

III. RESEARCH METHOD

This chapter discusses the design of this research and how to collect the data from the samples. The researcher describes the data collecting technique, the procedures and the instrument of this research. The researcher also gives the scoring system and how to analyze the data

3.1 Research Design

This research was quantitative research carried out by using *One Group Pretest-Posttest Design*. In this design, pretest and posttest was administered to find out whether Video Movie can be used to increase students' listening comprehension achievement. The researcher chose one class as the experimental class by applying probability sampling using lottery. In this research, the students were given the pre test before treatment and post test after the treatment. The treatments were given three times by teaching listening through vidoe movie. The design of this research is represented as follow:

T1 X T2

Where:

T1 : Pre-test

X : Trématent

T2 : Post-test

(Setiyadi, 2006:133)

3.2 Population and Sample

1. Population

The population of this research was the first grade students of SMAN 1 Natar Lampung Selatan. There were nine classes in this school of the first year students: X1, X2, X3, X4, X5, X6, X7, X8, X9. Each class consists of 35-40 students. The total number of the population were 342 students, consisting of 178 female students and 164 male students. In this research, the researcher chose the first year students to be investigated. The researcher choose X2 as the experimental class and it consists of 36 students.

2. Sample

From the population above one class were taken as the sample of this research, as the experimental class, that was given the treatment (teaching listening through video movie). The sample was selected by using random sampling technique. The experimental class (X2) was chosen based on the consideration that every student has low score in English. The sample of experimental class:

Experimental class	Female	Male	Total
	22	14	36

3.3 Data Collecting Technique

The data of the research were the students' listening comprehension achievement before the treatment (pre-test) and after the treatment (post-test) to

see whether or not there is significant increase of student's listening comprehension achievement after being taught by using video movie. To collect the data, the researcher uses the following ways:

➤ **Test**

The test was divided into pretest and posttest, it was elaborated in the following section:

a. Pre-test

The pretest was conducted before the treatment. It was used to know how far the students' listening comprehension achievement before treatment was given. The researcher gave pretest to the experimental class. The pretest was objective test in the form of multiple choices. The numbers of the items in the test were 30 items and each item had four options of the answer. One is the correct answer and the rests are the distracters. The total score was 100 points, so if the students answer the whole questions correctly they got 100 point.

b. Posttest

The posttest was conducted after the researcher had given the treatments. It was used to know how far the students have mastered listening comprehension after being taught through video movie. Similar to the pretest, in the posttest the researcher gave test of multiple choices. The questions were the same as it was used in pretest. But, the order of the questions and the distracters were changed from those in the pretest. The posttest consists of 30 items with four options of answer for each (A, B, C, D). One is the correct

answer and the rest were the distracters. This posttest had the same difficulty as the pretest.

3.4 Procedures of the Research

The procedures of the research were as follow:

1. Determining the population and sample of the research

To determine the population and sample of the research, the researcher chose two classes from nine classes at the first year students in SMAN 1 Natar Lampung Selatan which is divided in X1 until X9 and each of which consists of 35-38 students. Those classes one class was choosen as the try-out class and one class as the experimental class. In the try-out class, the researcher gave try-out test in the form of multiple choices. Then in the experimental class, the researcher gave pre-test, treatment, and post-test. The researcher took the sample randomly by using lottery, where all the classes have the same opportunities to be the sample.

2. Administering listening test (try out) to know the quality of the test

Try out test was carried out to know the quality of the test in order to take the data. The try out was conducted in the first meeting. The class was used for the try out test was the class which is not included in the experimental class. Try out was conducted to measure the reliability of pretest and posttest. It was administered for 40 items in 90 minutes. The aim of try out was to know the quality of the test which used as the instrument of the research, and determine which item should be revised for the pretest and posttest. This research used

the result of the try out test to measure the level of difficulty and discrimination power, to find out the validity and reliability.

3. Preparing the materials which will be taught

The researcher took the material based on youtube video (Fairy Tales Story Collection retrieved on 21 June 2012 from <http://www.appuseries.com/fairytales.php>). The video movie used in this research is flash video/FLV in form of narrative spoken. The topics of video movie were about Fairy Tales with the title are: Rapunzel, The Frog Prince, and The Proud Teacher. The reason why the researcher choose this material was because narrative text is one of the monologue texts that should be mastered by students in this levels.

4. Administering the pre-test

In this research, there was one pre-test that proper to the first grade students of SMAN 1 Natar. The researcher administered the pre-test in order to find out the student's basic ability before treatment. In this term the researcher asked the students to do multiple choice test which consist of 30 items and the student should choose the correct answer from four option, e.g. a, b, c, and d. The pre-test took 90 minutes.

5. Giving treatment

After giving the pretest to the students, the treatment was conducted for three times. The class was given treatments with three different lesson plans. Those three lesson plans consist of three different topics. It was necessary in order to increase the students' ability in comprehending the English text. During the process of treatment, the researcher used video movie as a media for teaching listening, then the English teacher observed the students' activity.

6. Administering the post-test

Post-test was done after giving the treatment. Post-test was very important in this research because the researcher need to know whether there was any significant increase of students' listening comprehension achivement after being taught through video movie. The post-test took 90 minutes which consist of 30 items and divided in four option a, b, c and d.

7. Analyzing the data (pre-test and post-test)

Both of the pretest and posttest results of the class treated by using repeated measures t-test (Repeated Measures t-test of SPSS (statistical package for social science) version 16.0 for windows). It would test in order to find out whether there was any significant increase of student's listening comprehension achievement after being taught by video movie.

8. Concluding the results

After analyzing the results of both pretest and posttest, the conclusion would be explained based on the result.

9. Reporting the results

In reporting the result, the data were arranged systematically based on the pretest and posttest to see whether there is an increase on the student's achievement in listening comprehension activity significantly or not.

3.5 Instrument Used for Collecting the Data

3.5.1 Listening Test

The instrument of this research is a set of listening comprehension tests that were used for try out, pretest and posttest. These tests were in the form of multiple

choices. There were 40 items of multiple choices (objective test) in the try-out. Then, there were 30 items of multiple choices in pretest and posttest. The multiple choices tests were used since its marking is rapid, simple, and most importantly reliable, not subjective or influenced by the marker's judgment (Heaton, 1975).

3.5.2 Observation

To know the process of teaching listening by using video movie the researcher used observation sheet that would be used during teaching learning process.

Researcher made lesson plan and taught the students based on the lesson plan. In collecting the data the researcher was assisted by English teacher. When the researcher acts as the teacher, the English teacher observed student's activities in the classroom.

3.6 Criteria of a Good Test

To know whether the test is good or not, some criteria should be considered. The criteria of a good test are: Validity (content validity and construct validity), Reliability, level of difficulty and discrimination power.

3.6.1 Validity

Validity refers to the extent to which the test measures what is intended to measure. It means that it relates directly to the purpose of the test (Shohamy, 1985:74). There are four types of validity, namely face validity, content validity, construct validity, and empirical validity or criterion-related validity. To measure

whether the test has a good validity, the researcher was used content validity and construct validity. Face validity only concerns with the layout of the test, while the criterion-related validity is concerned with measuring the success in the future, as in replacement test, (Hatch and Farhady, 1982:251). So these two validities are considered to be less needed. Therefore, To measure whether the test has a good validity, this research used content and construct validity. the two types of validity which will be used in this research, as follows:

- **Content validity** means that the test is good reflection of what has been taught and the knowledge which the teacher wants the students to know (Shohamy 1985: 74). It means that the items of the test should present the material being discussed. Then, the test is determined according to the materials that have been taught to the students. In other words, the test is based on the materials in the English curriculum, so that it can be said that the test has content validity since the test is good representation of material studied in the classroom.

- **Construct validity** was concerned with whether the test is actually in line with the theory of what it means to know the language (Shohamy, 1985; 74). It means that the test items should really test the students or the test items should really measure the students' ability in listening comprehension.

Since listening and reading have the same purpose in comprehending the message while listening comprehends the oral message, reading comprehend the written messages. Therefore, the researcher used the same aspects of listening comprehension namely determining the main idea of the text, finding specific information of the text, inference, similar or opposite meaning, referencing and vocabulary in the test instrument. The table specification of the instrument test could be seen on the table below:

Table 3.1 Table Specification of Listening Comprehension Test

Aspects	Objectives	Description	Number of items	Percentage
Macro Skill	Identifying main idea	- What is the main idea of the story? - What does the story tell us about?	2	5%
	Identifying specific information	- Who are involve in the story? - What is the characteristic of wicked witch?	22	55%
	Reference	- They ate, played, and slept together slowly. The underlined word “they” in the sentence refers to...	6	15%
	Vocabulary	- Rapunzel grew up to be a <u>beautiful</u> girl. The word “beautiful” has the opposite meaning to...	4	10%
	Inference	- What is the moral value of the story you heard? - What is the coclusion of the story you heard?	6	15%
Total			40	100%

3.6.2 Reliability

Reliability refers to extend to which test is consistent in its score and gives us an indication of how accurate the score test are. The concept of reliability stems from the ideas that no measurement is perfect even if we go to the same scale there would be differences in our weight which are a result of the fact that measuring instrument is not perfect. To measure the coefficient of the reliability

between odd and even group, this research will use the Pearson Product Moment formula as follows:

$$r_1 = \frac{\sum xy}{\sqrt{\sum x^2 \sum y^2}}$$

Where:

r_1 : coefficient of reliability between odd and even numbers items

x : odd number

y : even number

x^2 : total score of odd number items

y^2 : total score of even number items

xy : total number of odd and even number

(Lado : 1961 in Hughes, 1991: 32)

Then this research used “Spearman Brown’s prophecy formula” to know the coefficient correlation of whole items. The formula is as follows:

$$r_k = \frac{2 \cdot r_1}{1 + r_1}$$

Where:

r_k : the reliability of the test

r_1 : the reliability of half the test

(Hatch and Farhady, 1982:246)

The criteria of reliability are:

0.80- 1.00 : very high

0.50 – 0.79 : moderate

0.00 – 0.49 : low

(Hatch and Farhady, 1985:247)

3.6.3 Level of Difficulty

Difficulty level related to how easy or difficult the item is from point of view of the students who take the test. This was important since test items, which are too easy, tell us nothing about differences is discarded.

To see the level of difficulty, this research used the following formula:

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

Where: LD= Level of difficulty

R = the number of students who answer correctly

N = the number of then students

(Arikunto, 1992: 23)

The criteria are:

LD < 0.30 = difficult

LD 0.30 – 0.70 = average

LD 0.70 – 1.00 = easy

3.6.4 Discrimination Power

The discrimination power (DP) refers to the extent to which the item differentiates between high and low level students on the test. A good item according to this criterion is one which good students do well on and bad students fail. To know the discrimination power of the test, the researcher used the following formula:

$$DP = \frac{\text{correct Upper} - \text{correct Lower}}{1/2N}$$

Where:

DP : discrimination power

U : the proportion of upper group students

L : the proportion of lower group students

N : total number of students

The criteria are:

D: 0.00-0.20 : poor items

D: 0.21-0.40 : Satisfactory items

D: 0.41-0.70 : Good items

D: 0.71-1.00 : Excellent items

D: - (Negative) : bad items (should be omitted)

(Heaton, 1975:180)

1. If the value is positive discrimination a large number of more knowledgeable students then poor students got the item in correct. If the value is zero, no discrimination.
2. If the value is negative, it means that more low-students than high level students got the item correct.
3. In general, the higher the discrimination index, the better. In classroom situation most items should be higher than 0.20 indexes.

(Shohamy, 1985:81)

3.6.5 Scoring System

In scoring the students result of the test, this research used Arikunto`s formula. The ideal higher score is 100. The score of pretest and post tests are calculated by

using formula as follows: $S = \frac{R}{N} 100$

Where:

S : the score of the test

R : the total of the right answers

N : the total items

(Arikunto, 1997:212)

3.7 Data Analysis

3.7.1 Test

The data were analyzed in order to see whether the students' listening comprehension achievement is increase or not. The researcher examined the students' scores using the following steps:

1. Scoring the pretest and posttest.
2. Tabulating the results of the test and calculating the scores of the pretest and posttest.
3. Drawing conclusion from the tabulated result of the pretest and posttest administered, that is by statistically analyzing the data using statistical computerization i.e. *Repeated Measure t-test of Statistical Package for Social Science (SPSS) version 16.0 for windows* to test whether the increase of students' gain is significant or not, in which the significance is determine by $p < 0.05$. It is uses as the data come from the two samples. (Hatch and Farhady, 1982:111)

3.7.2 Process

Learning process is the students' activities during the listening comprehension test of narrative text through the use of video movie. The indicator of the student's activities is that if 80% students are active during learning process. In interpreting all data available by selecting material, the researcher selected the data in order to keep them relevant with the research question about the process of teaching listening comprehension. The process of teaching listening comprehension through video movie can be showed in appendix 14.

3.8 Hypothesis Testing

The researcher used the hypothesis testing to prove whether the hypothesis proposed in this research was accepted or rejected.

The hypothesis of the research was: “There was a significant increase of students’ listening comprehension achievement after being taught through video movie”

The hypothesis was statistically analyzed using repeated measures t-test that was used to draw the conclusion in significant level of 0.05 ($P < 0.05$) in which the hypothesis is approve if $\text{Sig} < \alpha$.

H_0 : There is no significant increase of students’ listening comprehension achievement after being taught through video movie

H_1 : There is a significant increase of students’ listening comprehension achievement after being taught through video movie.

3.9 Result of a Tryout test

Before conducting the pre-test and post-test, a try out test was carried out. This test was administered in order to determine the quality of the test as instrument of the research. To know the reliability of research instruments, the try out test was administered at the first meeting on August 28th, 2012 approximately at 07.30- 09.00 a.m. at class X3. It required 2 x 45 minutes to administer the try out test. There were 40 items of multiple choices in try out test. In this activity the researcher did not use video movie, the students only hear sound of the story without looking the media. After analyzing the data, the writer got that 30 items were good, 10 items were bad and should be dropped . The result of the reliability found through this research was 0.98 (**appendix 4**).