

III. RESEARCH METHOD

This chapter discusses the research design that was used by the researcher, population and sample of the research, how the researcher collected the data, research procedure, validity and reliability of the test, level of difficulty, discrimination power, scoring system, data analysis, data treatment and hypothesis testing.

3.1 Research Design

This quantitative research has two variables which are *teaching listening through English songs* as the independent variable and *listening ability* as the dependent one. The researcher used one group pre test post test design. The research used two classes, one as the try-out class and other class as the experimental class. The students were given pre test, the treatments and post test. The treatments of teaching listening comprehension through English songs were implemented three times. One group pre test post test design is represented as follows:

T1 X T2

T1 : Pre test

X : Treatments, is the application of using classic pop song. The researcher will conduct three treatments.

T2 : post test

(Setiyadi, 2006:143)

3.2 Population and Sample

The population of this research was the students of the eleventh grade of SMAN 12 Bandar Lampung at even semester academic year of 2012/2013. The sample was chosen by the English teacher of that school. The researcher chose this way to determine the sample because he did not want to disturb teacher's teaching learning program that has been being run at that time. The researcher took two classes of seven science study program class as the try-out and experimental class. Each class consisted of 30 students. There are twenty female students and ten male students in the try-out class. Then there are twelve female students and eighteen male students in the experimental class.

3.3 Data Collecting Technique

3.3.1. Administrating the Try-Out

The Try-out was administrated in eleventh class two as the try-out class in the first meeting before Pre-test of teaching listening comprehension through English songs implemented in the experimental class, as the consideration of the pre-test items.

3.3.2. Administrating the Pre-test

The pre-test was administrated in eleventh science 2 class as the experimental class before giving the treatment of teaching listening

comprehension through English songs to see the students' basic listening comprehension.

3.3.3. Administrating the Post-test

The post-test was administrated in eleventh science 2 class after the treatment of teaching listening comprehension through English songs implemented by using the same topic and three times repetition of one question like in the pre-test. The result of the post-test was compared with the result of the pre-test to analyze the improvement of the students' listening comprehension through English songs.

3.3.4. Conducting the Interview

The interview was conducted in eleventh science 2 class, in which the students' answers were classified and generalized as the resource. The interview was conducted to find out the problems the students face in learning macro skills of listening comprehension through English songs.

3.4 Research Procedure

The procedures of this research as follows:

1. Determining the population and chose the sample.
2. Administering try out

The try was used to know the quality of the test to take the data and to determine which item should be omitted for the pre-test and post-test. The questions were multiple choices. In this try out the researcher administered

the try out test in order to measure the level of difficult (LD) and discrimination power (DP), to find the reliability and the validity of the test. The test item was objective test. It was given in class out of the experimental. It was administered 25 items in 45 minutes.

3. Administering the pre test

Pre test was needed to find out the students' ability in listening. The students heard a song and then they tried to answer several questions connecting to the song. The questions represented the five aspects that measure in this research, that are determining main idea, finding the detail, reference, inference, and vocabulary. The questions were multiple choices.

4. Conducting the treatments

At the first time, the researcher who acts as the teacher told the student about the material. Then the students heard a classic pop song as the introduction to the material. Next, the researcher gave the students two kinds of tests they were; multiple choices and missing lyrics that was answered based on the song that they heard. The song was played three times. And the researcher conducted three treatments.

5. Administering post test

The post test was conducted after the treatments. The students answered several questions based on the classing song that was heard by them. The questions were multiple choices and missing lyrics. It was administered 25 items in 45 minutes.

6. Analyzing the result of pre test and post test

The researcher gave score for the students' answer in pre test and post test by seeing the three aspects that was measured in this research, these are identifying the main idea, identifying specific information and identifying inference. In this step the researcher used inter-rater to check the students' result in order to avoid the subjectivity in giving correction.

7. Comparing the result of pre test and post test

After analyzing the result of pretest and post test, the researcher compared those results to answer whether there is significant improvement made by the students after being taught about listening comprehension through classic songs.

8. Interview

The researcher gave some questions to the student considering whether the classic pop song was good or not for students' listening ability.

3.5 Validity and Reliability of the Tests

The test was considered as the valid one if the test measures the object to be measured and it was suitable with the criteria (Hatch and Farhady, 1982:250). According to the Hatch and Farhady (1982:281), there are two basic types of validity, such as content validity and construct validity. Therefore, to measure whether the test has a good validity, this research used content and construct validity. The validity of the instrument was presented as follows:

Content Validity

Content validity concerned with whether the test is sufficiently representative and comprehensive for the test. According to Hatch and Farhady (1982:251), since content validity is the extent to which a test measures a representative sample of the subject matter, the focus of content validity is adequacy of the sample of the appearance of the test. Therefore, since the test instrument is conducted to get the data of the students' listening comprehension achievement, the content validity of the test items were conducted by including listening materials which were arranged based on the materials already given and it was suitable with the curriculum. Thus, if the measuring instrument has represented all the ideas that connected with the materials that were measured, that measuring instrument has fulfilled the aspect of content validity.

Construct Validity

Construct validity concerned with whether the test was actually in line with the theory of what it means to know the language that was being measured. To achieve the construct validity, the test was adopted from the indicators of Basic Competence (kompetensi dasar) that have been formulated before in syllabus based on KTSP 2006 curriculum of senior high school.

According to Hughes (1991:134), in macro skill, to understand what someone says, a listener has to involve with listening for specific information, obtaining gist of what is being heard or the listener should get the general idea of the information, following instructions or directions. Therefore, in this

research, the researcher focused more on improving the students' macro skills of listening comprehension rather than micro skills of listening.

Besides, the test instrument was conducted based on the aspects of macro skills of listening comprehension above. It was tried-out first then was selected from 35 items. Then, the test instrument used to analyze the improvement of the students' listening comprehension through songs by measuring these aspects, such as *identifying the main idea, identifying specific information and identifying inference* (WH questions).

In order to fulfill the criteria of construct validity, the table of specification of listening aspect which was modified from the theory proposed by Hughes (1991:134), the test instrument can be seen below:

Table 3.5.1 Summary of specification of listening test

No	Types of Listening Comprehension	Items Numbers	Percentage (%)
1	Identifying the main idea	12,13,16,23,26,30,33	20%
2	Identifying specific information	2,3,7,8,10,15,18,19,20,22,24,28,29,32,34	43%
3	Identifying inference	1,4,5,6,9,11,14,17,21,25,27,31,35	37%
Total		35 items	100%

In order to measure the content and construct validity, *inter-rater* analysis was used to make the listening test instrument more valid. Thus, three English teachers of SMAN 12 Bandar Lampung were the raters. If the percentage of one item is >50%, it means that the item test will be taken.

Reliability

Reliability of the test can be defined as the extent to which a test produces consistent result when administrated under similar conditions (Hatch and Farhady, 1982:243). Split-half technique used to estimate the reliability of the test and to measure the coefficient of the reliability between odd and even group, *Pearson Product Moment formula* was used as follows:

$$r_l = \frac{\sum xy}{\sqrt{[\sum x^2][\sum y^2]}}$$

r_l : 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 numbers.

(Lado, 1961 in Hughes, 1991:32).

The criteria of reliability are:

- 0.80 – 1.00: high.
- 0.50 – 0.79: moderate.
- 0.00 – 0.49: low.

(Hatch and Farhady, 1985:247).

To know the coefficient correlation of whole items, *Spearman Brown's prophecy formula* was used. The formula was followed:

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

r_k : The reliability of the test.

r_1 : The reliability of the half test.

(Hatch and Farhady, 1982:246).

3.6 Level of Difficulty

Level of difficulty was used to look for the level of test that was given to the students.

For seeing the level of difficulty, the researcher used the following formula:

$$LD = \frac{U + L}{N}$$

In which:

LD : Level of Difficulty

U : The proportion of the upper group students

L : The proportion of the lower students

N : The total number of students who following the test

The criteria are:

< 0.30 : difficult

0.30-0.70 : average

> 0.70 : easy (Shohamy, 1985:79)

3.7 Discrimination Power

Discrimination power was the extent to which the items discriminate between the testers, separating the more able testers from the less able.

For determining discrimination power, the researcher used the following formula:

$$DP = \frac{U - L}{\frac{1}{2} N}$$

In which:

DP Discrimination power

U : The proportion of the upper group students

L : The proportion of the lower students

N : The total number of students

The criteria are:

DP: 0.70-1.00 : Excellent

DP: 0.40-0.69 = Good

DP: 0.20-0.39 = Satisfactory

DP: 0.00-0.19 = Poor

DP: - (Negative) = Bad item, should be omitted (Heaton, 1975:182)

3.8 Scoring System

The scoring system that was used in this research was driving the right answer by total items times 100. In scoring the students' result of the pretest and post-test, the formula by Arikunto (1993, 240):

$$S = \frac{R}{N} \times 100$$

Notes:

S : score of the best

R : the right answers

N : the total item

3.9 Data analysis

In order to see whether there was an effect in students' listening by using classic pop songs, the researcher analyzed the data by:

1. Scoring the pretest and posttest.
2. Tabulating the results of the tests and calculating the mean of the pretest and the posttest.
3. Drawing conclusion from the tabulated results of the pretest and posttest administered, that is by statistically analyzing the data using statistical computerization i. e. repeated measures T-test of SPSS (Statistical Package for Social Science) version 17.0 for windows to see whether or not the difference between pretest and posttest are significant, in which the

significance was determined by $p < 0.05$. It was used as the data come from the same sample or known as paired data

(Hatch and Farhady, 1982:114).

The researcher also used qualitative data to support this researcher in order to see whether or not there were some problems or difficulties faced by the students in the aspects of macro skill in listening during teaching learning process through classic pop songs. The researcher knew whether there were some differences between the previous technique taught by the teacher and the technique which was used by the researcher through classic pop songs. The researcher also intended to know whether or not the students enjoy in learning by using classic pop songs.

Interview conducted in experimental class in the form of open and formal questions (the questions must be in the form of explanation or description rather than “yes” or “no” answers, to avoid the students from being reluctant to answer the questions given). The interview was used to know the students’ opinions and respond toward the listening technique was used by the researcher. This interview was held after conducting the treatment and post test. The researcher asked the students’ opinion about effectiveness of using English song in teaching listening.

3.10 Data Treatment

1. Normality Test

The writer used normality test to know whether the data was distributed normally or not. The hypothesis was formulated as followed:

H_0 : The data is not distributed normally

H_1 : The data is distributed normally

In this script, the criterion for the hypothesis was that:

H is accepted if $\text{sign} > \alpha$. in this case, the writer used the level of significance 0.05. It proved that the H_0 was accepted and all the data was distributed normally.

3.11 Hypothesis Testing

In testing the hypothesis that the teaching learning through English songs would improve the students' listening comprehension significantly, *Repeated Measure T-Test* was used. The hypothesis was also statistically tested by using statistical computerization (SPSS 17), in which the significance is determined by $p < 0.05$. Therefore, the hypothesis which can be cited is as follows:

H1: There is significant improvement of students' listening comprehension ability after being taught through English songs.

Besides that, the interview was conducted. The interview in experimental class was in the form of open and formal questions (the questions must be in the form of explanation or description rather than "yes" or "no" answers, to avoid the students from being reluctant to answer the questions given); to analyze its qualitative data in order to find out the problems the students face in learning macro skills of listening comprehension through English songs.