III. RESEARCH METHODS

This chapter discusses about research design, population and sample, research procedure, data collecting technique, instrument of the research, instrument of the try out, validity and reliability, scoring system, measuring productive vocabulary, data analysis, normality test, hypotheses testing.

3.1 Research Design

This research was quantitative research. In conducting the research, the researcher applied one group pretest-posttest, pre-experimental design. The researcher used one class where the students received pretest before treatments and they received posttest after the treatments. The pretest was to find out the students' preliminary ability and the posttest was used to see how far the increase of the students' vocabulary mastery after the treatments. The treatments were given to the students by using flashcards.

In doing this research, the data was taken by having pretest (TI) and posttest (T2) on the sample. It could be demonstrated as follows:

TI X T2

Where:

T1 = Pretest

X = Treatment (Picture)

T2 = Posttest

(Setiyadi, 2006:135)

3.2 Population and Sample

The population of this research was the students of the first grade of SMPN 18 Bandar Lampung. The researcher selected one class, class VII A as the subject of the research. The researcher used lottery technique to choose the treatment class. So that those all classes got same chance to be sample.

3.3 Research Procedures

The procedures of this research were as follows:

1. Determining the population and selecting sample.

   The population of this research was the first grade students of SMP Negeri 18 Bandar Lampung. The sample of this research was VII A which consisted of 30 students.

2. Administering the try-out

   Try-out was administered with the aim to know the quality of its test and to determine the items that should be omitted for pre-test and post-test. The time allocation of the try out was 80 minutes with 50 questions.

3. Adminstering the pre test

Pre-test was administered to measure the students' vocabulary mastery before treatments. The time allocation of pretest was 80 minutes with 40 questions.

4. Conducting the treatment.

After giving pre test, the students were given three times treatments by using flashcards.

5. Administering the post test.

The post-test was administered after treatments. The post-test was conducted in 80 minutes with 40 questions.

6. Analyzing the data

The data compute through SPSS version 16.0.

3.4 Data Collecting Technique

1. Pretest for Vocabulary

The pretest was conducted before the treatment. It was used to know how far the students have mastered the vocabulary before treatment were given. The test used by the researcher was an objective test in the form of multiple choices. The number of the items in the test was 40 in which each item has four options of answer (A, B, C, D). The time allocation of this test was 80 minutes. In this research, the researcher chose content words (noun, verb, adjective, adverb).

2. Posttest for Vocabulary

The posttest was administered after the researcher conducted treatments. It was used to know the improvement of students' vocabulary mastery after

being taught through flashcards. Similar with the pretest, the researcher used a test in the form of multiple choice tests that consisted of 40 questions. The questions were the same as the pretest. The time allocarion of this test was 80 minutes. In this research, the researcher chose content words (noun, verb, adjective, adverb).

3. Pretest Active Vocabulary

Administered the active vocabulary test in order to find out the students' vocabulary use in writing before being taught through flashcards. It required 40 minutes.

4. Posttest Active Vocabulary

Administered the active vocabulary test in order to find out whether there was improvement of students' vocabulary use in writing after being taught through flashcards. It required 40 minutes.

3.5 Instruments of the Research

The research instrument were vocabulary test in the form of objective test, multiple choice and writing test. In this research, the researcher administered three kinds of the test: try out test, pre-test, and post-test. The total number of the items of the try out was 50 items, and the total number of the items for pretest and posttest were determined from the result of the try-out test. In those three tests contained content words, consist of nouns, verbs, adverbs, and adjectives.

1. **Instrument of Passive Vocabulary**

Table 1. Table Specification of Try Out Test

No.	Content Words	Number of Items	Percentage	Item Numbers
1	Noun	13	26%	10, 12, 14, 16, 17, 20, 21, 22, 28, 29, 43, 45, 50
2	Verb	13	26%	1, 2, 3, 4, 5, 11, 33, 35,36, 42, 46, 47, 48
3	Adverb	12	24%	6, 7, 15, 18, 19, 23, 27, 30, 31, 32, 34, 41
4	Adjective	12	24%	8, 9, 13, 24, 25, 26, 37, 38, 39, 40, 44, 49
	Total	50	100%	

2. Instrument of Active Vocabulary Test

In this research, the researcher made active vocabulary test in writing form. The researcher asked the students to describe a picture. In scoring the active vocabulary test, the reseracher calculated how many content words (Noun, Adjective, Verb, Adverb) that were used by the students. In measuring active vocabulary test the researcher used Lexical Density. Lexical density is the term most often used for describing the proportion of content words (Nouns, Adjective, Verb, and Adverb) to the total number of words. The result is a percentage for each text in the corpus. Ure (1971:452) concludes that a large majority of the the written texts have a lexical density of 40% or higher. This is the example of Lexical Density of written language.

In 1918, <u>when</u> the <u>chemical</u> <u>industry</u> was <u>first</u> <u>established</u> in the <u>area</u>, <u>Billingham</u> was a <u>village</u> <u>inhabited</u> by a <u>few</u> hundred <u>people</u> but <u>grew</u> <u>rapidly</u> as <u>ICI</u>'s <u>operations</u> <u>expanded</u>, <u>helped</u> by the <u>company</u>'s <u>reputation</u> for <u>providing</u> <u>secure</u> <u>employment</u>. The <u>wages</u>, <u>conditions</u> and <u>benefits</u> <u>offered</u> by ICI <u>were</u> <u>attractive</u> and the <u>company</u> <u>quickly</u> <u>gained</u> a <u>reputation</u> as a <u>good</u> <u>employer</u>. <u>Many</u> of those we <u>interviewed</u> <u>claimed</u> this <u>was</u> their <u>main</u> <u>reason</u> for <u>applying</u> for a <u>job</u>. Our <u>interviews</u> <u>also</u> <u>highlight</u> the <u>influence</u> of <u>family</u> when <u>making</u> <u>decisions</u> about <u>employment</u> and <u>ICI</u> <u>was</u> <u>certainly</u> <u>happy</u> to <u>recruit</u> the <u>sons</u> and <u>daughters</u> of <u>existing</u> <u>workers</u>.

Lexical density is calculated as follows:

Lexical density = (number of lexical words/total number of words) * 100

$= (60/102) * 100 =$ **58.8%**

3.6 Instrument of the Try Out

In this research, try out is an important thing. It measures the students' mastery of vocabulary before the students got pretest and posttest. The try out test can be said have a good quality if it has a good validity, reliability, level of difficulty, and discrimination power.

1. Validity of Try Out

Validity refers to the results of the test not to the test itself (Hatch and Farhady, 1982:251). The resrearcher only focuses on content validity and construct validity

because content validity and construct validity have represent another aspects to measure whether the test has a good validity.

- **Content Validity**

  Content validity cencerns with how well the test represents the subject matter content or behaviors to be tested. The procedure for determining content validity is to compare the test content with the universe of content or behaviors supposedly being measured (Hatch and Farhady, 1982:252).

  To fulfill this validity, the researcher should see all the indicators of the instrument and analyze them to see whether it has represented the material that measured or not. To get content validity, the test has been adapted from the students' book and has been determined according to the material that has been taught. To know whether the test has a good content validity, the items of the test discussed with the English teacher of SMPN 18 Bandar Lampung.

Tabel 2. Table Specification of Vocabulary Test

No.	Content Words	Number of Items	Percentage	Item Numbers
1	Noun	13	26%	10, 12, 14, 16, 17, 20, 21, 22, 28, 29, 43, 45, 50
2	Verb	13	26%	1, 2, 3, 4, 5, 11, 33, 35,36, 42, 46, 47, 48
3	Adverb	12	24%	6, 7, 15, 18, 19, 23, 27, 30, 31, 32, 34, 41
4	Adjective	12	24%	8, 9, 13, 24, 25, 26, 37, 38, 39, 40, 44, 49
	Total	50	100%	

- **Construct Validity**

   According to Setiyadi (2006:6) if the instrument just measures one aspect, for example vocabulary, the construct validity can be measured by evaluating all items in the test. If all items have measured vocabulary mastery, this instrument has fulfilled construct validity. This instruments would have fulfilled construct validity. Construct validity examines whether the test is actually in line with the theory, meaning whether the test is in line with the school curriculum. Content words is supposed to be comprehened by grade VII students of Junior High School.

2. Reliability

Reliability is a consistency of a measurement or how far that measurement can be measured the similar subjects in different time but showed the same result (Setiyadi, 2006:16). Knowing the reliability of the half of the test, Arikunto's formula is used to get the result of the computation of try out test. After the researcher has obtained the reliability of half of its test, the researcher used Spearman Browns Prophecy formula (Hatch and Farhady, 1982:246).

$$rxy = \frac{n\sum XY - (\sum x)\ (\sum Y)}{\sqrt{(n \sum X^2 - (\sum X)^2\ (n\sum Y^2 - (\sum Y)^2)}}$$

Where:

r : coefficient of reliability between odd numbers and even numbers items

x : total numbers of odd numbers items

y : total numbers of even numbers items

n : numbers of students who take part in the test

x^2 : square of x

y^2 : square of y

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

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

<div align="right">(Arikunto, 2002:157)</div>

The criteria of reliability:

0.80 – 1.00   : very high

0.60 – 0.79   : high

0.40 – 0.59   : average

0.20 – 9.39   : low

0.00 – 0.19   : very low

To compute the coefficient correlation of the whole items, the researcher used Spearmen Brown Prophecy Formula.

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

Where:

rk : reliability of a full test

rl : reliability of half test

the criteria of reliability are:

0.90 – 1.00   : high

0.50 – 0.89   : moderate

0.00 - 0.49   : low

<div align="right">(Hatch and Farhady, 1982:246)</div>

The result of the computation by using Arikunto's formula showed that reliability of the half test (r_{xy}) was 0.98 (See Appendix 11). Using Spearman Browns Prophecy Formula, it was found that the reliability of the whole of the test (r_k) was 0.98 (See Appendix 11). Based on the criteria of the reliability, the reliability 0.98 belongs to the high level. It indicated that data collecting technique instrument in this research was reliable.

3. Level of Difficulty

To see the level of difficulty, the researcher used the following formula:

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

Where:

FV   : level of difficulty

R   : numbers of students answers correctly

N   : total number of the students

Heaton (1975:182) states the criteria of the index of difficulty. The criteria are:

$FV < 0.30$: difficult

$FV = 0.30 - 0.70$: satisfactory

$FV > 0.70$: easy

There were three criteria of level difficulty; easy, satisfactory, and difficult. The items or the test can be said it was good if the items are not too easy and difficult. From the test, the data has been analyzed and the computation of this test (See Appendix 9), it was found that there were 2 items which were less than 0.30, it

means that the items were difficult. There were 28 items were more than 0.70 it was easy and 20 items were satisfactory (between 0.30-0.70).

4. Discrimination Power

Discrimination power is used to indicate the discrimination power of the fail and the success of the students. To find out the discrimination power the researcher used the following formula, adapted from Heaton (1975:182):

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

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 of discrimination power are:

0.00 – 0.19   : poor

0.20 – 0.39   : satisfactory

0.40 – 0.69   : good

0.70 – 1.00   : excellent

- (negative)   : bad items, must be omitted

From the computation of discrimination of try out test (See Appendix 9), it was found 7 items was poor the discrimination power of them were 0.00 – 0.19. There were 3 items was bad because the discrimination power of both were negative, 23

items were satisfactory (0.20 – 0.39), 16 items were good (0.40 – 0.69) and 1 item was excellent.

Based on the table of discrimination power (See Appendix 9), there were 10 items were omitted (1, 2, 4, 5, 9, 13, 16, 21, 44, 46). The researcher took 40 items were administered in pre-test and post-test.

3.7 Validity and Reliability of Active Vocabulary

1. Validity

The type of validity that used was content validity. Content validity is concerned with whether the test is sufficiently representative and comprehensive for the test. The test should be so constructed as to contain a representative sample of the course (Heaton, 1975:160). Researcher used the vocabulary that was supposed to be comprehened by the first grade students of Junior High School. The researcher made the test related to which is content words because content words can cover the material. The material was vocabulary of occupation, contain: verb, adjective, verb, and adverb.

2. Reliability

Shohamy (1985) states that reliability refers to the extent to which the test is consistents in its score. It can also give an indication of how accurate the test score. To measuring reliability of the test, the researcher counted the result of the test used lexical density method. The basic method used is that of counting the lexical density of the texts. This is term used by Ure (1971:445) to describe the percentage of lexical (content words) in the total number of words in any given

text, either written or spoken. The researcher used Lexical Density to determine the reliability of active vocabulary test.

$$L_d = (N_{lex}/N) \times 100$$

Where:

L_d = the analysed text's lexical density

N_{lex} = the number of lexical word tokens (nouns, adjectives, verbs, adverbs) in the analysed text

N = the number of all tokens (total number of words) in the analysed text

The criteria of lexical density:

<40% : Low

>40% : High

Ure (1971: 443)

To measuring the reliability of the result from the text analysis the researcher counted the result by using Lexical Density Formula.

Text in the Pre test

My English Teacher

I has <u>english</u> <u>teacher</u> in my <u>school</u>. He is <u>Mr. Bambang</u>. He <u>still</u> <u>young</u>. His <u>age</u> 30 <u>years</u>. He <u>very</u> <u>discipline</u> <u>person</u>. He <u>never</u> <u>late</u> <u>come</u> to the <u>class</u>. He <u>always</u> <u>use</u> <u>glasses</u>. He has <u>long</u> <u>mustache</u>. <u>Female</u> <u>students</u> in my <u>school</u> <u>like</u> him. He is <u>very</u> <u>diligent</u> <u>teacher</u>. I <u>think</u> he is <u>very</u> <u>creative</u>. He <u>always</u> <u>make</u> <u>joke</u> or <u>game</u> <u>when</u> he <u>felt</u> boring in the <u>class</u>. He <u>always</u> <u>smile</u> if he <u>meets</u> with the <u>students</u>. <u>When</u> he is <u>teaching</u>, he is <u>very</u> <u>serious</u>.

Lexical words   **: 45**
Total Numbers of words : 85
LD = (N_{lex} / N) X 100
   = (45/58) X 100 = 52.94%

Text in the post test

My Favorite Teacher

My favourite teacher is Mister Anton. He teaches Mathematics lesson. He is fourty one years old. He lives in Teluk Betung. He is tall and thin. He is very handsome. He has black hair, his hair is curly. Mister Anton has bright skin. His eyes is black.

Mister Anton is smart person. He is very patients and friendly. He always helps the students when we can not do the task. He also answer when his students ask to her. That's way many students like him very much. Mister Anton really good teacher in my school. Teacher is good job because teacher teach many knowledge for us.

Lexical words: 56

Total numbers of words : 101

LD = (N_{lex} / N) X 100

= (56/101) X 100 = 55.44%

The result from the computation above, we can see that the score in pre test is 52.94% and in the post test is 55.44%. Based on the criteria of the Lexical Density, it indicated that the text is over 40%. It means that the text has good reliability.

3.8 Scoring System

In scoring the students result of the test, this research uses Arikonto's formula. The ideal score is 100. The score of pretest and posttest are calculated by using formula as follows:

$$S = \frac{R}{N} \ 10$$

Where:

S   : The score of the test

R   : The total of the right answer

N   : The total items

<div align="right">(Arikunto, 1997:212)</div>

3.9 Measuring Active Vocabulary Test

To measure active vocabulary in writing the researcher used lexical density. Ure (1971: 443) defines lexical density is the term most often used for describing the proportion of content words (nouns, verbs, adjectives, and often also adverbs) to the total number of words. Various variants of lexical density have been proposed. A popular 'minor variant' is to calculate the *noun density,* the number of nouns divided by the total number of tokens in the text. Other options are for instance verb or adjective or adverb types per total lexical words. Ure (1971:445) distinguishes between words with lexical properties, and those without. Lexical density is then defined as the total number of words with lexical properties divided by the total number of orthographic words. The result is a percentage for each text in the corpus. In order to calculate lexical density we need to make a distinction between different types of words:

1. **Lexical words**

- lexical verbs (e.g. *run, walk, sit*)

- nouns (e.g. *dog, Susan, oil*)

- adjectives (e.g. *red, happy, cold*)

- adverbs (e.g. *very, carefully, yesterday*)

2. **Function words**

- auxiliary verbs (e.g. *can, will, have*)

- numerals (e.g. *two, three, first*)

- determiners (e.g. *the, those, my*)

- pronouns (e.g. *she, yourself, who*)

- prepositions (e.g. *in, to, after*)

- conjunctions (e.g. *and, but, if*)

Lexical density may be determined thus:

$$L_d = (N_{lex}/N) \times 100$$

Where:

L_d = the analysed text's lexical density

N_{lex} = the number of lexical word tokens (nouns, adjectives, verbs, adverbs) in

the analysed text

N = the number of all tokens (total number of words) in the analysed text

Ure (1971: 443)

3.10 Data Analysis

After conducting pretest and posttest, the researcher analyzed the data. It was used to know whether there is improvement of the student's english vocabulary mastery after being taught through flaschards in writing. The researcher examine the students score by using following steps:

1. Scoring pretest and posttest.

2. Calculating the total correct answer of pretest and posttest.

3. Tabulating the score of the student's English vocabulary test results using t-test. The formula manually is as follows:

$$\frac{\overline{X1} - \overline{X2}}{\overline{S D}}$$

In which

$$S\frac{N}{D} = \frac{S_D}{\sqrt{n}}$$

Where:

XI = Mean of the pretest

X2 = Mean of the posttest

$S\frac{N}{D}$= Standard eror of differences between two means (denominator)

SD = Standard deviation

N = number of students

(Hatch and Farhady, 1982:116)

4. Drawing conclusion from the tabulated result of the pretest posttest, that is statistically analyzed by usig SPSS (Statistical Program for Social Sciences) in order to examine whether increase of the students gain is significant or not.

5. Scoring active vocabulary pretest and posttest.

3.11 Normality Test

Normality test is used to know whether the data in pretest and post test are distributed normally or not. The hypotheses of the normality test is as follows:

H_1 : The data is distributed normally

H_0 : The data is not distributed normally

In this research, the criteria the hypotheses are:

H_1 is accepted if level of significance at 0.005 (Sig.$>\alpha$). Meanwhile, H_1 is rejected

if significant value does not exceed level of significance at 0.005.

3.12 Hypotheses Testing

The hypotheses are:

H_1 : There is improvement of students' vocabulary mastery by using

        flashcards in teaching vocabulary.

H_0 : There is no improvement of students' vocabulary mastery by using

        flashcards in teaching vocabulary.

H_1 : There is improvement of students' vocabulary use in writing after being

        taught through flashcards.

H_0 : There is no improvement of students' vocabulary use in writing after

        being taught through flashcards.