III. RESEARCH METHOD

3.1 The Research Design

To conduct this research, the researcher used *Pretest Posttest Control Group Design*. This design belongs to true experimental designs. True experimental designs have three basic characteristics: (1) a control group is present, (2) The sample are randomly selected and assigned to the groups, and (3) a pretest is administered to capture the initial differences between the groups (Hatch and Farhady 1982:22).

The researcher used this design because she wanted to give special treatment to the experimental class using retelling story in teaching reading comprehension and the control class which does not. There were two classes of this experimental study; one is experimental class which got treatment through retelling fairy tale and another as a control class which got treatment through translation.

The pretest administered first before the treatment. It was intended to measure the students' basic ability of both in order to ensure their entry point.

Control class was needed for comparison purposes because it lets the writer interpret her findings more confidently. Both of them got the same materials.

The research design can be represented as follow:

G1 (random) T1 X T2

G2 (random) T1 O T2

(Hatch and Farhady 1982: 22)

G1 : Experimental class
G2 : Control class
T1 : Pre- test
T2 : Post-test
X : Treatment (teaching reading through retelling a fairy tale)
O : Regular teaching (teaching reading through translation)

The result of post test of both classes were compared to prove whether or not there is difference of students' reading comprehension achievement between experimental class and control class in order to see whether teaching reading through retelling a fairy tale is effective or not

3.2 Population and Sample

3.2.1 Population

The population of this research was the 1st grade students of MA Ma'arif 4 Kalirejo Kab. Lampung Tengah period of 2011/2012. There are 2 classes in the 1st grade of MA Ma'arif 4 Kalireio. Each class consists of 30 students.

3.2.2 Sample

The samples of this research were two classes of the first year students in 2011/2012 academic year. The writer chose the first grade students because she faced that they could not comprehend in reading English reading fairy text. This material (Fairy tale/ narrative texts) is including in the first grade of Senior High School. One was the experimental class and the other was the control class. As the

sample, the class was selected randomly by using lottery, since the first year students of MA Ma'arif 4 Kalirejo Lampung Tengah was not stratified class.

There was no priority class. It was applied based on that consideration that every student in the population had the same chance to be chosen and in order to avoid the subjectivity in the research. Next, to determine which class is as the experimental class and as the control class, the researcher used a coin by flipping it.

3.3 Data Collecting Technique

To collect the data the writer will use the following techniques:

1. The Pretest

The pre-test will be administered before the treatment applied. It is done in order to know how far the competence of students in reading comprehension before the treatment. It is also needed to know whether both experimental class and control class are equal or not in the terms of their reading comprehension achievement. The test will be multiple choices.

2. The Posttest

Post-test will be administered after the treatment applied in order to find out whether there is difference of students' reading comprehension between those who are taught through retelling fairy tale and those taught through ordinary technique (translation). The post test will be done after six meetings of the treatments. The result of the post test of two classes will be compared in order to know whether teaching reading through retelling fairy tale is effective or not.

3.4 Research Procedures

The research will be conducted during normal class hour. The writer will follow the following procedures:

1. Determining the Problem

   The first step of this research was determining the problem. The writer determined what kind of problems appear in the classes.

2. Determining and selecting the samples

   The population of this research is the first year students of MA Ma'arif 4 Kalirejo, Lampung Tengah. The samples of this research were two classes which was chosen randomly.

3. Conducting try out

   There were 40 items. The try out was administered in 60 minutes. The aim of this try out was to know the quality of test, which would be used as instrument of the research.

4. Preparing the pretest and giving the pretest.
5. Preparing the treatment material and presenting the treatment material by implementing the technique.
6. Preparing the posttest material and giving the posttest.
7. Analyzing the data. The data of pretest and posttest were put into a score table and it was used to see the significant increase of students' score in reading comprehension.

3.5 Scoring System

In scoring the result of students' work, the researcher used Percentage Correct (Lyman, 1971:95). The percentage correct score is used in reporting the result of classroom achievement tests.

The researcher will calculate the average of the pre-test and post test by using this formula:

$$X_{\%c} = 100\frac{R}{T}$$

(Lyman, 1971: 95)

Where:

$X_{\%c}$ = percentage of correct score
R = number of right answers
T = total number of items on test.

3.6 Try Out

The tests are said to have good quality if it had a good validity, reliability, and level of difficulty and discrimination power.

3.6.1 Validity of the Test

Validity is the extent to which a test does the job desired of it; the evidence may either empirical or logical (Lyman, 1971:196). A test can be said valid if the test measures the object to be measured and suitable with the criteria (Hatch and Farhady, 1982: 251). According to Hatch and Farhady (1982: 251), there are four basic types of validity: face validity, content validity, construct validity and empirical or criterion-related validity.

The validity of the test is considered in this researcher. The writer will take content and constructs validity for this research. It will be considered that the test should be valid and in line with reading theory and material.

a. Content Validity

Content Validity is the extent to which a test measures a representative sample of the subject matter content, the focus of content validity is adequacy of the sample and simply on the appearance of the test (Hatch and Farhady, 1982:251). The writer uses content validity because she wants to know whether or not the content of the test is sufficiently representative and comprehensive for the test to be a valid measure of what it is supposed to measure. To know whether the test is good reflection of what will be taught and of the knowledge which the teacher wants the students to know, the researcher compares this test with table of specification. If the table represents the material that the researcher wants to test, then it is valid from that point of view. A table of specification is an instrument that helps the test constructor plans the test.

Table 1. Table of specification

NO	Objective	Number of items	Presentage
1	Identifying main idea	1.,5.,11.,18.,26.,32.,35.	17,5%
2	Inference	8.,9.,14.,15.,27.,38.	15%
3	Reference	2.,16.,25.,29.,34.,39.	15%
4	Specific Information	3.,4.,6.,13.,17.,19.,21.,22., 23.,30.,31.,33.,36.,40.	35%
5	Vocabulary	7.,10.,12.,20.,24.,28.,37.	17,5%
Total			100%

b. Construct Validity

Construct validity is concerned with whether the test is actually in line with the theory of what reading comprehension means (Hatch and Farhady, 1982:251).

To know the test is true reflection of the theory in reading comprehension, the researcher will examine whether the test questions actually reflected the means of reading comprehension or not.

3.6.2 Reliability of the Test

Reliability refers to the extent to which the text is consistent in its score, and gives us an indication of how accurate the test score are (Hatch and Farhady, 1982: 244).

To measure the coefficient of the reliability between the first and the second half, Pearson Product Moment is used, which is formulated as follows:

$$rxy = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{n(\sum x^2) - (\sum x)^2 n(\sum y^2) - (\sum x)^2}}$$

(Hatch and Farhady, 1982: 222)

Where,

n = number of students

r = coefficient reliability between first and second half

x = total number of first half

y = total number of second half

x^2 = square of x

y^2 = square of y

$\sum x$ = total score of first half items

$\sum y$ = total score of second half items

Then the researcher will use "Spearman Brown's Pharophecy formula" (Hatch and Farhady, 1982: 246) to determine the reliability of the test as follow:

$$Rk = \frac{2rl}{1+rl}$$

(Hatch and Farhady, 1982: 286)

Where:

Rk : the reliability of the test

rl : the reliability of half test

The criteria of reliability are:

0.90- 1.00 = high

0.50- 0.89 = moderate

0.0- 0.49 = low

3.6.3 Level of Difficulty

Level of difficulty of an item simply shows how easy or difficult the particular item proved in the test (Heaton, 1975: 182). Level of difficulty is generally expressed as the fraction (or percentage) of the students who answered the item correctly.

To see the level of difficulty, the writer uses the following formula:

$$LD = \frac{R}{N}$$

(Heaton, 1975: 182)

Where:

LD : level of difficulty

R : the number of the students who answer correctly

N : the total number of the students

The criteria are:

< 0.30 = difficult

0.30- 0.70 = average

> 0.70 = easy

3.6.4 Discrimination Power

The discrimination power (DP) is the proportion of the high group students getting the items correct minus the proportion of the low-level students who getting the items correct.

To see the discrimination power, the writer uses the following formula:

$$DP = \frac{correctUpper - correctLower}{{}^{1}/_{2}N}$$

(Heaton, 1975: 182)

Where,

DP = Discrimination Power

U = Number of upper group who answer correctly

L = Number of lower group who answer correctly

N = Total number of the students.

The criteria are:

DP: 0.00-0.19 = Poor

DP: 0.20-0.39 = Satisfactory

DP: 0.40-0.69 = Good

DP: 0.70-1.00 = Excellent

DP: - (negative)= Bad items, should be omitted

3.7 Data Analysis

To analyze the data, some following techniques will be used:

3.7.1 Normality Testing

Normality test is an idealized model which can be used to dealing with natural behavior (Hatch and Farhady 1982:64). This test is used to measure whether the

data in two classes are normally distributed or not. The normal distribution has three distinct properties that allow the researcher to make inferences about the population in general and our sample of the population in particular. That is:

1. The mean, median, and mode in a normal distribution are all the same.
2. The first property results in the second characteristic- the shape of the normal distribution is bell shaped and symmetric.
3. The normal distribution does not have a zero score; the tails never meet the straight line.

(Hatch and Farhady 1982:64)

The data are tested by One-sample Kolmogorov-Smirnov Formula (SPSS 15).

The criteria of normal distribution are:

The hypothesis is accepted if the result of the normality test is higher than 0.05 (sign > □). In this case, the researcher used level of significance of 0.05.

3.7.2 Homogeneity Test

This test is used to know whether the data of the pretest from the experimental class and from the control class are homogeneous or not. The data will be tested by Independent Sample Test (SPSS 15).

The criteria for the homogeneity of pre test are:

H_1: There is no significant difference in the level of ability (equal)
H_0: There is a significant difference in the level of ability (not equal).

The criteria for the hypothesis is: H_1 is accepted if the result of Homogeneity test of pre test is higher than 0.05 (Sign > □).

3.7.3 Hypothesis Testing

The researcher will test the hypothesis to prove whether the hypothesis proposed by the researcher is accepted or not. The writer used Independent Group T-test to know the level of significance of the treatment effect. The hypotheses is analyzed at the significant level of 0.05 in which the hypothesis is approved if sign <□. It means that the probability of error in the hypothesis is only about 5%.

The formulation is:

$$t_{obs} = \frac{\overline{X}_e - \overline{X}_c}{S_{(\overline{X}_e - \overline{X}_c)}}$$

With:

$$S_{(\overline{X}_e - \overline{X}_c)} = \sqrt{\left(\frac{S_e}{\sqrt{n_1}}\right)^2 + \left(\frac{S_c}{\sqrt{n_2}}\right)^2}$$

$\overline{X}_e$	: Mean from the difference pre-test and post-test of experimental class 1
$\overline{X}_c$	: Mean from the difference pre-test and post-test of experimental class 2
$S_{(\overline{X}_e - \overline{X}_c)}$	: Standard error of differences between means
n	: Subjects on sample

(Hatch and Farhady, 1982:111)

The criteria are:

If the t-observe is higher than t-table : H_0 is accepted
If the t-observe is lower than t-table : H_1 is accepted

H_0 = There is no difference of students' reading comprehension between students who are taught through retelling story and students who are taught through translating technique by using fairy tale.

H_1 = There is difference of students' reading comprehension between students who are taught through retelling story and students who are taught through translating technique by using fairy tale.

(Hatch and Farhady 1982:120)