vol. 18 Number 1, February 2011; page 5 – 25, Assessment in Education Principles, Policy &Practice.

Bervick and Ross. 1996. `Cross-Cultural Pragmatics in Oral proficiency Interview Strategies`.

Bloom. 1971. Handbook on Formative and Summative Evaluation of Student Learning.New York: McGraw-Hill

Bookhart. 2011. `Successful Student`s Formative and Summative Uses of Assessment Information`, Volume 8, Issue 2, 2011, Assessment in Education Principles, Policy and Practice

Borg and Gall. 2003. Educational Research an Introduction**.**Seventh Edition.Boston : Library of Congress Cataloging-in-Publication Data.

Brown. 2004. Language Assessment Principles and Classroom Practices.*S*anFrancisco : Pearson Education, 10 Bank Sxtreet, White Plains,Ny 10606

Calmorin. 2004. Measurement and Evaluation**.** Mandaluyong City: National Bookstore.

Canale and Swain. 1980. `Theoretical bases of communicative approaches to second language teaching and testing`. Applied Linguistics 1, 1 – 47

Celce-Murcia. 1995. `Communicative Competence: a paedagogically motivated model with content specifications`. Issue in Applied Linguistics 2, 5-36.

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

Chamot O`Malley. 1994. The CALLA `Handbook: Implementing the Cognitive Academic Language Learning Approach`

Clarke. 1988. `The Short Circuit Hypothesis of ESL, Reading –or When Language Competence Interfere with Reading Performance`. IASL 114 – 124

Cowie& Bell. 1999. *`A Model of Formative Assessment in science Education`*, Vol 6 Number 1

Davidson. 2007. `The Identity of Language Testing`. Language Assessment Quarterly1, 1, 85 - 88

Davidson and Lynch. 2002, Testcraft: A Teacher`s Guide to Writing and Using Language Test Specifications.New Haven, CT: Yale University

Derewianka.1995. Exploring How Texts Work.New South Wales: Primary Teaching English Assosiation.

Elatia. 2003. `History of the Baccalaureat: a Study of Interactional Between Educational legislation, government Policy in Language Examination`

Fielding and Pearson. 1994. `The Many faces of Performance Assessment`. *Phi Delta Kappan 74 (6) 478*

Fulcher. 2003. Testing Second Language Speaking. London: Longman/Pearson Education

Fulcher and Davidson. 2007. Language Testing and Assessment an Advanced Resource Book.London: Routledge Taylor &Frands Group,

Garcia.1994. Preventing Inappropriate Referrals of Language Minority Students to Special Education. Wheaton, Md.

Grabe. 1988. `Reassessing the Term "Interactive" in Interactive Approaches to Second Language Reading`, 56 - 70

Haris.1969. Testing English as a Second language.New York: McGraw-Hill Book Company.

Halliday. 1973. `Relevant Models of Language` in Haliday, Explorations in the Functions of Language**.** New York: Elsevier North Holland.

Hartoyo. 2011. Language Assessment. Semarang: PelitaInsani Semarang. of Language.

Hodgson. 2012. `Effective Formative Assessment of Students Learning: a Study of a Statistical Course`, *Vol. 37, Issue 2,2012, Page 215 – 225,* Assessment & Evaluation in Higher Education

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

Hornby. 2010. Oxford Advance Learner`s Dictionary, International Student`s Edition. 8th Edition. Oxford: Oxford University Press.

Hughes. 2003. Testing for Language Teachers Second Edition.Cambridge: Cambridge University Press.

Hughes Arthur. 2008. Testing for Language Teachers Second Edition**.** Cambridge: Cambridge University press.

Iseni. 2006. `Evaluating Reading Text in the Text of English Sky 2`. Semarang: English Language Education Study Post Graduate Program Semarang University.

Kehoe. 1985. `Writing Multiple-Choice Test Items` Practical Assessment , Research Evaluation, 4, 9

Kramsch. 1986. `From Language Proffiency to Interactional Competence`. Modern Language Journal 70, 4, 366 – 372

Macdonald. 2004.`Developing Competence e-learners: the role of assessment`**.**Volume 29, Issue2, 2004, Assessment & Evaluation in Higher Education.

Markee. 2000. `Conversional Analysis`. Mahwah, Nj: Erlbaum

McNamara. 1997. Measuring Second Language Performance. London: Longman

Nitko.1983. Educational Test and Measurement Introduction. New York: Hartcourt Brace Jovanovich, Inc

Nunan. 1992. `Communicative Tasks and the Language Curricullum` TSOL Quarterly 25:279-295

O`Malley and Pierce. 1996. Authentic Assessment for English Language Learners for Practical Approaches for Teachers.Ontario: longman.

Ortega, Perez. 2013.`Second Language Assessment`.www. Slassessment.com

Peregoy and Boyle. 1993. Reading, Writing and Learning in ESL: A Resource Book for K-8 Teachers. New York: Longman

Philida.The Oxford ESOL Handbooks for Language Teachers.Oxford: Oxford University Press

Rini. 2006. `The Correspondence Between the Final Test and English Program Objectives at SD Muhammadiyah I Surakarta`. Semarang: The Graduate Program English Language Education State University Semarang

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

Routman. 1994. Invitations:Changing as Teachers and Learners K12. Portsmouth, NH: Heineman

Ruch. 1924. The Improvement of the Written Examination.Chicago: Scott, Foresman and Company.

Tinanbunan. 1988. Evaluation of Student Achievement. Jakarta: Depdikbud

Tuckman. 1975. Measuring Educational Outcomes Fundamental of Testing.New York: Harcourt Brace Jovanovich, Inc.

Umiyatun. 2010. The Problems of Writing Recount Encountered by the students of the State Junior High School 2 Purworejo in the Academic *Year 2009/2010*. The Graduate Program English Language Education State University Semarang

Underhill. 1982. The great reability validity trade-off: prolems in assessing the productskills in Heaton Language testing. London: Modern English Publications.

Vickkerman. 2001. `Assessment & Education in Higher Education`, *Vol. 34, Issue 2, 2001, Page 221-230*

Victoria. 2008. Oxford Learner`s Pocket Dictionary fourth edition. Oxford: Oxford Eastwood, John. 2008. Oxford Learner`s Pocket Grammar fourth edition.Oxford: Oxford University Press.

Wood. 1977. `*Multiple Choice : A State of report Evaluation in Education*`**.** International Congrress, 1977. 1. 191 – 280

Young.1988. `Interactional Competence: Challenge for validity` Paper in joint sympocium `Interdisciplinary interfaces with language testing.

---- assess.psu.edu

------- Undang-UndangRepublikIndonesiaNomer 20 Tahun 2003 TentangSistemPendidikanNasional.

------ Lampiran Permen No 69 tahun 3013 tentang struktur program Kurrikulum 2013

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

APPENDICES

Table 1 : Core Competencies and Basic Competences of High School / Madrasah Aliyah

Core Competences	Basic Competences
3. Understanding, applying, analyzing factual knowledge, conceptual, procedural based ingintahunya sense about science, technology, arts, culture, and humanities with knowledge of humanity, national, state, and civilization-related causes of phenomena and events, as well as applying procedural knowledge in a specific field of study according to their talents and interests to solve problem	3.1 Analyze social functions, text structure and linguistic elements of the text presentation of identity, according to the context of its use. 3.2 Analyze social functions, text structure and linguistic elements in the expression of praise winged (extended), as well as the response, according to the context of its use. 3.3 Analyze social functions, text structure and linguistic elements in an expression of concern (care), as well as the response, according to the context of its use. 3.4 Analyze social functions, text structure and linguistic elements to represent and inquire about intention of doing an action / activity, according to the context of its use. 3.5 Analyze the social functions, text structure, and linguistic elements of expression of congratulations winged (extended), as well as the response, according to the context of its use. 3.6 Analyze social functions, text structure and linguistic elements to represent and inquire about the actions / activities / events that do / happened in the past, which refers to the time that refers to the end, according to the context of its use.

Table 2: The result of Try out I, Try out II, and Try out III

NO	NAME	Result of Try Out 1 March7, 2014	Result of Try OutII March 14, 2014	Result of Try Out III March 28, 2014
1	Adelia Ayu Larasati	52	85	89
2	Angger Bagaskara	68	75	79
3	Ani Handayani	70	79	82
4	Anindia Putri Ramadhani	67	78	79
5	An-Nisa' Nur Ash-Shiddiqqiyah	67	91	92
6	Arienda Novianne Devy	70	82	85
7	Audini Putri	80	74	78
8	Bima Bintang Ramadhan	68	82	85
9	Dewi Jasmine Sulistyorini	85	74	78
10	Dian Permata Budi Kurniawati	60	83	85
11	Dyah Ayu Kartika Ningrum	80	76	79
12	Elsa Meidifa	75	83	85
13	Fairuz Khalishah Izdihar	75	83	85
14	Fia Sari Kusumawati	83	89	89
15	Icha Adellia Safra	78	81	85
16	Kenny Desnandi Reyhan	65	82	85
17	M. Aji Ma'ruf Saputro	85	80	83
18	Miftakhun Naja	67	80	84

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

19	Muhammad Yaffi	83	85	85
20	Nanda Aries Pratama	67	83	85
21	Noviana Rizky	82	68	78
22	Nur Rika Tri Rachmawati	72	62	75
23	Puspita Anggraini	70	78	79
24	Raditya Rizqy Wibowo	47	78	80
25	Rahadyan Widya Ramandana	70	76	78
26	Ricky Armidha Putra	78	83	84
27	Rifky Setya Mahardika	73	74	78
28	Rinenggo Pulunggono Sulistiyo	78	80	83
29	Rizki Riza Putri	83	68	73
30	Sellyanindah Pertiwi	60	85	88
31	Septa Trisna Yudha	70	79	80
32	Setyo Adi Wibowo	80	67	74
33	Siti Aulia Nurul Aini	72	69	76
34	Ulva Anandya	73	67	73
35	Yumna Wahida Assa	67	71	74
	Means	**72**	**78**	**81.4**

Table 3: Reviewing of Existing Formative Assessment on Recount by English Teachers in Semarang Municipality

No	Name	School	Fomative Assessment	Number of items	Review
1.	**Budiyono, Spd, Mom**	**SHS 1 Smg**	**Multiple-choice 4 distractors**	**15**	**Grammar:80 MC: 60 Score: 70 =B**
2.	**Benny SDW, SPd**	**SHS 2 Smg**	**Open answer**	**20**	**Grammar:85 OA:80 Score:82,5 B+**
3.	**ArietSetyayoga, M.A**	**SHS 3 Smg**	**Open answer**	**20**	**Grammar:85 OA:80 Score:82,5 B+**
4.	**DraRahayuningtyas, MPd**	**SHS 4 Smg**	**Multiple Choice 5 distractors**	**30**	**Grammar:90 MC:90 Score:90 A**
5.	**Tri Lestari. Dra**	**SHS 5 Smg**	**Summary**	**15**	**Grammar:72 Sum:70 Score:71 B**
6.	**DrsSuharno, MPd**	**SHS 6 Smg**	**Multiple Choice 5 distractors**	**25**	**Grammar:80 Mc:90 Score:85=A**
7.	**Frans Budi, SPd**	**SHS 7 Smg**	**Multiple-choice 4 distractors**	**30**	**Grammar:90 MC:80 Score:85=A**

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

8.	**MochJohari, MSi**	**SHS 8 Smg**	**Multiple Choice 5 distractors**	**25**	**Grammar:90 Mc:70 Score;80=B+**
9.	**DraLastaria**	**SHS 9 Smg**	**Summary**	**15**	**Grammar:90 Sum:70 Score:80=B**
10.	**DraAnieSetyowati**	**SHS 10 Smg**	**Multiple Choice 5 distractors**	**25**	**Grammar:90 Mc:70 Score;80=B+**
11.	**Dian Arini Rosita, SPd**	**SHS 11 Smg**	**Multiple-choice 4 distractors**	**30**	**Grammar:70 Mc:90 Score 80=B+**
12.	**FitriMarkorini, SPd**	**SHS 12 Smg**	**Multiple-choice 4 distractors**	**12**	**Grammar:90 Mc: 70 Score: 80=B+**
13.	**NasriSunarsih, SPd**	**SHS 13 Smg**	**Open answer**	**15**	**Grammar:70 Oa:70 Score:70=B**
14.	**Sri Lestari, SPd**	**SHS 15 Smg**	**Summary**	**15**	**Grammar:70 Sum:70 Score 70=B**
15.	**Dian Milasari, SPd**	**SHS 16 Smg**	**Multiple Choice 5 distractors**	**15**	**Grammar:70 Mc:70 Score 70=B**

Marking

85 – 100 A

76 – 84 B+

67 – 75 B

60 – 65 C

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

Appendix 1. Final Product on Developing Formative Assessment on Recount

Exercise 1: Listening

For each question from 1 to 5, five options are given. One of them is the correct answer. Make your right choice (A, B, C, D, or E) in the brackets provided.

Man: Hi my name is Amir. I am 15 years old. How do you do?

Woman: ….

1. What is the best response for the woman to reply?

A. How do you do?

B. I am very well

C. I am Tina

D. I am fine

E. I am O.K.

2. Man : It is time for me to go home. Good bye all.

Woman ….

What is the best response to give?

A. Good afternoon

B. Good evening

C. Good morning

D. Good noon

E. Good night

3.Man: What was the result of your English test?

Woman: I got ninty-five. I only missed one.

Man: ….

What is the best response for the man to reply?

A. Great, I'm happy to hear it.

B. Oh no, I`m sad to hear it

C. I`m sorry to hear it

D. You are so foolish

E. You are joking

4. Man : What's the film about?

Woman : Well, it's about the fighting of young heroes in the beginning of independence. Would you like to come to the movie with me tonight?

Man: ….

What is the best response for the man to give?

A. It is my business

B. It is not funny

C. Of course

D. Not at all

E. Not now

5. Diana: Tomorrow is my sister`s birthday. Would you please come to my house to celebrate her sister`s birthday.

Dani :I am so sorry, my mother has asked me to accompany her to go shopping.

What can we infer from the dialogue?

A. The woman shows her attention

B. The woman refuses his offer

C. The man shows his attention

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

D. The man refuses her offer
E. The woman greets him

Exercise 2 Grammar section

6. Astrid: I went to the 'Java' mall yesterday.
Neisya: What did you buy there?
Neisya: I … a pair of shoes.
A. will have bought
B. have bought
C. had bought
D. will buy
E. bought

7. Teacher: Did you visit Kintamanilake during lat vacation to Balilast year?
Bagas: No, I did not as it was raining heavily when I was there so I … it.
A. hadn't visited
B. didn't visit
C. had visited
D. was visiting
E. visited

8. Amin: …your grandma last year John? I missed her so much.
John : Yes, I saw my grandma last year.
A. are you going to see
B. are you seeing
C. have you seen
D. did you see
E. do you see

9.Juvent: What did you send to Semaang last week?
Ivan: I … some papers to Semarang last week.
A. was sending
B. am sending
C. have sent
C had sent
E. sent

10. Toto: Where did you hide yourself thirty years ago?
Yanti: I …. in Indonesian policewoman, how about you?
A. was hidding
B. have hidden
C. had hidden
D. hide
E. hid

Exercise 3. Reading section

Questions 11 through 15 are based on the following passage

I found a spider in the bathroom yesterday. I shouted for help, but no one came. There was no one else in the house. I needed to take a bath, but I couldn't because of the spider. So, I thought I was very brave! I'll have to take care of myself! I went to the kitchen and got a glass. Then, I went to the study room and got a paper. Next I went back to the

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

bathroom. Then, I caught the spider in the glass. After that I went to the window and threw the spider outside.

11. What is the passage about?

A. The writer caught and killed a spider in the house

B. The writer could take a bath by himself

C. The writer could not take care of himself

D. The spider was in the glass

E. Finally the writer was sad

12. What's the first event of the story?

A. The writer found a spider in the bathroom a day ago.

B. The writer went to the kitchen and got a glass.

C. After that the writer went to the window

D. The writer had to take care of himself

E. The writer needed to take a bat

13. **Finally**, I had a bath. I did it myself! I was proud of myself

The synonym of the underlined word is ….

A. In the beginning

B. In the middle

C. In front of

D. At the end

E. Firstly

14. Where did the writer throw the spider?

A. Outside of the window

B. Inside the house

C. In the glass

D. In a paper

E. In a house

15. So, I thought I was very **brave**!

The opposite of the underlined word is ….

A. generous

B. similar to

C. proud of

D. afraid of

E. gentle

Exercise 4.The following questions for numbers 16 -20 Fill in the blanks by choosing A, b, C, D, or E.

Announcement

To .. 16.. the 23rd …17..of "SMA Indonesia Merdeka", Students Association will make some …18 .., that include:

- Sport : Basketball (M&F) and Soccer (M)
- Art : Singing and English Drama (M&F)
- Class : The smartest student and funniest (M&F)

All competitions will be …19..out on October 23 – 25. All member of Students Intra organization must come to the meeting on September the first. Further ..20.., please contact:

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

08882522529 (Amel, Secretary) 081326304278 (Juvent, The Chief), 085643317687 (Bagas, Vice), 0888234234 1 (OC) and 085577333123 (Registrar).
Anita – The Chief
16.
A.debate
B. update
C. dedicate
D. celebrate
E. communicate
17.
A. universary
B. birthday
C. holiday
D. treaty
E. unity
18.
A unions.
B. additions
C. simulations
D. registrations
E. competitions
19.
A signed
B. carried
C. threw
D. run
E. put
20.
A prohibition
B. legalization
C. information
D. promotion
E. motion
Questions 21through 30 are based on the following passage.

There is a famous story about George Washington, the first President of the United States, and a cherry tree. It seems that when George Washington was a boy, he had a little axe.

One day, he used it to cut down a cherry tree that his father was very proud of. When his father saw what had happened to his cherry tree, he was very angry. He asked who had cut it down. George said, "I cannot tell a lie." He admitted that he had cut down the tree. His father was proud of George's honesty.

This story was not actually true. However, it did demonstrate American's belief about Washington's honesty.

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

21. What does the text tell about?
A. Cherry tree
B. Father's dishonesty
C. George Washington's dishonesty
D. George Washington and the cherry tree
E. The first president of the United States
22. What is the main idea of paragraph 2?
A. George was liar
B. His father was very angry
C. George cut down his father`s cherry tree
D. George Washington, the first President of the United States,
E. It seems that when George Washington was a boy, he had a little axe.
23. "When his father saw what had **happened** to his cherry tree, he was very angry."
Which one of the following word has the same meaning as the underlined word?
A. Occurred
B. Handed
C. Seemed
D. Named
E. Threw
24. Who is George Washington?
A. First President of The United
B. Owner of a cherry tree
C. An American idol
D. A dishonest boy
E. A proud father
25. "….one day, he used **it** to cut down a cherry tree that his father was very proud of." (Par. 2)
The underlined word refers to ….
A. hoe
B. saw
C. nail
D. hammer
E. little axe
26. The first paragraph of a recount text belongs to
A. events
B. orientation
C. conclusion
D. reorientation
E. complication
27. What is the tense mostly used in a recount text?
A. Past tense
B. Past perfect
C. Present tense
D. Past continuous
E. Present continuous

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

28. What is the purpose of Recount?
A. To entertain
B. To amuse
C. To retell the past event
D. To inform a certain information
E. To state opinion from different point of view
29. What can you find in the first paragraph of a recount text?
A. a sequence of events
B. who, where, ,when
C. an explanation
D. pros and cons
E. description
30. Once upon a time there lived a beautiful princess, named Halimah, princess who lived with a poor family. She was very kind.
She refers to ….
A. A beautiful princess
B. A beautiful woman
C. An old princess
D. An old woman
E. A woman

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

APPENDIX 2

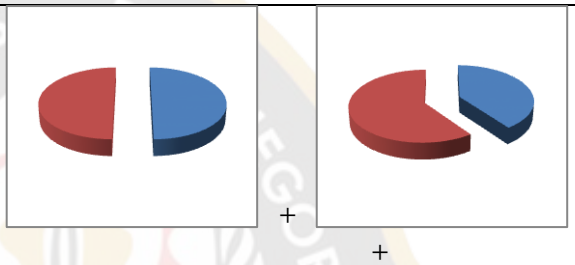
GUIDELINE OF CONCTRUCTING BEST MULTIPLE-CHOICE ITEMS

No.	Guideline of Constructing Formative Assessment on Recount	Item of Formative Assessment
1.	Test item must comply with indicator	Juvent: What did you send last week? Ivan: I … some papers to Semarang last week. A. was sending B. am sending C. have sent C had sent E. sent The indicator must be "The students can choose the past tense form of "send"
2.	The arrangement of option should be constructed from the largest characters or from the least characters	Exercise 1: Listening For each question from 1 to 5, five options are given. One of them is the correct answer. Make your right choice (A, B, C, D, or E) in the brackets provided. Man: Hi my name is Amir. I am 15 years old. How do you do? Woman: …. What is the best response for the woman to reply? A. How do you do? (**14 characters)** B. I am very well (**13 characters)** C. I am Tina **(9 characters)** D. I am fine **(9 characters)** E. I am O.K. (**7 characters)**
3.	Capital letter must be used in the initial position, names, or after full stops.	There is a new student in our class this week. The new student is John. He is from Semarang, Central Java. In what province is Semarang city? semarang belongs to …. A. Jakarta B. East Java C. West Java D. Yogyakarta E. Central Java semarang in the stem question is not correct. The name of town or city the first letter must be capital. It must be Semarang.
4.	Distracters must be grammatically correct	Amin: …your grandma last year John? I missed her so much. John: Yes, I saw my grandma last year. A.are you going to see B. are you seeing C. have you seen D. did you see E. do you saw -→ (do you saw is grammatically

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

		incorrect because do (Present) saw is (past) so the right is E. do you see
5.	The stem of questions should be formulated clearly and firmly.	Toto: Where did/do you hide yourself thirty years ago? Yanti: I …. In Indonesian policewoman, how about you? A. was hidding B. have hidden C. had hidden D.hide E.hid The students will be confused between did (Past tense) and do (Present tense). It is better did means Past tense is suitable for recount.
6.	The stem of questions do not give clues to the correct answer.	Recount is retelling the past events. Recount refers to…. events. A. last B. now C. next D. nowadays E. tomorrow Past gives a clue to last.
7.	Options should be parallel and logical in terms of material.	They were tired after swimming in the beach and then they went home happily. The paragraph above belongs to …. Of recount, A. tired B. event C. happy D, orientation e. re-orientation Options A & D are not pararellel because tired and happy belong to adjective, while event, orientation, and re-orientation belong to noun. Tired must be being tired, happy must be happiness
8.	The length of response options should be relatively the same.	Woman: Tomorrow is my sister`s birthday. Would **you please to come** to my house to celebrate my sister`s birthday, please come. Man :I am so sorry, my mother has asked me to accompany her to go shopping. What can we infer from the dialogue? A. The woman shows her attention B. The man shows his attention C. The woman refuses his offer because she must accompany her mother to go shopping D. The man refuses her offer E. The woman greets him **Option C is not relatively the same with A, B, D, and E.**
9.	The option do not contain the	The girl … to the market early in the morning for buying the need in a week. A. go

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

	statement "All options above answer is wrong / right"	B. went C. goes D. All statements above are right E. will go The option D is not correct. It is better omitted
10.	The options in the form of numbers or time must be arranged in order of the size of the numeric value or chronological time.	Sastro is the son of Mr. Dinata. Ali, Tomi, and Adi are the brothers of Sastro, while Emi, Tety, and Cecil are the sisters of Sastro. How many children does Mr. Dinata have? He has … children. A. 3 B. 4 C. 6 D. 6 E. 7 Option C is incorrect because C should be 5.
11.	Figures, graphs, tables, diagrams, and the like contained in the questions should be clear and functional.	 Mr. Ali received a half of the portion (blue figure) from his parents and Mrs. Ali received a quarter portion (blue figure) from her parents. How much are Mr. and Mrs. Ali receive from their parents? A. 1/4 B. 1/2 C, 3/4 D. 1 1/4 E, 1 1/2
12.	The formulation of stem does not use the phrase or words must not be ambiguity.	The genre of the text which explains the past events of the process of something happen is called a/an ….. A. recount B. narrative C. discussion D. hortatory E.Explanation The past event is ambiguity because it refers to a recount, while explains the process of something happen is an expalanation. It is ambiguity.
13.	The stems do not depend on the answer to the previous question.	1. Recount is … past events. A. arguing B. retelling C. describing

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University

		D. discussing E. explaining 2. Recount means retelling …. Events. A. past B. now C. next D. tomorrow E. nowadays Question number two can be answered by looking stem number one. Recount is retelling past event. The word "past" is clearly stated in stem question number one.

www.eprints.undip.ac.id © Master Program in Linguistics, Diponegoro University