

III. RESEARCH METHOD

This chapter discusses about the design of the research, subject of the research, research procedure, analyzing the data, research instrument (validity, and reliability of the instrument, level of difficulty, discrimination power), data analysis, scoring system and hypothesis testing.

3.1 Research Design

This research was intended to find out the most interesting topic from three topics given and to know the students' problem during the teaching learning process using Clustering Technique. Therefore, Quasi Experiment was employed to answer the quantitative data and questionnaire was used to collect the qualitative one. In conducting the research, the researcher used One Way Repeated Measures Anova Design by giving different materials in every treatment and every test. The researcher used two different classes, one class where the students got one time pre-test, three times treatment, one time post-test and one more class for try out.

In the treatments, the students got three different topics. Each topic was evaluated to know the most effective topic for utterance production in teaching vocabulary. The criterion used to determine whether there was an effective topic was

determined by comparing the mean of score in every post –test. The research design is described as follows:

$$T1 \quad X \quad T2$$

Where:

T1 : pretest

T2 : post test

X : treatments

(Hatch and Farhady, 1982: 20)

To see the students' basic vocabulary ability, pre-test was administered before the treatment using Clustering Technique. Then the treatment of Clustering Technique in teaching vocabulary. To analyze how the improvement of their vocabulary achievement through Clustering Technique and to know the effective topic for utterance production, the post-test was administered afterward.

3.2 Subject of the Research

The research was conducted at the first year of SMA N 1 Natar at the second semester of 2012/2013. From nine classes available, the researcher used class X-2 that consisted of 39 students as the experimental class and class X-1 as the try out class by using purposive sampling. Purposive sampling is a sampling technique on purpose. Means researchers determine their own samples to be taken because there are certain consideration. So, samples are not taken randomly, but determined by the researchers.

3.3 Data Collecting Technique

The data of the research was the students' vocabulary achievement in form of multiple choice test before and after the treatments and matching test in every treatments. The instrument of the research was multiple choice tests, where the researcher gave pretest and posttest in order to evaluate, to measure the vocabulary achievement. In collecting the data, this research used the following procedures.

1. Administrating the Pre-test

The pre-test was administered to the students before the treatment implemented. Pre-test was conducted to know the students' basic vocabulary ability before the students being taught by using clustering technique. Meanwhile, before administering pre-test, the researcher explained the material that would be tested. The test was in the multiple-choice form with four options (A, B, C, and D) and the sum of items was 30 items. The time allocation was 50 minutes. The result of the pretest was compared with the post test result to find out their achievement. The scoring system was the load of each correct answer was thirty. Therefore, if one participant answers all the items correctly, she got 100 points ($30/3 \times 10$).

2. Administering the Post-test

The post-test was administered to the students after the treatment. The post test consisted of 30 items in the form of multiple choices with four options and time allocation was 50 minutes. It was done in order to know the students'

vocabulary achievement after having the treatment. The questions or the items in post test were the same as the pretest. However, the researcher changed the questions number and the distracters than those in pretest. This test had the same difficulty as pretest.

3. Questionnaire

The researcher administered questionnaire to collect the data that deals with students opinion, attitudes on the teaching learning process through Clustering technique. The written form instrument was arranged and spread to the students. The students filled the questionnaire. They were provided a set of questions or statements and was asked to answer the questionnaire freely. The items was about their motivation in learning English, response of learning vocabulary through Clustering technique, and difficulties in learning vocabulary through Clustering technique. The researcher also provides the reason for the students to give for in order to get more valid data.

3.4 Research Procedure

This research was conducted based on the following procedures:

1. Determining the subject

The researcher took two classes of the first grade students at SMA Negeri 1 natar. Those classes were chosen by using non probability sampling technique as an experimental class and another one as a try out class.

2. Preparing the try-out test and conducting it

The try out test was administered at the first meeting. The researcher used objective vocabulary test in form of multiple choice test that consisted of 50 multiple choice items in 90 minutes. 20 items of 50 items provided was eliminated based on the result of the try-out-test by considering its difficulty index and its discrimination power. Those multiple choice items was analyzed and revised to find out the good items to be used in pre test and post test. When the items had been analyzed whether it was good or not for the pre test and post test, it was divided into three parts for three pre tests and post tests.

The aim of try out test was to find out the quality of the items of the test used in the research, such as, validity, reliability, level of difficulty, and discrimination power of the test.

3. Selecting the Materials

The researcher chose the appropriate materials based on the syllabus.

There was three meetings by using different materials related to vocabulary comprehension.

4. Conducting Pre-test

The pre test was administered in order to measure the students' mastery of vocabulary before being taught through clustering technique. The pre test was in form of multiple choices with 30 items. The questions and the answers of this test were the same as the try out test. In addition, it was

administered for 50 minutes in the experimental class. So that, the researcher was able to compare the result with the post test result later.

5. Treatment

In this research, the treatment was done in three meetings with 90 minutes in every meeting. The chosen class was taught by using clustering technique. The procedure of implementation of clustering technique in teaching vocabulary predicate as follows:

Pre Activities

While Activities

Post Activities

6. Administering Post-test

Post-test was conducted after the treatment. Post-test was administered in order to know the progress of vocabulary skill after using clustering technique. Based on the design of this research that is one way repeated measures anova, post-test was conducted in three times, after each meeting or after each treatment. The post test was administered to the experimental class. The result of the test was compared with the pre test to find out whether there was different score or increase from the pre test and the post test result. It was in objective test in form of multiple choices that consists of 30 items. The questions and the answers of this test were same with the try out test. The post test was administered in 50 minutes.

7. Analyzing Data

After conducting the final test, the researcher analyzed the data. After collecting the data, the students' worksheet was analyzed subjectively by both researcher and teacher. Then, The researcher analyzed the mean of every test by compared from the two raters based on the test. The mean of pre-test and post-test was used to know the improvement of students' vocabulary skill through Clustering Technique. Meanwhile, the mean of every tests in every treatments were used to know which one is the effective topic for utterance production.

3.5 Instrument of the Research

The instrument for collecting data was test in form of multiple choice test that consists of try out, pre test and post test. Try out test was given to know how the quality of the test which was used as the instrument of the research. Pre test was given in order to know the students' vocabulary before the treatments. Post test was given in order to know the students' vocabulary achievement after the treatments. The total number of the pre test and post test was 30 items with four alternative options (A, B, C, D) for each item. The try out was administered about 90 minutes, and pre test was administered about 50 minutes.

3.6 Analysis of Research Instrument

In order to get the good quality of the test, the researcher used the result of the try out test to measure the validity, reliability, difficulty level, and discrimination power. There are some elements that were tested as follows:

Validity and Reliability

Setiyadi (2006: 29) states that in order to make the research valid and reliable quantitative research focuses on collected data. Therefore, as this research based on quantitative research, some considerations are also going to be taken as follows:

1. Validity of the Test

Validity refers to appropriateness, meaningfulness, and useful of the inferences a researcher makes Fraenkel and Wallen (1993). It means that validity refers to the extent to which an instrument gave us the information that we want. Meanwhile, Setiyadi (2006: 24) mentions that the test should reflect all the areas to be assessed in suitable proportions and represent a balanced sample. Other source says that validity is a matter of relevance; it means that the test measures what is claimed to measure. To measure whether the test has good validity, it has to be analyzed from content and construct validity.

a. Content Validity

Content validity is the extended to which a test measures representative sample of the subject matter contents, the focus of the content validity is adequacy of the sample and simply on the appearance of the test (Hatch and Farhady, 1982:251). It means that the test should be correct and represent the materials that had been taught such as vocabulary of occupation. To get the content validity, the material and the test were composed based on indicators and objectives in syllabus of KTSP curriculum. The materials that were taught based on the students' handbook

for Senior High School. Beside that, the writer also made a table of specification to judge the content validity already good or not.

Table 1. Specification that was used to judge the content validity

Aspect	Spread of Item	Sum
Form	3, 4, 5, 6, 9, 11, 15, 18, 20, 21, 25, 26, 27, 29, 30, 32, 33, 34, 36, 37, 38, 39, 41, 45, 46, 48, 49	27
Meaning	12, 17, 23, 40, 42, 43, 47	7
Use	1, 2, 7, 8, 10, 13, 14, 16, 19, 22, 24, 28, 31, 35, 44, 50	16
Total		50

(Adopted and developed from Nation, 2001: 27)

b. Construct Validity

Construct validity is 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 whether the test is in line with the school curriculum. In this research, the researcher used the vocabulary that was supposed to be comprehended by the first grade students of Senior High School. The materials was about real objects, occupations, and sports.

Table 2. Specification of used to judge the construct validity

Word Classes	Spread of Item	Percentage
Noun:	1, 6, 9, 12, 15, 17, 21, 23, 27, 29, 30, 32, 33, 39, 40, 42, 43, 45, 47, 49. (20 items)	40 %
Verb	2, 7, 8, 10, 13, 14, 16, 19, 22, 24, 28, 31, 35, 44, 50. (15 items)	30 %

Adjective	3, 5, 11, 20, 25, 36, 37, 46, 48. (9 items)	18 %
Adverb	4, 18, 26, 34, 38, 41. (6 items)	12 %

(Taken from Arikunto, 2005:196)

2. Reliability of the Test

Reliability refers to the extent to which the test is consistent in its scores, and it given an indication of how accurate the test score are. Heaton

(1991)states that reliability is a necessary characteristic of any good test:

for it to be valid at all, a test first be reliable as a measuring instrument.

To find the reliability of this test the writer used Split Half Method, which

has two steps, they are :

First, using Pearson Product Moment Correlation, we should find the

coefficient correlation between Odd and Even Number of the items.

$$r_{xy} = \frac{\sum xy}{\sqrt{(\sum x^2)(\sum y^2)}}$$

Where:

r : Coefficient of reliability between odd and even number.

$\sum x$: Total Square of x (total score of odd number).

$\sum y$: Total square of y (total score of even number).

$\sum xy$: Total score of odd and even number items.

$\sum x^2$: total score of odd number items

$\sum y^2$: total score of even number items

2. Second, after we got the coefficient Correlation between Odd and even Number, we continue to put them into the Reliability Formula.

The reliability Formula is below:

$$r_{11} = \frac{2 \times r_{xy}}{1 + r_{xy}}$$

Where:

r_{11} : reliability coefficient

r_{xy} : coefficient between odd and even number

The criteria of reliability

0.00 – 0.20 : very low

0.21 – 0.39 : low.

0.40 – 0.59 : average.

0.60 – 0.79 : high

0.80 – 1.00 : very high (Hatch and Farhady, 1982:246)

3. Difficulty Level.

In determining the level of difficulty of the test, the writer used the following formula:

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

Where :

LD : Level of Difficulty.

U : The number of upper group who answer correctly.

L : the number of lower group who answer correctly

N : The total number of the students.

The criteria are :

< 30 : difficult.

0.30 – 0.70 : average.

> 70 : easy. (Shohamy, 1985: 79)

4. Discrimination Power.

To estimate the discrimination power of the tests, the writer used the following formula:

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

Where :

DP : Discrimination power.

Upper : The number of upper group students who answer correctly.

Lower : The number of lower group students who answer correctly.

N : The total number of the students. (Shohamy, 1985: 82)

The criteria are :

0.00 – 0.20 : poor.

0.21 – 0.40 : satisfactory.

0.41 – 0.70 : good.

0.71 – 1.00 : excellent.

(Negative) = bad items (should be omitted) (Heaton, 1975: 182)

3.7 Data analysis

Data analysis is a process for organizing the data in order to get the explanation form. After collecting the data, the researcher calculated the student's achievement by:

1. Scoring pre test and post test.

In scoring the students' test result, researcher used Arikunto's formula. The ideal higher score is 100. The scores of pretest and posttest were calculated by using formula as follows:

$$X\%c = \frac{100 R}{N}$$

Where :

X% c = percentage of correct score

R = the total of the right answer

N = the total items (Lyman, 1971:95)

2. Tabulating the result of the test and calculating the mean of the pre test and post test.

To compute the average score or mean of the pre test and post test, the writer used a very simple formula as follows:

$$\frac{M}{N} = \sum x$$

In which:

M : (Mean) Average score

$\sum x$: Total student's score

N : Total number of the students.

The average is total student's score divided by total number of the students.

(Hatch and Farhady, 1982:55)

3. Drawing conclusion from the tabulated result of the pretest and post test

administrated. The data was analyzed by using statistical computerization

Repeated Measures t-test of SPSS 17 for Windows i.e. $t = \frac{\bar{x}_1 - \bar{x}_2}{\bar{S}_D}$ to test whether

the difference between pretest and post test is significant or not, in which the

significance is determined by $p < 0.05$ (Hatch and Farhady, 1982: 114). Whereas:

$$\bar{S}_D = \frac{SD}{\sqrt{n}}$$

$$SD = \frac{\sqrt{\sum D^2 - (1/n) - (\sum D)^2}}{n-1}$$

Where

t = hypothesis test

\bar{X}_1 = mean score pre test

\bar{X}_2 = mean score post test

\bar{S}_D = standard error of differences between two means

SD = standard deviation

n = number of students (Hatch and Farhady, 1982:116)

Since, the data was gained from one group and the research was intended to find out whether there is a significant difference of the students' English vocabulary achievement.

3.8 Data Treatment

Repeated Measures t-test for hypothesis testing has three basic assumptions, namely:

The data is interval or ratio

The data is taken from random sample in a population

The data is distributed normally.

(Setiyadi, 2006: 168-169)

Therefore, the researcher employed these following procedures:

1. Random test

Run test is used to make sure whether the data random or not. The researcher uses SPSS version 17.0 to help her. In this case, the researcher uses the mean as the cut point run t-test. The hypothesis for the random test is formulated as follows:

H_0 : The data are random

H_1 : The data are not random

In this research, the criteria for the hypothesis are:

H_0 is accepted if Sig. $> \alpha$. In this case, the researcher uses the level of significance 0.05.

2. Normality test

Normality test is used to know whether the data in pretest and post test are distributed normally. The hypothesis of the normality test is as follows:

H_0 : The distribution of the data is normal

H_1 : The distribution of the data is not normal

In this research, the criteria for the hypothesis are:

H_0 is accepted if significant value exceeds level of significance at 0.05

(Sig. > α). Meanwhile, H_0 is rejected if significant value does not exceed level of significance at 0.05.

3.9. Testing the Hypothesis.

Hypothesis testing is intended to see whether the hypothesis that is proposed in this research is accepted or not. To test the hypothesis, *repeated measures T-test* was conducted at the significant level of 0.05 ($P < 0.05$). The hypotheses are:

H_0 :

- There is no significant increase on students' vocabulary achievement after being taught through Clustering Technique.
- There is no difference among the students' achievement of each topic.

H_1 :

- There is significant difference of the students' vocabulary achievement after being taught through Clustering Technique.
- There is difference among the students' achievement of each topic.

(Hatch and Farhady, 1982: 111)

The criteria are:

1. If the t-value is lower than T-ratio: H_0 is accepted there is no significant increase on students' vocabulary achievement after being taught through Clustering Technique.

If the f-value is lower than F-table: H_0 is accepted there is no difference among the students' achievement of each topic.

2. If the t-value is higher than T-ratio: H_1 is accepted there is significant increase on students' achievement after being taught through Clustering Technique.

If the f-value is higher than F-table: H_0 is accepted there is difference among the students' achievement of each topic.