CHAPTER III

RESEARCH METHOD

This chapter presents the research method which consists of research design, research hypothesis, variable of the research, research subject, research procedure, data collection technique, instrument, and data analysis technique.

A. RESEARCH DESIGN

This study was conducted to find out the answer of two research problems. First, to find out whether hypertext glosses is effective to improve the eight graders' reading comprehension of recount text at SMPN 2 Surabaya. Second, it is to find out the advantages and the disadvantages the eighth graders at SMPN 2 Surabaya obtain from leaning reading recount text using hypertext glosses.

Based on the research questions and objectives of study, the appropriate research design of this study was experimental research. The experimental research was the way to find out a causal relationship between two factors that are intentionally made by the researcher by setting aside other factors that influence. Quasy-experimental research design looks the same as true experimental research design, but it lacks the key ingredient random assignment.¹ It is the design which does not provide full control. As stated by Ary that full experimental research is

¹ William M. K. Trochim, *The Research Methods Knowledge Base*, 2nd Edition (Ithaca, N.Y: Cornell Custom Publishing, 1999). 215

not always possible to conduct a random selection of subjects in education research especially learning. This is because the subjects that have been formed are naturally in groups. Those are group of students in one class. In addition, it seems to be impossible for the researcher to manage all the students based on her desire.² There were two groups in this research; experimental and control group. Kothari stated that control group is the group which is exposed to usual condition while experimental group is a group which is exposed to special condition.³ In this design, the subject could not be randomly assigned to either the experimental or the control group.

The procedure of experimental research design was presented in the following chart:



Figure 3.1 The Procedure of Experimental Research Design

² Daniel Muijs. *Doing Quantitative Research in Education*. (London :SAGE Publication Ltd. 2004), 1.

³ C.R.Kothari. *Research Methodology: Methods and Techniques*, 2nd rev ed. (New Delhi : New Age International. 2004), 30.

Explanation:

- A : Experimental Group (8G Class)
- B : Control Group (8H Class)
- Y1 : The pre-test conducted before treatment (8G and 8H Classes)
- X : Treatment of reading recount text using Hypertext Glosses (8G Class)
- Y2 : The post-test conducted after treatment (8G and 8H Classes)

Based on the procedures of experimental research design, the researcher needed two classes as the sample of the research, the control and experimental groups. Here, the selection of the classes was based on the recommendation of the English teacher at SMPN 2 Surabaya. Then, the researcher gave pre-test (reading test) to the experimental and control group. The pre-test was a test before the treatment was applied. It was used to make sure that both groups had the similar ability in reading comprehension. If the result of pre-test showed that their reading score was similar, the next steps of this research could be continued. Yet, if their reading score was very different, it had to look for another group that had similar ability in reading comprehension.

After finding the groups that had similar ability in reading, the next step was conducting treatments to the experimental groups only. The treatments were conducted 3 times. During the treatment, the experimental group was taught reading recount text using hypertext glosses while control group was taught as usual (using traditional method).

Finally, the researcher gave post-test to the experimental group and control group. The post-test was used to measure the significant difference in reading ability between the experimental group who got treatments using hypertext glosses and control group who did not. The result of the post-test brought the researcher to answer whether hypertext glosses is effective or not in improving the eighth graders' reading comprehension of recount text.

Since this study tried to answer two research problems and the second was about what the advantages and disadvantages the eighth graders at SMPN 2 Surabaya obtain from reading recount text using hypertext glosses, so the researcher also gave questionnaire to the eighth graders in experimental group. They were given questionnaire after they were taught reading recount text using Hypertext glosses in the last of the third treatment (see appendix 1).

B. POPULATION AND SAMPLE

The important part of a research is population and sample. Population is all subjects that have certain quality or characteristics which is determined by the researcher. Meanwhile, sample is a part of the number of characteristics owned by the population⁴. As stated by Ary, there are three steps in sampling.⁵ Those are: population, accessible population, and sample.

1. Population

The first step in sampling is the identification of the target population. Population is subject/ object chosen by the researcher that has special quality

⁴ Sugiyono, *Metode Penelitian Kuantitatif Kualitatif dan R & D* (Bandung: Alfabeta, 2010), 80-81.

⁵ Donald Ary, Introduction to Research in Education (USA: Wadsworth. 2010), 149.

and characteristics to be learned and made conclusion.⁶ It is the large group to which the researcher wishes to generalize the result of this research. The target population of this research was the students of junior high school widely in Indonesia who have the same problem as the researcher has found.

2. Accessible Population

The accessible population was then determined from the target population. The selection of accessible population was conducted because it is difficult, expensive, and time-consuming to research the total population of the students of senior high schools in Indonesia. Here, SMPN 2 Surabaya was selected as the accessible population. The reason why this school was chosen as the accessible population was because this school has the characteristics as the total of population has. Moreover, the researcher chose Junior High School 2 Surabaya because it was considered as favorite school in Surabaya. Unfortunately, based on the researcher preliminary study when she did direct teaching practice (PPL) in this school, most of the students are in low quality in English especially in reading comprehension.

3. Sample

Sample is the process of selecting a number of individual or group (objects of research) for a study such the individual or group (the object of study) is representative of a larger group on the selected object (population).⁷

⁶ Sugiyono, *Statistika untuk penelitian* (Bandung : Alfabeta, 2010), 61.

⁷ Sumanto, Metodologi Penelitian Sosial & Pendidikan, (Yogyakarta: Andi Offset, 1995), p. 39

The researcher chose sample using random sampling where all of the samples had the same chance of being selected. The researcher used cluster random sampling. Using this method, the researcher divided population into groups and the sample was chosen per group rather than per individual.⁸ The sample of this research was the eighth graders of SMPN 2 Surabaya in the first semester in 2013-2014 academic years. In this school, there were 8 classes for eighth graders. The number of students in the eighth grade was for about 360 students and each class contained for about 36 students. But, only two classes were chosen as control group and experimental group. They were 8G and 8H. They were chosen because of teacher's suggestion and data of English score which showed that both classes were having the same low quality English and low comprehension of reading. Moreover, those classes were taught by the same teacher to avoid extraneous variable or variable that may compete with independent variable in explaining the outcome.⁹

C. RESEARCH HYPOTHESIS

This research was conducted to test the hypothesis which consists of two major hypotheses. They are null hypothesis and alternative hypothesis.¹⁰

1. Null hypothesis: Hypertext gloss is not effective to improve the eighth graders' reading comprehension of recount text at SMPN 2 Surabaya.

 Alternative hypothesis: Hypertext gloss is effective to improve the eighth graders' reading comprehension of recount text at SMPN 2 Surabaya.
Explanations:

 H_o will be accepted if $t_{value} < t_{table}$

 H_a will be accepted if $t_{value} > t_{table}$

 T_{table} is the score gotten from $t_{distribution}$, while t_{value} is the score gotten from calculation using the formula of t-test.

D. VARIABLE OF THE RESEARCH

Variable is any feature or aspect of an event function or process that, by its presence and nature, affects some other event or process which is being studied.¹¹ Variable is a construct or a characteristic that can take on different values or scores¹². There are two variables in this research. They are independent and dependent variable.¹³

1. Independent variable

Independent variable is a variable which is selected and manipulated by the researcher.¹⁴ It is a variable that is presumed to cause changes to occur in another variable, a causal variable. In this research, hypertext glosses is independent variable.

¹¹ Yogesh Kumar Singh, *Fundamental of Research Methodology and Statistic* (Publish New Age International Publishers, 2006), 136.

¹³ Donald Ary, Introduction to Research 266

2. Dependent variable

Dependent variable is the variable which is the effect of the changes is observed by the researcher, but it is not manipulated.¹⁵ It is a variable that changes because of another variable, the effect or outcome variable. The dependent variable in this research is the eighth graders' reading comprehension of recount text.

E. RESEARCH PROCEDURE

According to Yogesh, there are some steps of the experimental research. They are as follows:¹⁶ (Detail research procedure, see appendix 1)

1. Selecting and delimiting the problem.

Based on background of study, there were two research problems which were important to be researched and answered in this study. First, to find out whether hypertext glosses is effective to improve the eighth graders' reading comprehension of recount text and second, to find out the advantages and disadvantages the eighth graders at SMPN 2 Surabaya obtain from leaning reading recount text using hypertext glosses. They were important to be researched since there was no research about it yet and it was important to know the result as reference of alternative strategy in teaching reading.

2. Reviewing the literature

Literature which involved the framework of the study and previous research of the study was very important as the foundation of this study. In this study, the researcher collected as many literatures related to Hypertext glosses such as the definition, example, the advantages and disadvantages and the explanation about how to teach reading comprehension of recount text using hypertext glosses. The literatures are taken from book, journal, thesis, dissertation and other sources.

3. Preparing the experimental design

The researcher had prepared the proposal and some instruments to do the research. The researcher also prepared the schedule for the research, place and subject of study. Moreover, the researcher did a preliminary study in the school chosen.

4. Defining the population

In this phase, the researcher did the preliminary research to obtain the required information dealing with the way and the form of teaching and learning English at SMPN 2 Surabaya. The writer narrows down his focus on reading comprehension.

5. Carrying out the experiment

In this research, the researcher did some steps:

a. Conducting try out

The researcher had administered try out to examine the reliability and validity of the instrument before conducting pre-test, treatment and post-test,. Try out was held on July 18th, 2013. The subject of tryout was the eighth graders at SMPN 2 Surabaya, out of experimental (8g class) and control group (8h class). The researcher chose 8D class based on English teacher recommendation.

b. Giving pre-test to the experimental and control groups

The pre-test was held for both experimental group (8g class) and control group (8h class). It was held before giving any treatments to the experimental group. Pre-test in experimental group was conducted on Tuesday, July 23rd, 2013 in the class room and pre-test for control group was conducted on Wednesday, July 24th, 2013.

c. Conducting treatments to the experimental group using hypertext glosses

There were three times treatments for experimental group using hypertext glosses. In the treatments, students who were in experimental group were taught using hypertext glosses as media in teaching reading of recount text. The first, second and third treatment were actually the same. The difference was in the story of recount text. The treatments were held for experimental group while control group were taught as usual using traditional media. d. Giving post-test to the experimental and control groups

The post-test was held for both experimental group (8g class) and control group (8h class). It was held after giving treatments to the experimental group. Post-test in experimental group was conducted on Friday, August 23rd,2013 in the class room and post-test for control group was conducted on Thursday, August 22nd, 2013.

e. Giving questionnaire to the experimental groups

Since this research tried to answer two research questions, the researcher had to distribute questionnaire to find out what the advantages and the disadvantages the eighth graders at SMPN 2 Surabaya obtain from learning reading recount text using hypertext glosses. It was given after the third treatment on Friday, August 23rd, 2013, to the experimental group only.

6. Measuring the outcomes

After conducting the pre-test, post-test from both groups and questionnaire from experimental group, the researcher measured the data to answer the research questions stated in chapter 1.

7. Analyzing and interpreting the outcomes.

The researcher was analyzing the first data from pre-test and post-test using t-test formula. This formula was considered appropriate to be used for quantitative research and to answer the first research question. To analyze the data from questionnaire, the researcher calculated the frequency of students' answers in questionnaire then give statistic description.

8. Drawing up the conclusion

After analyzing the data, the researcher drew the conclusion of this study.

F. DATA AND RESOURCE OF DATA

To answer research questions stated in chapter 1, there were two data needed in this study. They are the result of test and questionnaire:

1. The Result of Tests

The result of test (English reading score) was data needed to answer the first research question about whether hypertext glosses is effective to improve the eighth graders' reading comprehension of recount text or not. There were two data gotten from test. The first was reading pre-test score and the second reading post-test score. The score of reading pre-test was gotten from experimental group and control groups before any treatments meanwhile the score of reading post-test was gotten from experimental group (after getting treatments using hypertext glosses) and control groups (after learning reading as usual/using traditional method).

2. The result of Questionnaire

The result of questionnaire (students' answer) was data needed to answer the second research questions about the advantages and the disadvantages students obtain from learning reading using hypertext glosses. The questionnaire's answer was gotten from the eighth graders in experimental group after they were taught reading recount text using hypertext glosses.

G. DATA COLLECTION TECHNIQUE

Data collection technique is the way or procedure to collect data needed to answer research questions.¹⁷ In this research, the data was collected from the tests and questionnaire. In order to collect all of the data needed in this study, the researcher has done some phases:

1. Collecting data from test

- a) To collect data from test, the researcher has conducted try out to examine the validity and reliability of the test instrument. It was hold once to different class (out of experimental and control group). The researcher has changed the items of the instruments which were not valid and reliable. Then, data collection was started by pre-test and ended by post-test.
- b) To obtain the data dealing with the origin of the students' reading comprehension, pre-test was administered to both control and experimental groups. Pre-test was to measure the eighth graders' reading comprehension of recount text before getting treatments. In addition, it aimed to know whether both of the groups had balance skill in reading before the application of hypertext glosses. Pre-test was also as consideration whether the researcher could continue doing the research in the classes chosen or

change the class. The data from pre-test was in the form of students' score dealing with reading comprehension of recount text.

- c) To give the treatment, the researcher implements Hypertext glosses as media to teach reading of recount text in experimental group. In this case, the researcher gave three times treatments. The output of this research was the improvement of the eighth graders' reading comprehension. On the other hand, the control group was not treated by hypertext glosses but the traditional media.
- d) To find out whether hypertext glosses is effective in improving the eighth graders' reading comprehension of recount text or not, the researcher gave post-test to measure reading ability of experimental and control group. The data was in the form of students' reading score. Then, the score of pre-test and post-test of control and experimental groups were calculated and compared, if the score showed that experimental group who were taught using hypertext glosses got better score than those who were not, it indicated that hypertext glosses is effective to improve students' reading comprehension. Yet, if the calculation of pre-test and post-test of both groups were stable (no significant difference) or even the experimental group was getting lower score, it meant that hypertext glosses was not effective as media to improve students' reading comprehension.

2. Collecting data from questionnaire

- a) The researcher has conducted test and retest (tryout) to examine the validity and reliability of the instrument before collecting data from questionnaire. It was held once to the class out of experimental and control group. The researcher has changed the items of the instruments which were not valid and reliable.
- b) The researcher distributed the questionnaire to the experimental group after they got treatments using hypertext glosses, to obtain data dealing with the advantages and disadvantages the eighth graders obtain from learning reading recount text using hypertext glosses.

H. INSTRUMENT OF THE RESEARCH

To get the data needed to answer research questions in this study, the researcher used some instruments. Instrument is the measurement tool in the test which potentially made the researcher easier in collecting data and analysis.¹⁸ They were: Test and Questionnaire.

1. Paper Based Test

The researcher will use Test to answer the first research question about whether hypertext glosses is effective to improve the eighth graders' reading comprehension of recount text or not. Brown stated that a test is a method of

measuring a person's ability or knowledge in a given domain.¹⁹ It was divided into two, pre-test and post-test:

- a. Pre-test : It is preliminary test administered to determine a students' baseline knowledge for an educational experience or course of study.²⁰
- b. Post-test : It is a test given to students after completion of an experiment that used to measure the students' ability after getting treatments.

The pre-test and post-test were about reading comprehension of recount text. The researcher gave pre-test and post-test to the both control and experimental groups. The pre-test and post-test were in multiple choices with 4 options. They did pre-test or post-test in 60 minutes. The amount of questions was 40 items contained of 5 instructional objectives: (see appendix 3)

- Identifying references/ meaning/ synonym/ antonym based on context of word or phrase of the text
- 2) Identifying the generic structure of the text
- 3) Identifying main idea
- 4) Identifying explicit information of the text
- 5) Identifying implicit information of the text

The test as an instrument needs to be valid and reliable. As stated by Weigle that a good instrument should be valid and reliable.²¹ Therefore, it is

¹⁹ H DoughlasBrown. *Teaching by Principles: An Interactive Approach to Language Pedagogy, 2nd ed.* (New York: Addison Wesky Longman, Inc.2001), 146.

²⁰ Oxford Learners' Pocket Dictionary. *Dictionary*, 3^{rd} ed. (New York : Oxford University Press.2008)

²¹ Sara Cushing Weigle, Assessing Writing, (Cambridge: Cambridge University Press. 2002), 49.

important to know the validity and the reliability of the test. The researcher needed some tests to measure whether the reading test made by the researcher was valid or not. First, the researcher made reading test instrument based on indicator/ content validity which was suitable with English standard competence. After that, the researcher administered the try out to the eighth graders at SMPN2 Surabaya, out of experimental and control group. And then, she analyzed item difficulty and item discrimination of reading test instrument and the last calculated the reliability. The concept of item difficulty, item discrimination, validity and reliability are discussed below:

a. Item difficulty

Item difficulty can be also called as item easiness. It expresses the proportion or percentage of students who answered the item of instrument (test) correctly. Item difficulty can range from 0.0 (none of the students answered the item correctly) to 1.0 (all of the students answered the item correctly). The average level of difficulty for a four-option multiple choice test should be between 60% and 80%. The average level of difficulty value, less than .25, there are several possible causes: the item may be too challenging relative to the overall level of ability of the class; the item may be ambiguous or not written clearly; there may be more than one correct

²² OMET. 2013, *Item difficulty and Item Analysis* (http://www.omet.pitt.edu/docs/OMET%20Test%20and%20Item%20Analysis.pdf accessed on May 4th 2013)

answer. On the contrary, if an item has high difficulty value, more than .80, the item can be too easy for students.²³

For the instrument (tests) in this research, the researcher had administered tryout to the eighth graders (out of experimental and control group) at SMPN 2 Surabaya to find out the average level of difficulty of the test. From the tryout result, then the researcher calculated using this following pattern:

$$p = \frac{B}{T}$$

Explanation:

P = Level of difficulty

B = The amount of students who answer correctly

T = The amount of students who answer the test

Here is the classification of item difficulty which is intended for this research:²⁴

Level of difficulty	
Range Result	Level of Difficulty
0.00 - 0.20	Very Difficult
0.21 - 0.29	Difficult
0.30 - 0.70	Enough
0.71 - 0.80	Easy
0.81 - 1.00	Too Easy

Table 3.1

From the calculation (see appendix 4), the result showed that the level of students difficulty was about 0.56 - 0.64. It means that the entire item of test which was tested was enough for students. It was not too easy or too difficult to do. However, it still had to be analyzed per item to find out item discrimination of the test.

b. Item Discrimination

The point-biserial correlation is an index of item discrimination. The point-biserial showed how good the item serves to discriminate between students with higher and lower levels of knowledge. It reflects the degree of relationship between scores on the item - 0=incorrect, 1=correct - and total test scores. Thus the point-biserial will be positive if better students answered the item correctly more than poorer students did, and negative if the opposite occurred. A negative point-biserial is showed by a minus sign in front of the value.²⁵

The value of a positive point-biserial discrimination index can range between 0 and 1; the closer the value is to 1, the better the discrimination. (The value of a negative point-biserial discrimination index can range between -1 and 0, but positive values are desirable). Item discrimination is really influenced by item difficulty. As a general rule, point-biserial values of .30 and above are considered to be desirable. Items with negative

²⁵ OMET. 2013, Item difficulty and Item Analysis (http://www.omet.pitt.edu/docs/OMET%20Test%20and%20Item%20Analysis.pdf accessed on May 4th 2013)

discrimination values should be revised or eliminated because maybe the item was ambiguous or misleading. The cause of negative discrimination can often be gained by examining the distribution of student responses.²⁶

In this study, the researcher had calculated item discrimination of the test from tryout result (see appendix 4). Based on the calculation, the result showed that there were 30 items from 40 questions (multiple choices) had discrimination score more than .30. Since the discrimination of test item values of .30 and above are considered to be desirable, the test item was considered appropriate to be used and could be used to differ students who have high and low ability in reading. However, the researcher needed to revise the items which were not applicable.

c. Validity of the Test

The instrument administered to the students must be valid. An instrument is valid if it can successfully measure what is intended to measure.²⁷ There are three kinds of validity. They are face validity, content validity and construct validity.²⁸ In this study, the content validity might be used. It was because the test should be constructed as the representative of the course. Content validity depends on careful analysis of the language being tested and of the particular course objective. The basic competence of

²⁶ Angelica Hotiu. Thesis: "The relationship between Item Difficulty and Discrimination

²⁷ Arthur Hughes," *Testing for Language Teachers*", (Cambridge: Cambridge University Press. 2003),

teaching English for the eighth grade based on Standard of Content (Standar Isi) is responding meaning and rhetoric in simple essay accurately and fluently related to neighborhood in Descriptive and Recount text.²⁹

The test items were accordance to the basic competence with five indicators (see appendix 2). Since the test items fulfilled the basis of indicators stated in basic competence of teaching reading for the eighth graders, it can be concluded that the test items are valid. In this research, recount text was used, because it related to standard curriculum. In the test, the students of both experimental and control were asked to answer reading test of recount text. Besides, the instrument was validated by the English teacher in the school as well. Finally, the test can be said as a valid one after the English teacher agreed that the test represented the curriculum.

d. Reliability of the Test

Reliability is a necessary characteristic of any good test. The instrument is reliable if the instrument result is stable when it is scored at different administration to the same group. Reliability is the extent to which a test measures consistently.³⁰

After calculating item difficulty and discrimination of the test, the researcher still needed to calculate the reliability because the reliability depends on the result of item discrimination. If the amount of item which

²⁹ National Education Department, Depdiknas, *Kurikulum Pendidikan Sekolah Dasar*. (Jakarta: Depdiknas, 2006).

were appropriate and could be divided by three, i.e. 30 items had value discrimination more than .30, the researcher could use Kr-21 than Kr-20 to calculate the reliability.

In order to measure the reliability of the test, in this study, the researcher used internal consistency reliability. It measures the consistency within the instrument (questions of the test), whether it is reliable or not. The procedure of internal consistency reliability is by administering tryout once and using formula. The formula could be used appropriately for this tests was Kuder-Richardson formula 21 or it is known as Kr-21. The formula is as

follow:

 $r11 = \left[\frac{K}{K-1}\right] \left[1 - \frac{x(K-x)}{K.Sdx^2}\right]$

Explanation:

K = the number of items in the test

 \bar{x} = mean of the scores

 $K.Sdx^2 = Variance$

Internal consistency coefficients can take value from 0 to 1. Higher

values represent higher level of internal consistency. Here are criteria of the

reliability of the test according to Bartz:³¹

Criteria of reliability	
The criteria	The Description
Very High r	.80 or above
Strong <i>r</i>	.60 to .80
Moderate <i>r</i>	.40 to .60
Low r	.20 to .40
Very low <i>r</i>	.20 or less

Table 3.2

³¹ A.E.Bartz, *Basic Statistical Concept in Education and the Behaviour Science* (Minneapous: Burgess Pub.Co, 1976). p.293

The criteria above were used to measure the reliability of test items in this research. Those criteria indicate that the test can be considered good or reliable if its range of value is .60 until .80. In the contrary, the test is considered poor or unreliable if the value less than .40.The test which is not reliable cannot be used and must be revised.

In order to know whether the test items were reliable or not, the researcher had administered tryout. The tryout was given to the eighth graders in SMPN 2 Surabaya, VIII A class. Based on the calculation of the tryout, the tryout score showed the strong criteria of reliability (see appendix 6). The reliability was 0.76. Based on the criteria on table 3.2, the test was reliable to be used in this research. Thus, it could be used for the instrument of this research (see appendix 5 and 6).

2. Questionnaire

To answer the second research question about the advantages and disadvantages the eighth graders obtain after learning reading recount text using hypertext glosses, the researcher used questionnaire.

A questionnaire is a set of questions on a topic or group of topics designed to be answered by a respondent.³² In this case, the questionnaire was used to get students' respond of advantages and disadvantages they get from

³² Jack C.Richards and Richard Schmidt, *Longman Dictionary of Language Teaching and Applied Linguistic*, 4th ed (London: Pearson Education Limited, 2010), 593.

learning reading recount text using hypertext glosses. Questionnaire which was used for this study was open-ended questionnaire. It was a questionnaire which gave students opportunity to give their own opinions. There were 16 items which involved students' reading comprehension before treatment, the strength of hypertext glosses on students' reading comprehension and the weakness of hypertext glosses. The students had to answer the questionnaire in 15 minutes. (see appendix 8 or 9)

a. Validity of the Questionnaire

The validation is also needed for questionnaire. In this study, the researcher used content validity to measure whether the questionnaire items are valid or not. Content validity depends on careful analysis of the language being tested and of the particular indicator. Because of the second research questions of this study was about advantages and disadvantages students obtain from learning reading recount text using hypertext glosses, therefore the items of questionnaire involved the advantages and disadvantages during and after the implementation of hypertext glosses such as the strength, the weakness and the improvements students got from learning reading using hypertext glosses (see appendix 7).

b. Reliability of the Questionnaire

The researcher had measure the questionnaire using test and retest reliability (tryout) to measure whether the questionnaire was valid or not. The researcher administered the questionnaire to the eight graders at SMPN 2 Surabaya twice (out of experimental and control group). Based on English teacher' recommendation, the questionnaire was administered in 8D class in language laboratory. They were given questionnaire after the researcher introduced hypertext glosses to them and showed the use of hypertext glosses. Then, they were asked to answer the questionnaire. It was used to know whether the questions could be understood well by the students, whether the choices are ambiguous or not, whether the questions were related to the indicator or not. The items which were not reliable would be revised then it would be administered again (retest) to the same respondents (8D) to measure whether the questionnaire was really reliable to be used as instruments in this study.

I. DATA ANALYSIS TECHNIQUE

There were two data which would be analyzed in this study. The first was data from test to answer the first research question about whether hypertext glosses is effective to improve the eighth graders' reading comprehension of recount text at SMPN 2 Surabaya. The second was data from questionnaire to answer the second research question about what the advantages and the disadvantages the eighth graders obtain from learning reading using hypertext glosses.

After getting the data from tests (pre-test and post-test of experimental and control group) and the result of questionnaire, the researcher began analyzing the data. Here are the techniques to analyze the data gotten.

1. Test Analysis Techniques

The researcher would use some steps / tests of data to analyze data gotten from the test (reading score in pre-test and post-test). They are Normality test, Homogeneity Test and T-test. Here are the explanations:

a. Normality Test

Normality test is a test that used to determine whether a data set is well-distributed or not.³³ In this research, the purpose of it is to check whether the test scores of the experimental and control groups have normal distribution or not.

In this study, the researcher analyzed the normality of data (reading score) from pre-test and post-test of experimental and control group using One Sample Kolmogorof – Smirnof Test in SPSS.16.³⁴

b. Homogeneity Test

Homogeneity test is a test that used to determine whether the data is from the same variance or not.³⁵ In this research, the purpose of homogeneity test is to find out whether the test scores of both groups have same variance or not.

³³ Sugiono, Statistik untuk Penelitian 80

³⁴ Drs. Syahri Alhusin, Ms. Aplikasi Statistik Praktis Dengan menggunakan SPSS 10 for windows.

In this study, the researcher analyzed the homogeneity data from pretest and post-test of experimental and control group using Chi Square Test with manual calculation. The patterns are:

a) Find the biggest variant score and the smallest variant score, the formula is:

Fhit = $S1^2$ $\overline{S2^2}$ Explanation: S_1^2 = the larger of variance S_2^2 = the smaller of variance b) Find the F score *dkpembilang* :n-1 = 25*dkpenyebut* : n-1 = 25

c. T- Test

The aim of t-test is to know whether hypertext glosses is effective to improve the eighth graders' reading comprehension of recount text at SMPN 2 Surabaya. Based on the key term in this research, effective means the students who were taught using hypertext glosses got higher score than those who were not.

There should be an alternative hypothesis (H_a) and null Hypothesis

- (H_o) to answer the questions above.
- H_a: Hypertext glosses is effective to improve the eighth graders' reading comprehension of recount text at SMPN 2 Surabaya.
- H_o: Hypertext glosses is not effective to improve the eighth graders' reading comprehension of recount text at SMPN 2 Surabaya.

Next, the students' score of the test is calculated using the T-test formula. T-test formula which is appropriate for this kind data is T-test paired sample.

T-test paired sample is used to measure the effect of independent variable to the dependent variable. The pattern is below:

$$t = \frac{Mx - My}{\sqrt{\left[\frac{\sum X^2 + \sum Y^2}{Nx + Ny - 2}\right]\left[\frac{1}{Nx} + \frac{1}{Ny}\right]}}$$

Explanation:

t = the t-test Mx = Mean deviation of experimental group My = Mean deviation of control group $\sum X^2$ = The sum of mean of experimental group $\sum Y^2$ = The sum of mean of control group Nx = The sum of students in experimental group Ny = The sum of students in control group

2. Questionnaire Analysis Technique

In this research, there are sixteen items of questionnaire with 5 answer choices. One of them is blank column which is used to put the students' opinion. The researcher would use Statistic Descriptive Frequency to analyze this questionnaire. Statistic Descriptive Frequency is one of analysis technique which emphasizes in presenting the frequency of questionnaire' answers. The researcher used this analysis technique because the researcher would like to describe the result of questionnaire, given to the experimental group, based on the frequency of the respondents who answered the questionnaire. The result will be described from the most answer to the fewest to know which answer choices which students agreed or suitable with their condition before, during and after the implementation of hypertext glosses.