

III. RESEARCH METHODS

This chapter presents research design, population and sample, data collecting techniques, research procedures, try-out, scoring system, procedures of collecting data, and hypothesis testing.

3.1. Research Design

In this research, the researcher used a quantitative research based on the experimental design. A quantitative research was used to measure how far the dictation technique influences the students in improving their listening ability. The researcher used *one group pre-test and post-test design* which took two classes as the experimental class and the try-out class.

Before administering the pre-test, the researcher firstly administered a try-out in the try-out class to measure the quality of the test which was used in taking the data. After that, the pre-test was administered to measure the students' listening ability before the treatments. After the treatments were finished, the post-test was administered to know the progress of students' listening ability. The researcher used one class as the sample of the research where the researcher conducted the pre-test, the treatments, and the post-test to see whether there is better gain of students' listening ability after being taught through dictation.

In this research, the researcher was interested in investigating whether there is a significant improvement of students' listening ability after being taught through dictation technique. The design was presented as follows:

T1 X T2

Note:

- T1 is the pre-test.
- X is the treatment.
- T2 is the post-test.

(Hatch and Farhady, 1982: 24)

3.2. Population and Sample

3.2.1. Population

The population of this research was the first year students of SMA Negeri 1 Tanjungbintang in academic year of 2012/2013. There are seven classes of the first year students. Each class consists of 34-38 students.

3.2.2. Sample

The researcher selected the sample by using random sampling technique through lottery drawing because the whole classes had the same rights to be chosen as the sample of the research. The researcher took two classes; the first was X-3 as the try-out class and the other was X-1 as the experimental class.

3.3. Data Collecting Techniques

To carry out the research, the researcher used the following techniques to collect the data, as follows:

3.3.1. Pre-test

The pre-test was administered to the experimental class before applying the treatment. This test was done in order to know how far the students' competence in listening before the treatment. This test consisted of two parts; writing simple sentence and multiple choice with three options. The test was conducted within 60 minutes.

3.3.2. Post-test

The post-test was administered to the experimental class in the last program of this research after applying treatment. This test was done in order to know the improvement of students' listening ability. The test items were the same as in the pre-test but in random order.

3.4. Research Procedures

The procedures of this research were as follows:

1. Determining the problems

In determining the problem, the researcher did pre-observation at SMA Negeri 1 Tanjungbintang to investigate the students' problems of listening in the class.

2. Determining the population and sample

The population of this research was the first year students of SMA Negeri 1 Tanjungbintang. There are seven classes which consist of 34-38 students per each class. The researcher took two classes as the experimental class (X-1) and

another one as the try-out class (X-3). The researcher selected randomly in determining the experimental class and try-out class through lottery.

3. Finding and selecting the materials

This research applied the materials which were taken from the English handbook based on the teaching and learning syllabus and the School Based Curriculum 2006 (KTSP).

4. Administering the try-out test

After designing the research instruments, the try-out test was carried out. The try-out was administered before giving the pre-test. It was expected to find out whether the test items used in the research were good or not. It refers to validity, reliability, level of difficulty, and discrimination power. This test consisted of two part; part one was writing simple sentence (listen, repeat, and write) and part two was multiple choice with three options (A, B, and C) and one of them was as the correct answer. The number of the test items is 50; 15 items for part one and 35 items for part two. The score of each correct answer was 2 points. Therefore, if one participant answered all the items correctly, she/he got 100 points. This test was administered for about 80 minutes.

After analyzing the data, the researcher got that 25 items were good and 25 items were bad and should be dropped. To know the reliability of the test, the researcher used *Spearman Brown Formula*. The computation showed that the reliability of the test was 0.84 (see appendix 3). From the computation of level of difficulty, the researcher got that there were 6 items difficult items, 14 easy items, 30 items satisfactory items. From the computation of discrimination

power (see appendix 4), the researcher got that there were 3 bad items, 22 items were poor, and 25 items were satisfactory and good. Eventually, the items that were administered for both of the pre-test and post-test were 25 items.

5. Administering the pre-test

The pre-test was aimed for gaining the data of the students' listening ability. The test was administered in the experimental class, class X-1. It was administered as the first test for all testiest before applying the treatment. The total of test items were 25 for 60 minutes. The score of each correct answer was 4 points. If one participant answered all the items correctly, she/he got 100 points.

6. Implementing the treatment using dictation technique

The treatment in the form of dictation technique was implemented in the experimental class. The dictation technique was given for three times. The time allocation was 2 x 40 minutes per meeting.

7. Administering the post-test

The post-test was conducted in the last program of this research after applying the dictation technique to the experimental class. The post-test was given to X-1. The test was the same as in the pre-test but in random order. The purpose of the post-test was to find out how far the improvement of students' listening achievement after the treatment.

8. Analyzing the test result

The data of the research (pre-test and post-test) were statistically analyzed by using SPSS 16.0. The researcher used *Paired Samples T-test* to analyze the data in order to find out whether dictation technique is able to improve the students' listening ability. The gain from the score of pre-test and post-test of the experimental class was then compared to see the difference.

9. Making the report about the findings.

3.5. Try-out

The try-out test was administered to gain appropriate instruments for this research. This test was administered in another class beyond the experimental class. There are four terms used to analyzed the try-out score. They are validity of the test, reliability of the test, level of difficulty, and discrimination power of the test. If the test has met the four criteria, it indicates that the test can be used as the base of arranging pre-test and post-test. These are some elements tested as follows:

3.5.1. Validity of the Test

To determine the validity of the test, the researcher only analyzed the test from content validity to measure whether the test has good validity or not. A test can be said valid if it measures the object to be measured and suitable for the criteria (Hatch and Farhady, 1982: 251).

Content Validity means that the test is a good reflection of the material that will be tested. It means that the items of the test should represent the material that has

been discussed. To get the content validity, the test is adopted from students' handbook which is based on the curriculum. It means that the test has a good content validity since the test is a good representation of material studied in the classroom by the teacher.

Table 1. Specification of Micro Skill in Pre-test

No	Objective	Number of Items	Percentage
1.	Intonation pattern	2, 6, 9, 14, 18, 25	24%
2.	Recognition of function and structure	1, 8, 15	12%
3.	Cohesive devices	10, 11, 19, 21, 23, 24	24%
4.	Detect sentence constituents	3, 4, 5, 7, 20, 22	24%
5.	Recognizing discourse markers	12, 13, 16, 17	16%
	Total	25	100%

Table 2. Specification of Micro Skill in Post-test

No	Objective	Number of Items	Percentage
1.	Intonation pattern	1, 7, 10, 11, 14, 16	24%
2.	Recognition of function and structure	3, 5, 13	12%
3.	Cohesive devices	9, 17, 18, 20, 22, 24	24%
4.	Detect sentence constituents	2, 4, 6, 8, 15, 25	24%
5.	Recognizing discourse markers	12, 19, 21, 23	16%
	Total	25	100%

3.5.2. Reliability of the Test

Besides validity, reliability is also an important characteristic of a good test to identify whether the test as an instrument for collecting data is stable or not. As Shohamy (1985: 70) states that reliability refers to the extent to which the test is consistent in its score, and it gives an indication of how accurate the test score. To find out the reliability of the test, the researcher used split-half technique which requires her to split the test into two similar parts, odd (x) and even (y) of the test

items. To measure the coefficient of the reliability between odd and even group, the researcher used *Spearman Brown Formula* which was formulated below:

$$R_{xy} = \frac{N \sum XY - (\sum X)(\sum Y)}{\sqrt{N \sum X^2 - (\sum X)^2} (N \sum Y^2 - (\sum Y)^2)}$$

Note:

R : Coefficient of reliability between odd and even numbers

N : Number of the students

X : Square of x

Y : Square of y

$\sum X$: Total score of odd number

$\sum Y$: Total score of even number

The criteria of reliability are:

0.80-1.00 = Very high

0.60-0.79 = High

0.40-0.59 = Average

0.20-0.39 = Low

0.00-0.19 = Very low

(Hatch and Farhady, 1982: 198)

In this research, the result of reliability of the try-out test was 0.84 (see appendix 3). It could be inferred that the test had very high level of reliability, in the range 0.80-1.00. It indicated that this instrument would produce consistent result when it was administered under similar condition, to the same participants, and in different time (Hatch and Farhady, 1982: 286). So, it can be concluded that the test was reliable.

3.5.3. Level of Difficulty

Level of difficulty is used to know whether the test items are easy or difficult. To see level of difficulty, it was calculated by the following formula:

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

Note:

LD = Level of difficulty

R = The number of students who answer correctly

N = The total number of students who take the test

The criteria are:

0.0 up to 0.29 = Difficult

0.30 up to 0.69 = Satisfactory

0.70 up to 1.00 = Easy

(Heaton, 1986: 178)

Based on the criteria above, there were 14 easy items in the try-out test (10, 13, 16, 19, 21, 22, 27, 30, 35, 36, 38, 40, 45, and 49). There were 6 difficult items (6, 7, 12, 15, 31, and 43). And, there were 30 satisfactory items. (see appendix 4)

3.5.4. Discrimination Power

Discrimination power 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 researcher used the following formula:

$$D = \frac{\text{correct } U - \text{correct } L}{1 / 2N}$$

Note:

D = Discrimination index

correct U = The number of upper group who answer correctly

correct L = The number of lower group who answer correctly

N = The total number of the students who take the test

The criteria are:

0.00-0.20 = Poor items

0.21-0.40 = Satisfactory items

0.41-0.70 = Good items

0.71-1.00 = Excellent items

- (negative) = Bad items, should be omitted

(Heaton, 1975:180)

Based on the criteria above, there were 25 items in the try-out test which did not fulfill the standard of discrimination power, since those items had discrimination index under 0.20 which meant that the items had poor discrimination power. By looking discrimination power and level of difficulty, the total items that were administered were 25 items (1, 3, 4, 5, 8, 9, 11, 13, 17, 18, 21, 23, 26, 28, 29, 32, 34, 36, 39, 40, 44, 45, 46, 47, and 48). Those items had discrimination power more than 0.21 with the criteria satisfactory to excellent items. (see appendix 4)

3.6. Scoring System

In scoring the students' test, the researcher used percentage correct by Lyman (1971). The percentage correct score was used in getting the result of the students' achievement test. The researcher calculated the pre-test and the post-test by using this formula:

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

Note:

$X_{\%c}$ = Percentage of correct scores

R = Number of right answers

T = Total number of items on test

(Lyman, 1971: 95)

3.7. Procedures of Collecting Data

The procedures of collecting data were used to know the conclusion of students' listening comprehension achievement. The data analysis consisted of quantitative analysis which was based on the students' score. The data of the research were analyzed by using *Paired Samples T-test*. The researcher analyzed the data by doing this steps, as follows:

1. Scoring the pre-test and post-test.
2. Tabulating the result of the test and finding the mean of the pre-test and post-test by using this formula:

$$\bar{X} = \frac{\sum X}{N}$$

Note:

\bar{X} = Mean

$\sum X$ = Total score of the students

N = Number of students

(Hatch and Farhady, 1982: 5)

3. Drawing conclusion by comparing the means of the pre-test and the post-test.
4. Analyzing the data used t-test. It was important to find out whether the data from experimental class were random and normally distributed or not.

- Random Test

In this research, random test was used to know whether the data in the experimental class are random or not. The researcher used SPSS 16.0 for Windows with level of significant 0.05. The researcher used mean as cut point. The hypothesis was formulated as follows:

H_0 : The data are random.

H_a : The data are not random.

The data are determined random if it gets the criterion; the H_0 is accepted if the result of random test is higher than 0.05 (Sig.> α). From the result of the test, it showed that the data were random. (see appendix 10)

- Normality Test

The researcher used normality test to measure whether the data were distributed normally or not. The hypothesis were formulated as follows:

H_0 : The data are distributed normally.

H_a : The data are not distributed normally.

The H_0 is accepted if the result of normality test is higher than 0.05 (Sig.> α). In this case, the researcher used level of significant 0.05. The result of the normality test showed that the data were distributed normally. (see appendix 11)

3.8. Hypothesis Testing

Hypothesis testing was used to prove whether hypothesis that proposed by the researcher was accepted or not by using t-test. The researcher used SPSS (Statistical Package for Social Science) version 16.0. The researcher used *Paired Samples T-test* in order to know the significance of the treatments' effect by comparing the mean of pre-test and post-test. The hypothesis was analyzed at significant level of 0.05 in which the hypothesis was approved if sig.< α .

The criteria are:

H_0 (null hypothesis) is accepted if t-ratio is lower than t-table (t-ratio<t-table). It means that there is no significant improvement of students' listening ability after being taught using dictation technique.

H_a (alternative hypothesis) is accepted if t-ratio is higher than t-table ($t\text{-ratio} > t\text{-table}$). It means that there is a significant improvement of students' listening ability after being taught using dictation technique.

The hypothesis testing (see appendix 12) showed that t-ratio is higher than t-table ($10.607 > 2.042$). It meant that H_0 was rejected and H_a was accepted. It can be concluded that there is a significant improvement of students' listening ability after being taught through dictation technique at SMA Negeri 1 Tanjungbintang.