CHAPTER III
RESEARCH METHODS

This chapter prominently labors the point about the methods of research used in the study, covering research design, subject of the research, variables, data collecting technique, steps in collecting the data, instruments of the research, criteria of a good test, and data analysis drawn as follows:

3.1 Research Design

This research was a quantitative study. It was aimed at finding out whether there was a significant difference of speaking achievement between students taught through silent viewing and those taught through sound-only technique and determining in what elements of speaking most of the students differ when they are assigned to have oral performance. During the research, a true experimental research (two group-pretest-posttest design) was applied. There were two classes as the subject of the research. One class was the experimental class 1 while the other one was the experimental class 2.

The design can be illustrated as follows:

G1 (random) T1 X1 T2

G2 (random) T1 X2 T2

In which:

GI : **Experimental Class 1**

G2 : **Experimental class 2**

X1 : **Treatment (**using silent viewing technique)

X2 : **Treatment** (using sound only technique)

T1 : Pretests (before treatment)

T2 : Posttests (after treatment)

(Setiyadi, 2006: 143)

The pretest was administered to find out the students' initial speaking ability, represented by their speaking achievement, before treatments. Subsequently, the students' in the experimental class 1 was given three treatments by using silent viewing technique, while the students in the experimental class 2 accepted the same number of treatments by using sound-only technique. Eventually, a posttest was administered to find out the students' speaking achievement after treatments.

3.2 Population and Sample

The population of this research was grade X of SMAN 1 Natar in the school year 2011/2012 consisting of nine classes. Each class consists of 37-40 students. Two classes, selected by using purposive sampling, were taken as the subject of this research, in which one class became the experimental class 1, taught by using silent viewing, while the other one became the experimental class 2, taught through sound only technique. They were purposively sampled as a consideration from the English teacher in the school that they have relatively similar ability (Setiyadi, 2006: 44).

After the pretests for both experimental class 1 and 2 were conducted, the results were compared. It is necessary to measure whether the two classes had a relatively equal basic ability in speaking. Thus, the result of the homogeneity of the scores between the two classes was carried out by using T-test. The hypothesis for the homogeneity of the variance test is:

H_0 : There is no significant difference

H_1 : There is a significant difference

$H_{0,}$ indicates that there is no significant difference. This means that the students both in the experimental class 1 and 2 have a relatively equal initial ability in speaking. Meanwhile, the H_1 is proved if the students of the two classes do not have similar initial speaking ability. Perhaps, the students of the experimental class 1 have better initial speaking ability before treatments or vice versa.

3.3 Variables

Based on the problems of the research, the variables can be defined as follows:

1. The dependent variable is the students' speaking achievement.

2. The independent variable is the treatments proposed in the research (silent viewing and sound-only technique).

Dependent variable, in this case the students' speaking achievement, was the main variable in the research. This variable was measured after all treatments in the research were done. It was the product of interaction among all variables. On the other hand, independent variable is the one in a research which takes role as the cause or functions to affect the dependent variable.

3.4 Data Collecting Technique

In the process of data collecting, speaking tests were used as the instrument. There were two types of tests; pretest and posttest which are elaborated as follows:

1. Pretest

   Pretest was conducted before treatments were given. It was carried out in order to know how far the students had mastered speaking before treatments. The data collected in this research were used to measure whether there is a significant difference or not between the students' speaking achievement, in the experimental class 1 and 2.

2. Posttest

   Posttest was conducted after the employment of treatments. It was carried out in order to know the increase of the students' speaking achievement after being given treatments. The form of the test was the same as that in the pretest. It was conducted in 80 minutes, 10 minutes for each group since the oral performance was done in group which consists of five students respectively.

3.5 Procedures of Collecting Data

Related to data collecting process, there were some procedures applied for this research, they were:

1. Determining the subject of the research

   Two out of nine classes of grade X of SMA N I Natar were selected as the subjects of the research. One class became the experimental class 1 and

one class was the experimental class 2, determined by using simple

probability sampling using coin.

2. Preparing materials for pretest

In this research, there were two pretests conducted both in the

experimental class 1 and experimental class 2. The topic was about

narrative story. The materials were taken from students' handbook based

on the School-based Curriculum.

3. Conducting the pretest

Pretest was conducted in order to measure students' basic speaking ability

before being given treatments. It was in form of oral test of monologue,

carried out in 80 minutes.

4. Giving treatments

There were three times of treatments conducted in this research. Each

treatment was held for 90 minutes, consisting of procedures of teaching

speaking through silent viewing and sound-only technique.

5. Conducting the posttest

The posttests were conducted both in the experimental class 1 and 2. They

were held in order to measure the increase of the students' speaking

achievement. The same as the pretest, posttest was conducted in 80

minutes, in form of oral test of monologue presentation.

6. Analyzing the test result (pretest and posttest)

After pretest and posttest were conducted, the data of these two tests were

analyzed by using independent groups T-test. The mean of the pretests and

posttests both in the experimental and control class were compared. It was

done to know whether there was a significant difference between the two techniques in contributing increase toward the students' speaking achievement.

3.6 Criteria for Evaluating Students' Speaking Performance

The form of the test was subjective, since there was no exact and fixed answer for the speaking performance. Therefore, to minimize any subjectivity as much as possible, inter rater was occupied to assess the students' performance which was also documented in form of recording. The raters were the researcher herself and an English teacher at the school, Dian Noviyanti, who is the English teacher in the experimental class 1 and 2, graduated from English Department of Teacher Training and Education Faculty of Lampung University. She has been teaching those classes since 2009.

Meanwhile, the students' utterances were recorded since it helped the raters to evaluate their performance more precisely and objectively. Furthermore, the speaking test was measured based on two principles, reliability and validity, as explained below:

1. Reliability

Reliability much deals with how far the consistence as well as the accuracy of the scores given related to the students' speaking performance. The concept of reliability stems from the idea that no measurement is perfect; even if one goes to the same scale there will always be differences in the weight which become the fact that measuring instrument is not perfect. Since this was a subjective test, inter

rater reliability was occupied to make sure and verify that both the scoring

between raters and that of the main rater herself (the researcher) is reliable or not.

The statistical formula for calculating inter-rater reliability is as follows:

$$R= 1- \frac{6.(\sum d^2)}{N.(N^2-1)}$$

In which:

R   = Reliability

N   = Number of Students

D   = the difference of rank correlation

1-6   = Constant number

After the coefficient between raters was found, the coefficient of reliability was

analyzed based on the standard of reliability below:

a. a very low reliability   : ranges from 0.00 to 0.19

b. a low reliability   : ranges from 0.20 to 0.39

c. an average reliability   : ranges from 0.40 to 0.59

d. a high reliability   : ranges from 0.60 to 0.79

e. a very high reliability   : range from 0.80 to 0.100

Slameto (1998: 147).

Statistical computation of SPSS 15 was used to measure the inter-rater reliability

in this research. The results gained were reported as follows:

a. Inter-rater reliability of Pretest in Experimental Class 1.

$$R= 1- \frac{6.(\sum d^2)}{N.(N^2-1)}$$

$$R= 1-\frac{6.(1500)}{35.(35^2-1)}$$

$$R= 1-\frac{9000}{35.(1225-1)}$$

$$R= 1-\frac{9000}{35.(1224)}$$

$$R= 1-\frac{9000}{428\ 40}$$

$$R= 1-0.21$$

$$R= 0.79$$

This shows that both two raters have a high inter-rater reliability (0.79)

b. Inter-rater reliability of Posttest in Experimental Class 1.

$$R= 1-\frac{6.(\sum d^2)}{N.(N^2-1)}$$

$$R= 1-\frac{6.(912.5)}{35.(35^2-1)}$$

$$R= 1-\frac{547\ 5}{35.(1225-1)}$$

$$R= 1-\frac{547\ 5}{35.(1224)}$$

$$R= 1-\frac{547\ 5}{428\ 40}$$

$$R= 1-0.13$$

$$R= 0.87$$

The calculation above indicates that the two raters have a very high reliability

(0.87).

c. Inter-rater reliability of Pretest in Experimental Class 2.

$$R = 1 - \frac{6.(\sum d^2)}{N.(N^2-1)}$$

$$R = 1 - \frac{6.(612.5)}{35.(35^2-1)}$$

$$R = 1 - \frac{367\ 5}{35.(1225-1)}$$

$$R = 1 - \frac{367\ 5}{35.(1224)}$$

$$R = 1 - \frac{367\ 5}{428\ 40}$$

$$R = 1 - 0.09$$

$$R = 0.91$$

It implies that the two raters have a very high inter-rater reliability (0.91).

d. Inter-rater reliability of Posttest in Experimental Class 2

$$R = 1 - \frac{6.(\sum d^2)}{N.(N^2-1)}$$

$$R = 1 - \frac{6.(113.5)}{35.(35^2-1)}$$

$$R = 1 - \frac{668\ 1}{35.(1225-1)}$$

$$R = 1 - \frac{668\ 1}{35.(1224)}$$

$$R = 1 - \frac{668\ 1}{428\ 40}$$

$$R = 1 - 0.16$$

$$R = 0.84$$

This indicates that both of the raters have a very high inter-rater reliability (0.84).

Those calculations of coefficient between the first and the second rater depict that both raters have a close divergence varying from high to very high inter-rater reliability. Besides, it also points out that the first rater's scoring is representative and reliable to be proceeded forward.

2. Validity

In essence, any test should portray a proper reflection of what has been taught as well as what knowledge or ability students are intended to know and master. This is validity which obviously meets this rationale. It makes reference to a state that the test measures what is intended to measure. It means that it relates directly to the purpose of the test. Furthermore, this research focused on two types of validity, content validity and construct validity. Content validity can be best examined by the table of specification which matches the syllabus used by the teacher. From the table of specification, it is shown that the content validity is achieved.

Meanwhile, construct validity concerns with whether or not the test is actually in line with the theory of what it means to the language that is being measured. It would be examined whether or not the test actually reflects what it means to know a language (Shoamy, 1985: 74). The indicator of five speaking elements (pronunciation, grammar, vocabulary, fluency, and comprehension) is used in this research. It implied that the test measured those intended aspects based on the indicator, meaning that the construct validity has been fulfilled.

This research occupied a comparison of the test to the table of specification to know whether or not the test was a good reflection of what has been taught and the knowledge that the teacher wanted the students to know. A table of specification is an instrument that helps the test constructor plans the test. Below is the comprehensive depiction of it:

Table 1.1 Table of Specification of Speaking

Aspects of speaking elements	Percentage	Aspects of narrative speaking performance	Percentage
1. Pronunciation	15%	1. Opening: self and group introduction, orientation	20%
2. Vocabulary	15%	2. Evaluation	20%
3. Grammar	15%	3.Complication	20%
4. Fluency	25%	4. Resolution	20%
5. Comprehension	30%	5. Closing: moral message and group performance closing	20%

3.7 Speaking Test

In this study, the speaking tests were conducted in 80 minutes respectively. In the test, the students both in the experimental class 1 and 2 had to perform their speaking test in front of the class in form of monologue. It was a group performance, consisting of five students respectively. Yet, the scoring of the speaking performance was done individually. The record of students' performance

was used in order to score the students' speaking test accurately. Then, it was assessed together with the rater-the English teacher of the class.

In the students' speaking scoring, the record was used as one of considerations for scoring assessment since it helped both the researcher and the rater to evaluate the students' speaking performance more objectively. Furthermore, it could also be replayed as is needed by both the researcher and the rater to view it back in order to get the precise and representative result of scoring. The rating sheet modified from Harris (1974:84) was used as a guide. Based on the oral rating sheet, there are five aspects to be tested and evaluated, namely pronunciation, grammar, vocabulary, fluency, and comprehension which were completely illustrated as follows:

    a.   Pronunciation

5 the students' pronunciation is the same as the native speaker.

4 It is easy to be understood though sometimes unclear.

3 the pronunciation can be understood by the listener even though there is a
   difficulty.

2 Pronunciation is difficult to understand and there is often a repetition.

1 Pronunciation cannot be understood by the listener.

    b.   Grammar

5 grammatically correct sentence seen form the pattern.

4 Bit errors in sentence patters.

3 Usage of patterns so that they can blame the wrong sentence meaning.

2 at few mistakes, with no pattern of failure.

1 Incorrect grammar

    c.   Vocabulary

5 Use of vocabulary is appropriate to the material being discussed.

4 The use of inappropriate words

3 Using wrong words, conversation is rather limited because of inadequate

vocabulary.

2 A very limited vocabulary makes comprehension difficult.

1 Vocabulary limitation is very extreme to make a virtual conversation impossible.

    d. Fluency

5 Fluently and easily as is done by native speakers

4 Speed seems to be rather strongly influenced by the language problem.

3 The speed and smoothness rather strongly influenced by the language problem.

2 There is a lot of repetition.

1 Speech is so halting and fragmentary to take conversation impossibly virtual.

    e. Comprehension

5 It is easy to understand by the listener.

4 It is easy to understand even though sometimes the repetition may be necessary.

3 Can be understood even though a bit difficult.

2 It cannot be understood.

1 It cannot be said to be understandable even it is a very simple utterance.

The scores of each point are multiplied by four. Hence, the highest score is 100.

Here is the identification of the scores:

If the students get   5, then 5x4= 20

                         4, then 4x4= 16

                         3, then 3x4= 12

                         2, then 2x4= 8

                         1, then 1x4= 4

For example:

A student gets 4 in pronunciation, 3 in vocabulary, 3 in fluency, 4 in comprehension, and 3 in grammar. Therefore, his total score will be:

Pronunciation   4x4= 16

Vocabulary   3x4= 12

Fluency   3x4= 12

Comprehension   4x4= 16

Grammar   3x4= 12

Total   = 68

It means he gets 68 for his speaking performance.

3.8 Data Analysis

In order to see whether there was an increase of students' speaking achievement and to determine in what speaking element students most differ when they are assigned to have oral performance, the students' scores were examined by using some steps as follows:

a. scoring the pretest and posttest,

b. tabulating the scores of the students' speaking result using rating scale,

c. calculating and comparing the means of students' speaking scores both in the experimental class 1 and 2,

d. calculating and comparing the means of posttests scores in each element of speaking in experimental class 1 to those in experimental class 2,

e. drawing conclusion from the tabulated result of the pretest and posttest.

The data sample of score of pretest (T1) and posttest (T2) can be seen in the table below:

Table 1.2 Interraters' Scoring Scale

Students' name	Aspects of Speaking										Total
	Pronun-ciation		Vocabu-lary		Fluency		Compre-hension		Grammar		
	R1	R2	R1	R2	R1	R2	R1	R2	R1	R2	
A	0.3	0.45	0.45	0.45	0.3	0.45	0.5	0.75	0.9	0.9	
B	0.3	0.3	0.3	0.3	0.3	0.3	0.5	0.75	0.6	0.6	
$\sum N =$											

Table of score inter-rater reliability of pretest and posttest is:

No.	Students' Code	Rater 1		Rater 2	
		Pretest	Posttest	Pretest	Posttest
1	A	43	49	46	60
2	B	20	40	20	45

3	C	52	46	60	52

Since the data were from different groups of students, they were analyzed by using statistical computation i.e. independent group T-Test of SPSS version 15. This was carried out to test whether or not the significant increase of the students' gain was there as well as to determine in what element of speaking students most differed when they were assigned to have oral performance.

3.9 Data Treatment

In this research, the procedures to treat the data were as follows:

a. Normality of the Test

The normality of the test is used to measure whether or not the data in the experimental and control class are normally distributed. The hypothesis for the normality test is as follows:

H_0 : The data are normally distributed.

H_1 : The data are not normally distributed.

In this study, H_0 is accepted if it signs $> \alpha$, with the level of significance 0.05. To analyze the data, SPSS version 15 was used. Parametric Analysis was occupied, by using Independent T-Test.

b. Homogeneity Test

This test was used to determine whether the data fulfilled the criteria of the equality of variance or not. T-test was occupied to analyze the data. The hypothesis for the homogeneity of variance is as follows:

H_0 : There is no significant difference in the level of ability (equal)

H_1 : There is a siginificant difference in the level of ability (not equal)

In this case, the criteria for the hypothesis are:

H_0 is accepted if it signs > α, with the level of significance α= 0.05.

H_1 is accepted if it signs < α, with the level of significance α= 0.05.

c. Hypothesis Testing

The hypothesis testing was used to prove whether the hypothesis proposed in this research was accepted or not. SPSS (Independent T-test) was utilized. The hypothesis was analyzed at the level of significance 0.05. This was operated to draw the conclusion and is approved if it signs < α.

The hypotheses proposed were:

   H_0 : There is no significant difference of students' speaking achievement between those taught through *silent viewing* and *sound only technique.*

   H_1 : There is no significant difference of students' speaking achievement between those taught through *silent viewing* and *sound only technique.*

Independent T-test in SPSS version 15 was used to investigate the significant difference of students' speaking achievement with the level of significance 0.05. The criteria are:

   If the t-ratio is higher than t-table   : H_1 is accepted

   If the t-ratio is lower than t-table   : H_0 is accepted

3. 10 Research Schedule

This research was planned to be conducted on dates enlisted below:

Table 3. 12. 1 Research Schedule

No.	Date	Agenda
1.	January 9th 2012	1. Pre-research, introductory meeting
2.	January 10th 2012	2. Conducting pretest for experimental class 1 3. Conducting pretest for experimental class 2
3.	January 14th 2012	4. Conducting the first treatment of sound-only technique for experimental class 2
4.	January 16th 2012	5. Conducting the first treatment of silent viewing technique for experimental class 1
5.	January 17th 2012	6. Conducting the second treatment of silent viewing technique for experimental class 1 7. Conducting the second treatment of silent viewing technique for experimental class 2
6.	January 21th 2012	8. Conducting the third treatment of sound only technique for experimental class 2
7.	January 23th 2012	9. Conducting the third treatment of silent viewing technique for experimental class 1
8.	January 24th 2012	10. Conducting posttest for experimental class 1 11. Conducting posttest for experimental class 2