CHAPTER III

RESEARCH METHOD

In this chapter the researcher describes the research method. It consists of research design, population and sample, research instrument, validity and reliability testing, normality and homogeneity testing, data collecting method, and data analysis.

A. Research Design

To understand the meaning of research is truly necessary, because it is impossible to come to the points of research without it. According to Hornby (1995:996) research is careful study on investigation, especially in order to discover new facts or information such as scientific historical research. It means that a study is done carefully and accurately on investigation of an event, problem, and phenomenon about scientific to find out new information. According to Soekanto (1986) a research, especially in the empirical science, in general, aims to discover, develop, or test the truth of knowledge (Tanzeh. 2009:12). It means that a research is conducted to get answers of certain questions or to get solutions of the problems. A design is the general plan for carrying out the experimental research study (Ary et al, 2006:325).

Research design is a plan on how to collect and process data that can be implemented to achieve the research objectives. This study is conducted in an experimental design. Experimental research design is to enable researcher to estimate the effect of an experimental treatment (Ary, et al, 2006:325).

According to Ary, (2002:22) states that quantitative research uses objective measurement and statistical analysis of numeric data to understand and explain phenomena. “There are many kinds of the experimental researches, such as pre-experimental, randomized experimental or quasi experimental” (Ary et al, 2006: 326). Experimental research involves a study of the effect of the systematic manipulation of one variable on another variable and non-experimental research.

According to Ary, (2002:276) an experimental is a scientific investigation in which the researcher manipulates one or more independent variables, controls any other relevant variables, and observes the effect of the manipulations on the dependent variables. An experimental design serves two functions:

1. It establishes the condition for the comparisons required to test the hypothesis of the experiment.
2. It enables the experimenter through statistical analysis of the data to make a meaningful interpretation of the result of the study.

This experimental design used pre-experimental research design (one-group pretest-posttest design) that consist of pre-test, treatment and post-test. The pre-test and post-test were given to take the score of the student’s ability before and after being taught using Snowball Throwing.

Table 3.1. The design of one group pre-test post-test

Pretest	**Treatment**	**Posttest**
Y1	**X**	**Y2**

The procedures of experimental research that use one group pre-test post-test design applied in this study are:

1) Administrating a pre-test with a purpose of measuring speaking skill mastery of eight grade students at SMPN 5 Tulungagung.
2) Applying the experimental treatment in teaching speaking by applying Snowball Throwing Technique.
3) Administrating a post-test with a purpose of measuring speaking skill mastery of eight grade students at SMPN 5 Tulungagung.

Difference attributed to the application of the experimental treatment was determined by comparing the pre-test and post-test scores. In this research, the researcher wanted to know *the effectiveness of using Snowball Throwing Technique* in teaching speaking by conducting pre-experimental research. The impact was assessed by providing a specific treatment. The effectiveness of the technique was known after knowing the significant differences between the students who were taught before and after applying Snowball Throwing Technique.

B. Population, Sampling and Sample

1. Population

Population is the whole of subject in the research. According to Ary et al (2010:148) "a population is defined as all members of any well defined

class of people, events, or subjects". In this research, the population were the eight class grade students of SMPN 5 Tulungagung in academic year 2015/2016.

2. Sampling and sample

In the research, the researcher used purposive sampling. Purposive sampling refers to as judgment sampling sample elements judged to be "typical" or representative are chosen from population. To select sample was very important step in conducting a research. According to Ary et al (2010:148) a sample is portion of population. The researcher, choose one class is 8 H because the students have homogenous speaking ability compared to other classes. The meaning of having homogenous ability in this study is the students have similar score when they got examination especially in speaking exam. It was proved by the researcher when she conducted since Practice Teaching in the school and she taught the class, therefore she knew the students' speaking level of the class.

C. Research instrument and Data Collection Method

Instrument is a tool used to collect the data in order to overcome the problem of the research (Moehnilabib, et al, 2003: 71). The researcher used test as the instrument in collecting data. In this research the researcher used pre-test and post-test as the instruments.

1. Pre-test

Pre-test was administered before the students were taught using Snowball Throwing Technique. The researcher gave pretest on February, 10 2016. The pretest was administrated to know the students' speaking

skill before being taught by using Snowball Throwing Technique. The pretest the story was Malin Kundang. Each student was given 5 minutes to tell the narrative story.

2. Post-test

Post-test was administered on February, 20 2016. This post-test was intended to measure students' speaking ability after being given a treatment. In this test, the students were asked to tell a narrative story about The Mouse Deer and Crocodile. The students were given 5 minutes to tell the narrative story.

In this research, the researcher conducted the tryout of the test. The purposes of trying out the instrument, in this study were to know whether or not the instruction was clear and to convince that the students were familiar with the story. The researcher did the tryout twice, it was tryout of it should pretest and posttest.

The criterion of success of the students speaking ability adapted and matched from Hughes (1989). They were as follow:

Table 3.2 Analytic oral language scoring rubric

Aspect	**Need Improvement 1 (1-8)**	**Satisfactory 2 (9-17)**	**Good 3 (18-25)**	**Excellent 4 (26-35)**
Content (35)	Content was not clear the student was not using the generic structure.	Content was clear only one of generic structure are fulfilled, a few detail.	Content was clear, some generic structure are fulfilled, only some details.	Content was very clear, all of generic structure are fulfilled, so the listener can easy to understand .

Aspect	Need Improvement 1 (1-7)	Satisfactory 2 (8-15)	Good 3 (16-23)	Excellent 4 (24-30)
Language features (30)	Weak language control, a few use past tense, and did not use conjunction and adverb of time.	Adequate language control, a few use past tense, conjunction and adverb of time.	Good language control, some use past tense, conjunction and adverb of time.	Excellent control of language features, use of past tense, use of conjunction and use adverb of time.
Aspect	**Need Improvement 1 (1-5)**	**Satisfactory 2 (6-10)**	**Good 3 (11-15)**	**Excellent 4 (16-20)**
Fluency (20)	Speak was very slow, stumbling, nervous, and uncertain with response, except for short or memorized expressions difficult for a listener to understand	Speak was slow and often hesitant and irregular sentences may be left uncompleted, but the student was able to continue	Speak was mostly smooth, but with some hesitation and unevenness caused primarily by rephrasing and groping for words	Speak was effortless and smooth with speed that comes close to that of a native speaker
Aspect	**Need Improvement 1 (1-3)**	**Satisfactory 2 (4-7)**	**Good 3 (8-11)**	**Excellent 4 (12-15)**
Vocabulary (15)	Student had inadequate vocabulary to express his/her idea properly.	Student was able to use a few vocabularies, but was lacking, and can't expand his or her idea.	Student was able to use a lot of vocabulary and he or she can expand his/her idea.	Student was able to use rich precise vocabulary in a good manner, and he or she can expand his/her idea.

There are limitations of total score here to categories students' skill, the maximum total score of speaking was 100 and the minimum total score was 1-59. Passing score was score that have to be reached by

students in order to pass the test. The score can be categorized in the table below:

Table 3.3 Standard performance:

Excellent	80-100
Very Good	70-79
Good	60-69
Need Improvement	1-59

D. Validity and reliability testing

The researcher used test as instrument to collect data. A good test must fulfill and consider standardized of test itself. Measuring a good test, there were some aspects to make a good test, those are: validity and reliability.

1. Validity

Validity is the extent to which inference made from assessment result is appropriate, meaningful, and useful in terms of the purpose of assessment (Gronlund in Brown, 2004: 22). Thus, a test should test what the tester wants to test.

Content validity was used in this research. This kind of validity depends on careful analysis of the language being tested and the particular treatment activity. The test should be constructed to contain

representative sample because the relevancy of the objective and the content of the test items showed the content validity of the test. The instruction of test can be seen in Appendix 2. From the score achieved by the students in tryout test showed that the students performed their ability as being measured. Therefore, it can be concluded that speaking test administrated in tryout has met the criteria of content validity.

To measure students' speaking ability, the test used was speaking test. It means that test is said to have content validity if it is represented the content of universe. Ary et al (2006:226) stated that to have a content validity, the instruments are representative of some defined universe or domain of content. It meant that the items of the test must really test the domain that was speaking skill.

In order to judge whether or not the test has content validity, we need a specification of the skills or structure being tested. A comparison of test specification and test content is the basis for judgment for content validity. The researcher made this test based on the course objectives in the syllabus of second years. Therefore, this test is valid in term of content validity.

A test is said have face validity if it measures what is supposed to the measure. Face validity is hardly scientific concepts, yet it is very important a text which does not have face validity may not be accepted by test-takers, teachers, education authorities or employers (Isnawati 2011:30). There are some considerations used in this study to have a good test based on the validity:

1. The instructions given to the students were clear.
2. In this test the students were asked to retell the story. If activity used syllabus and their level.
3. The time allocation must be clearly. The researcher give invited about 5 minutes each of students to tell the story.

From the students score in pretest and posttest showed that the students performed their ability as it is measured. Therefore it can be concluded that speaking test administrated in pretest and posttest has met the criteria of face validity.

2. Reliability Testing

Ary (2002:250) states that reliability is concerned with the effect of such random errors of measurement on the consistency of scores. Actually, the ideal test should be both reliable and valid. In this research, the researcher also used SPSS 16.0 for windows to know the reliability of test instruments.

To measure the reliability of the instrument, the researchers conduct before pretest and posttest. The researchers used inter-rater reliability where the two scorers did the scoring and two sets of scores gotten from the two scores and they were calculated to get the correlation coefficient. The two scores were the English teacher of eight grade students was rater 1 and the researcher itself was rater 2. The researchers took two scores from rater 2, one score from the scores of tryout pretest and one score from tryout from posttest. The tryout test was done before pretest on Feb 9th 2016 and before posttest on Feb 18th 2016.

According to Riduan (2004:118), the criteria of reliability instrument can be divided into 5 classes as follows:

a. If the *alpha cronbach* score 0.00-0.20: less reliable
b. If the *alpha cronbach* score 0.21-0.40: rather reliable
c. If the *alpha cronbach* score 0.41-0.60: enough reliable
d. If the *alpha cronbach* score 0.61-0.80: reliable
e. If the *alpha cronbach* score 0.81-1.00: less reliable

The researcher used Alpha Cronbach Reliability Coefficient in SPSS 16.0 to analyze the data. The computation can be seen in Appendix 3. The result showed that *alpha* is 0.957 and *r* table was 0.349. It can be concluded that the instrument was very reliable and be used to collect the data.

E. Normality Testing

Normality tests are used to determine whether a data set is well-modeled by a normal distribution or not, or to compute how likely an underlying random variable is to be normally distributed. Normality test is intended to show that the sample data come from a normally distributed population.

To know the normality, the researcher used *kolmogorov-smirnove test* with SPSS. The hypothesis for testing normality are:

a. H_0 : Data is in normal distribution
b. H_a : Data is not in normal distribution

Critic area is H_0 was rejected when the significance value was lower than 0.05 ($\alpha = 5\%$). The analysis was as follow:

Testing data from pre-test and post-tets score using SPSS 16.00

Table 3.8 One Sample Kolmogorov-Smirnov Test

One-Sample Kolmogorov-Smirnov Test

		PRE	POST
N		32	32
Normal Parameters[a]	Mean	62.09	71.88
	Std. Deviation	5.567	6.374
Most Extreme Differences	Absolute	.217	.203
	Positive	.164	.203
	Negative	-.217	-.157
Kolmogorov-Smirnov Z		1.229	1.150
Asymp. Sig. (2-tailed)		.098	.142

a. Test distribution is Normal.

Based on the output from SPSS above it was known that the significance value from pre-test was 0.98 and from the post-test was 0.142. Both value from pre-test and post-test were bigger than 0.05. The significant value on pre-test was 0.98 and it was bigger that 0.05 (0.98 > 0.05). It means that H_0 was accepted and H_a was rejected and the data was in normal distribution. Then, for post-test score the value of sig/p was 0.142 and it was bigger than 0.05 (0.142 > 0.05). It means that H_0 was accepted and H_a was rejected and the data was normal distribution. So, it can be interpreted that both of data (pre-test and Post-test score) are normal distribution.

F. Data analysis

In this research, the researcher used quantitative data analysis. The quantitative data of this research were analyzed by using statistical method. This technique was used to find significant different on the students' speaking ability before and after being taught by using Snowball Throwing Technique. To know the signifficant diference of the speaking ability between before and

after being taught by using Snowball Throwing Technique paired sample Ttest of SPSS 16.0 for windows was used.