

III. METHOD

This chapter provides an overview of research design, population and sample, instrument, criteria of good vocabulary learning strategies questionnaire and vocabulary size test, research procedure, data analysis, hypothesis testing that would be applied in this research.

3.1. Research Design

The design used in this research was *ex post facto design* because there was no control group and no treatment to the sample but collecting the data by analyzing the correlation between cause and effect among the variables in the research. This research was quantitative descriptive since the data gathered was in form of number. Vocabulary learning strategies questionnaire based on Schmitt's (1997) taxonomy of L2 vocabulary learning and adapted from related research conducted by Kallayanasute (2011) was distributed and also a vocabulary size test conducted by Sutarsyah (2006) was administered to the sample of the research. This study dealt with three kinds of variables, they were gender as the independent variable, vocabulary learning strategies as the independent variable, and vocabulary size as dependent variable. According to Setiyadi (2006), the research design could be represented as follow :

Where :

| | | |
|-----------|-----------|-----------|
| G1 | T1 | T2 |
| G2 | T1 | T2 |

1. Gender { G_1 (male) and G_2 (female) } and Vocabulary Learning Strategies (T_1) as independent variables because it is assumed that gender and vocabulary learning strategies preferences influences students' vocabulary size.
2. Vocabulary Size (T_2) as a dependent variable because it is assumed that vocabulary size is affected by students' gender and vocabulary learning strategies preferences.

3.2. Data Collecting Technique

In collecting the data, the researcher used:

1. Questionnaire

After deciding the subject, the researcher gave the questionnaire to the sample. The questionnaire items were intended to measure the VLS preferences under five categories of strategies, i.e, determination, social, memory, cognitive, and metacognitive strategies. The questionnaire instruments consisted of 30 items, there were 6 questions for each strategy. It was designed specifically based on Schmitt's (1997) taxonomy of L2 vocabulary learning and used by the previous researcher by Kallayanasute (2011) (see Appendix 5). In addition, the result of the questionnaire was scored based on Likert Scale. The scores ranged from 1-5. There are five options to answer the question in one item of questionnaire. They are 1 (never use it), 2 (seldom use it), 3 (sometimes use it), 4 (often use it), and 5 (always use it).

2. Vocabulary Size Test

This step had been done after the researcher gave the questionnaire to the sample. The researcher gave the word-definition-matching format (WDMF) which was arranged by Sutarsyah (2006). The vocabulary test was consisted of four levels and each level consisted of 40 items as a sample (see Appendix 2). This level was derived from Bauer & Nation (1993) word level (1st 1000 words, 2nd 1000 words, 3rd 1000 words which was basically derived from the list of frequency word order of *General Service List* (GSL) and 800 words of University Word List (UWL). Thus, up to the 3rd level word levels consisted of 1000 base words and for UWL consisted of 800 base words.

The scoring was based on correct answer, that was, one correct answer was scored 1 and for the wrong answer was scored 0. The students' vocabulary size was gained by counting the number of correct answers to be divided by the number of items in every vocabulary level test. Then, we could estimate the students' vocabulary size by multiplying it with the population. At last, to get the final score the total vocabulary size of the four levels was divided by 38, because the maximum vocabulary size for all levels would be 3800.

3.3. Population and Sample

The problem of this research was particularly focused on investigating the difference of vocabulary learning strategy preferences between males and females

of the second grade students in learning vocabulary of English. The population of this research was the second grade students of SMAN 1 Bandar Lampung. The total sample of students involved in the study were 40 students. It consisted of 20 males and 20 females students. In determining the sample the researcher used *purposive sampling*. The aim of this purposive sampling was the data which has been gotten by the students will represent the cases that can solve the research questions.

3.4. Instruments

1) Questionnaire

Questionnaire is a list of questions and statements to be answered by the students. The questionnaire is used to measure the students' vocabulary learning strategy preferences. This questionnaire has two parts. The first part is designed to know the student's identity. The data in this section included the respondents' details about their age, gender, and class. Then, the second part is the questionnaire about students' frequency of the use of vocabulary learning strategies implemented by the respondents.

The questionnaire instruments consisted of 30 items. It was designed specifically based on Schmitt's (1997) taxonomy of L2 vocabulary learning and used by the previous research by Kallayanasute (2011) (see Appendix 5). Those questionnaire items measured the VLS preferences under five categories of strategies, there are determination, social, memory, cognitive, metacognitive

strategies. To be more specific, here is the table of specification for the questionnaire :

Table 3.1. Specification of the questionnaire

| Strategy Measured | Number of Questionnaire |
|-------------------|-------------------------|
| Determination | 1-6 |
| Social | 7-12 |
| Memory | 13-18 |
| Cognitive | 19-24 |
| Metacognitive | 25-30 |

The result of questionnaire was scored based on Likert Scale. The score ranged from 1-5. There were five options to answer the question in one item of questionnaires. They are 1 (never use it) , 2 (seldom use it), 3 (sometimes use it), 4 (often use it), and 5 (always use it). In addition, the scoring system of this questionnaire was based on the average of each strategy. To get the VLS preferences of each strategy, the following scales were used to indicate the frequency of the usage of each strategy. According to the five-point rating scale, the values in the level of using could be explained as follows:

Table 3.2. Average value of the questionnaire

| Average value | Result |
|---------------|----------------|
| 1.00-1.50 | Never used |
| 1.51-2.50 | Seldom used |
| 2.51-3.50 | Sometimes used |
| 3.51-4.50 | Often used |
| 4.51-5.00 | Always used |

The overall average indicated the preferences of the learners' vocabulary learning strategies. The average for each subscale of the inventory shows which group of strategies (i.e, determination, social, memory, cognitive, and metacognitive) which learners preferred to use. Then, by having this information learners would know if they were in very high or very low in any of these strategy groups.

2) Vocabulary Size Test

The Instrument used in this research was called WDMF or *word definition matching format* (Sutarsyah, 2006). The test belongs to the type of *breadth* vocabulary test. The students would use word definition matching format , translation. The vocabulary test consisted of four levels and each level consisted of 40 items as a sample. This level was derived from Bauer & Nation (1993) word level (1st 1000 words, 2nd 1000 words, 3rd 1000 words which was basically derived from the list of frequency word order of *General Service List* (GSL) and 800 words of University Word List (UWL). Thus, up to the 3rd level, the word levels consisted of 1000 base words and for UWL consisted of 800 base words. The students only need to match the two list, left and right, as shown in the example below :

Table 3.3. Word Definition Matching Format

| | |
|------------|-----------------------------|
| A. Private | |
| B. Blame | 1. _____ melompat |
| C. Jump | 2. _____ pertama |
| D. slow | 3. _____ pribadi/tidak umum |
| E. sorry | |
| F. total | |

3.5. Criteria of Good Test

In this research, to prove whether the VLS questionnaire and vocabulary size test have good quality, it must be tried out first. As Heaton (1991) states that a test would be said to have a good quality if it has good validity, reliability, level difficulty and discrimination power. The questionnaire is also called as a good test if it has good validity and reliability.

3.5.1. Validity

Generally, the validity of a test shows how far the test measures what is supposed to be measured (Setiyadi, 2006). Validity can be defined as the degree to which a test actually test what it is intended to test. To measure whether the instruments have a good quality or not, the researcher used the content validity and construct validity. Content validity is the extent to which a test measures a representative sample of the subject matter content (Hatch and Farhady, 1982). According to this validity, the items of the test should represent the material that has been taught in the class. The items in the test of vocabulary size are based on KTSP curriculum of general high school. Then, the construct validity is concerned with whether the test is actually in line with the theory (Hatch and Farhady, 1982).

3.5.1.1. Validity of the questionnaire

The validity of questionnaire is also measured to find if the components are suitable and related to the relevant theories of vocabulary learning strategy. Hatch and Farhady (1978) states that there are at least two validities should be fulfilled namely content and construct validity. Since the questionnaire was adopted from Kallayanasute (2011) the researcher considered that the construct validity of the questionnaire has been standardized. Therefore the researcher measured the constructed validity only. Construct validity is necessary for measurement instrument which has several indicators in measuring one aspect or construct (Setiyadi, 2006).

The composition of the questionnaire items was presented in the table of specification below :

Table 3.4. Specification of VLS Questionnaire.

| Strategy Measured | Number of Questionnaire |
|--------------------------|--------------------------------|
| Determination | 1-6 |
| Social | 7-12 |
| Memory | 13-18 |
| Cognitive | 19-24 |
| Metacognitive | 25-30 |

Based on the table above, it can be seen that all of aspects of vocabulary learning strategies by Schmitt's taxonomy have been included in the questionnaire. The researcher considered that the construct validity of the questionnaire has been standardized.

3.5.1.2. Validity of the vocabulary size test

Validity refers to the extent to which the test measures and to what is intended to measure. There are two basic types of validity; content validity and construct validity (Hatch and Farhady, 1982:250). A test can be considered to be valid if it can precisely measure the quality of the test. There are four kinds of validity that are :

- 1) Face validity, refers to the way the test looks to examinees, teachers, and experts in testing (Jafarpur, 1987). It usually concerns with the lay out of the test.
- 2) Content validity, concerns with the systematic examination of the content of the test.
- 3) Construct validity, measures certain specific characteristic in accordance with a theory of language learning.

- 4) Criterion-related validity, concerns with measuring the success in the future as in replacement test.

According to the types of validity above, the researcher used content and construct validity. Both of them are explained as follows:

3.5.1.2.1. Content Validity

Content validity is intended to know whether the test items are good reflection of what will be covered or not. The test items which are adapted from the materials that have been taught to the students should be constructed as to contain a representative sample of the course (Heaton, 1988). Content validity essentially concerns with the systematic examination of the test content. The test should cover a representative sample of the behavior domain to be measured (Jafarpur, 1987). WDMF was prepared with a balance proportion of the levels that were going to be used.

3.5.1.2.2. Construct Validity

Construct validity deals with the extent to which the test is based on theory of the trait under consideration (Jafarpur, 1987). The WDMF test was designed based on the frequency count that classified vocabulary levels and it is also used to make sure if learners of English have mastered the threshold level of vocabulary.

Based on the analysis of validity by using *Pearson Product Moment* formula, all items in the questionnaire of vocabulary learning strategies were valid (Appendix 14). The r values of the 30 statements were above the r table (2,048). Therefore,

the questionnaire was used completely as what has originally been conducted from the previous research by Kallayanasute (2011).

3.5.2. Reliability of the Instruments

Reliability is the overall consistency of a measure. A measure is said to have a high reliability if it produces similar results under consistent conditions. In this research there were two instruments used. The first is vocabulary learning strategy questionnaire and the second is the vocabulary size test. In order to construct a good research, the data collecting techniques should be try-out first whether the instruments has good quality to be used.

3.5.2.1. Reliability of the questionnaire

The questionnaire was scored according to Likert scale whereas the reliability of the questionnaire was measured by using *Cronbach Alpha Coefficient*. The researcher used this because it is the most common scoring to assess the consistency of the indicators in the questionnaire. The alpha ranges between 0 and 1. The alpha 1 shows the perfect reliability of the test items. This means that the higher the alpha, the more reliable the questionnaire will be (Setiyadi, 2006: 167). The statistical calculation on validity and reliability of questionnaire was done by using SPSS program version 16.

Having analyzed the items of the questionnaire by using *Cronbach Alpha* formula, the computation showed the reliability coefficient of the questionnaire was 0,809.

Table 3.5. The reliability of questionnaire

| Reliability Statistics | | |
|------------------------|--|------------|
| Cronbach's Alpha | Cronbach's Alpha Based on Standardized Items | N of Items |
| .809 | .800 | 30 |

The researcher found out that all were valid. that the questionnaire could be used to investigate the students vocabulary learning strategies see (Appendix 10).

3.5.2.2. Reliability of the Vocabulary Size test

Reliability refers to whether the test is consistent in its score and gives us an indication of how accurate the test score are (Shohamy, 1985: 70). A test is called reliable if the score gained by the examiners is constant whenever and by whomever the test is conducted. A test would not be a good parameter unless the test is suitable or constant. To measure the reliability of the test, the researcher used the split-half method to find the reliability of the vocabulary size test by dividing the number of the tests items into two groups for each test, odd and even. Based on *Pearson Product Moment* formula, the formula can be seen as follows:

$$r = \frac{\sum XY - \frac{(\sum X)(\sum Y)}{n}}{\sqrt{\left[\sum X^2 - \frac{(\sum X)^2}{n}\right] \left[\sum Y^2 - \frac{(\sum Y)^2}{n}\right]}}$$

r = coefficient reliability between odd and even number

X = odd number

Y = even number

n = numbers of students who take part in the test

X² = square of X

Y² = square of Y

ΣX = total score of odd items

ΣY = total score of even items

(Arikunto, 1997)

After the reliability of the half test was calculated, the researcher used Spearman Brown's Prophecy formula to measure the reliability of the test as a whole as follows:

$$r_k = \frac{2r_{xy}}{1 + r_{xy}}$$

r_k = the reliability of the whole test

r_{xy} = the reliability of half test (Hatch and Farhady, 1982: 247)

According to Arikunto (1998: 260), the standard of reliability of the instrument can be described as follows :

- 1) 0.80 - 1.0 : very high reliability
- 2) 0.60 - 0.79 : high reliability
- 3) 0.40 - 0.59 : medium reliability
- 4) 0.20 - 0.39 : low reliability
- 5) 0.0 - 0.19 : very low reliability

After conducting the try out test, the researcher used split-half method to estimate the reliability of the test to make sure whether the test was appropriate for testing or not. In using the split-half method, the researcher classified the test items into two similar parts (usually the odd- numbered items and the even-numbered items). By splitting the test into two similar parts, it was supposed as if the whole tests had been taken twice. The calculation showed that the reliability coefficient of the test was = 0,847 (Appendix 8). It can be assumed that, this instrument has very high reliability and

proper to be used in the research since the range of very high criteria in the criteria of reliability was 0.80 – 1.00 (Arikunto, 1998: 260).

3.5.3. Level of Difficulty

Level of difficulty relates to how easy or difficult the item is from the point of view of the students who took the test. It is important since test items which are too easy (that all students get right) can tell us nothing about differences within the test population. (Shohamy, 1985: 79). Level of difficulty was calculated by using the following formula:

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

LD = level difficulty

R = number of students who answers it right

N = total number of students

The criteria are:

LD < 0.30 = difficult

LD = 0.31- 0.70 = satisfied

LD > 0.71- 1.00 = easy (Arikunto, 1997:214)

Then, the result of difficulty level of the vocabulary size try-out test showed that there were 32 easy items, 120 average items, and 8 difficult items in this instrument. Below is the table displaying the difficulty level of the vocabulary size try-out test.

Table 3.6. Difficulty Level of the Vocabulary Size Try-out Test

| Classification | Items | % |
|----------------|--|-----|
| Easy | 5,8,13,25,28,49,50,53,60,63,67,75,80,92,99,101,104,108,109,110,118,120,122,123,125,126,128,140,142,144,147,160 | 20% |
| Average | 1,2,3, 6, 7, 9,10, 11, 12, 13, 14,15, 16, 17, 18, 19, 20, 21, 22, 23,24, 26, 29, 30, 31,32,33,34,36,37,39,40,41,42,43, | 75% |

| | | |
|------------------|---|------|
| | 44,45,46,47,48,51,52,54,55,56,57,58,59,61,62,64,66,68, 69,70,71,72,73,74,76,77,79,81,82,83,84,85,86,87,90,91,94 ,95,96,97,100,102,103,105,106,107,110,111,112,113,114, 115,117,119,121,124,129,150,131,132,133,134,135,136,1 38,139,141,143,145,146,148,149,159 | |
| Difficult | 4,27,78,88,89,93,127,137 | 5% |
| Total | 160 items | 100% |

3.5.4. Discrimination power of the Test

Discrimination power 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 in which good students did well, and bad students failed” (Shohamy, 1985: 81). To calculate the discrimination power (DP) of the test items, the researcher used the following formula:

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

In which,

DP: Discrimination Power

U: the total of correct answer of the higher group

L: the total of correct answer of the lower group

N: total number of students

(Shohamy, 1985)

The criteria are:

0.00- 0.20 : Poor

0.21- 0.40 : Satisfactory

0.41- 0.70 : Good

0.70- 1.00 : Excellent (should be omitted)

Based on the try out test analysis, the discrimination Power of the vocabulary size would be explained as follows. In the vocabulary size try-out test, there were 54

good items, 89 satisfactory items, and 17 excellent items. The following table presents the distribution of discrimination power of this instrument.

Table 3.7. Discrimination Power of Vocabulary Size Try-out Test

| Classification | Items | % |
|----------------|--|--------|
| Good | 1,2,6,7,8,12,13,14,16,17,18,19,20,22,23,24,25,26,28,27,30,36,37,39,40,41,45,46,49,52,57,58,60,66,67,71,72,74,75,112,119,123,128,133,135,136,144,147,148,150,151,155,156,157 | 33,75% |
| Satisfactory | 4, 5, 9, 15,38,42,43,44,50,51,53, 54,55,59,61,62,63, 64,65,68,69, 73,76,77,78,79,83,84,86,88,89,91,92, 93,94,95,97,98,100,101,103,104,105,111,113,114, 115,116,117,118,120,121,123,124,125,126,127,129, 130,131,132,134,137,139,140,141,142,143,145,146, 149,153,154,150,159,160 | 55,62% |
| Excellent | 3,10,11,21,27,31,32,33,34,35,36,47,48,70,85,87,152 | 10,63% |
| Total | 160 Items | 100% |

Due to the analyses on the difficulty level and the discrimination power described above, the researcher would administered all the items of this instrument to measure the students' vocabulary size.

3.5.5. Scoring System

Since the pupose of the test was to measure learners' vocabulary size, the scoring was based on the correct response, i.e., one correct answer was scored 1; one wrong answer was scored 0. As already explained, the learners' vocabulary size was obtained by counting the number of correct answers to be divided by the number of items (sample) in every vocabulary level test. From this, we could find the proportion of correct answers. Then, we could estimate the learners' vocabulary size by multiplying it with the population; 1.000 words or each level with additional 800 words from UWL. For example, the WDMF test was used 40 items for every level. In order to know the students' vocabulary size in vocabulary levels test and the students' score were computed by doing these activities:

- 1) Scoring the vocabulary test
- 2) Calculating the result of the vocabulary test

$$SS = \frac{\sum s}{SMPL} \times 1000$$

SMPL = Subject's Vocabulary size at a given level

SS = a subject's score on the Levels Test at the relevant level

$\sum s$ = Total students' vocabulary score

For instance, if a student can answer 20 items correctly in the first level (1st 1000 words), this means the student estimated to have vocabulary of 500 words in the first level. That is by calculating the correct answer divided by 40 (sample), then multiply by 1000.

$$\text{For example : } \frac{20}{40} \times 1000 = 500$$

Therefore, the students' vocabulary size is considered to have 500 words in the first 1000 words level, etc. To get the final score, the total vocabulary size of four level is divided by 38, because the maximum score for all level would be 3800.

$$\frac{3800}{38} = 100$$

3.6. Research Procedure

Below, there were the steps of the procedure in conducting the research :

- a) Determining the research question and determining the focus of the research. In order to determine the problem, the researcher read the books and E-journals that relate to the topic.
- b) Determining the instruments

In this research, the researcher used the appropriate instruments in order to be able to be interpreted. The instruments were:

1. The Vocabulary Learning Strategies Questionnaire
2. The Vocabulary Size Test

c) Finding the sample

The sample of this research was one class of the second year students of SMA Negeri 1 Bandar Lampung. The class had been chosen purposively.

d) Trying out the instruments

Before distributing the instruments, the researcher conducted the try out to the students first in order to guarantee the result to be more valid.

e) Analysing the result of the try-out test

After both the questionnaire and the vocabulary size test were administered, the researcher analyzed the validity and the reliability of the instruments.

f) Distributing the instrument

Two of the instruments were distributed to the students.

g) Scoring the data

The researcher counted the score of questionnaire and vocabulary size test

h) Analyzing the data

The researcher in this step analyzed the tabulated data. The analysis was done by using SPSS 16.0 computer program, the researcher analyzed the data by using *One way Anova*.

i) Drawing conclusion

As the last step, conclusion was drawn up carefully based on the result of the data analysis.

3.7. Data Analysis

In the attempt to investigate the difference of vocabulary learning strategy preferences between males and females of the second grade students at SMA Negeri 1 Bandar Lampung, the researcher passed the following steps for analyzing the data.

- 1) Tabulating the score of the questionnaire and the Vocabulary Size Test.
- 2) Analyzing the data

The tabulated data were analyzed by using *one way Anova* in SPSS 16.0 program. In this research, some analyses were done by the researcher which involve the analysis of these following steps :

- a) The reliability and validity of the vocabulary learning strategies questionnaire and the vocabulary size test.
 - b) Level of difficulty and discrimination power of the vocabulary size test.
 - c) The difference of the vocabulary learning strategies questionnaire preferences based on gender.
- 3) Making inference

The researcher inferred the comparison of vocabulary learning strategies used by the second grade students of SMA N 1 Bandar Lampung based on gender and vocabulary size to the result of the *One Way Anova* statistical package.

3.8. Hypothesis Testing

The hypothesis testing was used to prove whether the hypothesis which was proposed in this research was accepted or not. In the effort to measure the hypotheses, *Pearson product-moment* of SPSS 16.0 was used considering that the data came from the same sample. The following were the hypotheses took by the researcher :

H_0 = there is no significant difference of vocabulary learning strategy preferences
between males and females of the second grade students at SMA Negeri 1
Bandar Lampung.

H_1 = there is significant difference of vocabulary learning strategy preferences
between males and females of the second grade students at SMA Negeri 1
Bandar Lampung.

In this research the hypotheses were analyzed by *One way Anova*, since it is a statistical technique which aimed to test the comparative hypothesis and to examine whether there are significant differences between the means of the two groups or more (Martono, 2011). The criteria for accepting the hypotheses based on *One Way Anova* are explained as follows :

H_0 is accepted if the F-count is lower than F-table.

H_1 is accepted if the F-count is higher than F-table.

In addition, the criteria H_0 (null hypothesis) is accepted if alpha level is higher than 0.05 ($\alpha > 0.05$).

This chapter has explained the overview of research design, population and sample, instrument, criteria of good vocabulary learning strategies questionnaire and vocabulary size test, research procedure, data analysis, hypothesis testing that would be applied in this research. The next chapter discusses about result and discussion of this research.