CHAPTER III

RESEARCH METHOD

This chapter describes the method used in this study, including the research design, subject of the study, research instruments, data collection and data analysis techniques.

A. Research Design

The research design of this study is correlation. Correlation is one of the types of quantitative research. According to Creswell, correlation study is used when researcher relates two or more variables to see if they influence each other.¹ The degree of relationship is expressed by the coefficient of correlation. The coefficient of correlation is usually given the symbol r and it ranges from

-1.00 to +1.00, which is interpreted as follows²:

1. A correlation coefficient of -1.00 indicates that there is a perfect negative relationship between the two variables. The scatterplot shows a downhill pattern from left to right because X decreases as Y increases. The plots are presented in the figure below:

¹ John W. Creswell, *Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research, 4th Edition* (USA: Pearson, 2012), 338

² Jim Higgins, *The Radical Statistician: A Beginners Guide to Unleashing the Power of Applied Statistics in The Real World*, 5th Edition (Jim Higgins Publishing, 2006), 10

Figure 3.1 Negative Correlation



2. A correlation coefficient of 0.00 indicated that there is a zero correlation, or no relationship, between the two variables. The scatterplot does not show a downhill or uphill pattern because Y does not tend to increase or decrease as X increases. The plots are presented in the figure below:



3. A correlation coefficient of +1.00 indicates that there is a perfect positive relationship between the two variables. The scatterplot shows an uphill pattern from left to right because X increases as Y increases. The plots are presented in the figure below:



Figure 3.3 Positive Correlation

Correlation is an effect size. Therefore, the strength of the correlation

can be described as³:

- a. 0.00 0.199 : very weak
- b. 0.20 0.399 : weak
- c. 0.40 0.599 : moderate
- d. 0.60 0.799 : strong
- e. 0.80 1.0 : very strong

³ Syofian Siregar, Statistik Parametrik untuk Penelitian Kuantitatif : Dilengkapi dengan Perhitungan Manual dan Aplikasi SPSS Versi 17 (Jakarta: Bumi Aksara, 2013), 337

B. Research Variables

There are two variables analyzed in this study, they are:

1. Independent variable

According to Creswell, an independent variable is an attribute or characteristic that influences or affects the dependent variable.⁴ In this study, the independent variable is vocabulary size.

2. Dependent variable

Meanwhile, a dependent variable is an attribute or characteristic that is dependent or influenced by the independent variable.⁵ In this study, the dependent variable is spoken word recognition in Literal Listening.

C. Subject of the Study

1. Population

Population is the larger group to which a researcher wishes to generalize; it includes all members of a defined class of people, events, or objects.⁶ In this study, the population is 108 2nd semester students academic year 2014-2015 of English Teacher Education Department in Sunan Ampel State Islamic University. The students are divided into 5 classes of Literal

⁴ John W. Creswell, *Educational Research*...115

⁵ John W. Creswell, *Educational Research*...116

⁶ Donald Ary, *Introduction to Research in Education*, 8th Edition (Wadsworth: Cengage Learning, 2010), 647

Listening, they are: A class (23 students), B class (18 students), C class (23 students), D class (24 students) and E class (20 students).

2. Sample

According to Arikunto, if the subject is less than 100, it is better to take all of the subjects so it becomes population research. But if a large number of the subjects (more than 100) exist, the sample can be taken 15%, 25% or more.⁷ A larger sample is much more likely to be representative of the population. In this study, the population is more than 100 or more precisely 108 students, thus the researcher takes some samples from the population. Sample itself refers to a set of elements selected in some way from a population. The aim of sampling is to save time and effort. Sampling is also beneficial to obtain consistent and unbiased estimates of the population status in terms of whatever is being researched.⁸ Slovin's formula is used to calculate the sample size. It is computed as:

$$n = \frac{N}{1 + N(e)^2}$$

Where:

n =sample size

N = population size

e = margin of error (in this study the margin of error is 5% or 0.05)

⁷ Suharsimi Arikunto, *Prosedur Penelitian Suatu Pendekatan Praktik* (Jakarta: Rineka Cipta, 2004), 134

⁸ Roger Sapsford – Victor Jupp, *Data Collection and Analysis*, 2nd Ed. (London: SAGE Publications, 2006), 26.

Therefore, the sample size of this study can be calculated as follows:

$$n = \frac{108}{1 + 108(0.05)^2} = \frac{108}{1 + 0.27}$$
$$n = \frac{108}{1.27} = 85 \text{ Students}$$

Since there are 5 Literal Listening classes, proportionate stratified sampling is used to decide the sample size from every class. Proportionate sampling refers to the selection from each sample that is proportionate to the size of population.⁹ It is calculated as:

$$n_x = \frac{N_x}{N} \times n$$

Where:

 n_x = the sample size for x class

 N_x = the population size from x class

N = the total of population size

n = the total of sample size

⁹ Yogesh Kumar Singh, *Fundamental of Research Methodology and Statistics* (New Delhi: New Age International Publishers, 2006), 88

Class	Number of Participants	Percentage (%)
А	$n_a = \frac{23}{108} \times 85 = 18.1 \approx 18$ students	21.18
В	$n_b = \frac{18}{108} \times 85 = 14.2 \approx 14 \text{ students}$	16.47
С	$n_c = \frac{23}{108} \times 85 = 18.1 \approx 18 \text{ students}$	21.18
D	$n_d = \frac{24}{108} \times 85 = 18.9 \approx 19 \text{ students}$	22.35
Е	$n_e = \frac{20}{108} \times 85 = 15.7 \approx 16 \text{ students}$	18.82
Total	85 Students	100%

Therefore, the sample size for every class is calculated as follows:

Table 3.1 Number of Participants in Each of Literal Listening Class

The number of students who participates in this study is 85 students as presented in the table above. 18 students (21.18%) are taken from A class, followed by 14 students (16.47%) from B class, 18 students (21.18%) from C class, 19 students (22.35%) from D class and 16 students (18.82%) from E class. The students from every class are randomly chosen. It means that each individual in the population has the equal and independent chance of being chosen.¹⁰ Here the word 'independent' means that the selection of one element does not influence the selection of any other. All members of the population have essentially the same probability of being selected.

¹⁰ Yogesh Kumar Singh, Fundamental of Research Methodology...84

D. Research Instruments

There are two instruments used in this study, the first instrument is Vocabulary Levels Test (VLT) version 2 developed by Schmitt et.al, and the second one is listening comprehension test. The details are described below:

1. Vocabulary Levels Test (VLT)

Vocabulary Levels Test (VLT) is a test based on the frequency-based sampling. This study uses Vocabulary Levels Test version 2 developed by Schmitt et.al to assess the students' vocabulary size (see Appendix 1). It consists of five sections, divided into four levels of vocabulary frequency (2,000, 3,000, 5,000 and 10,000-level) and one level of academic words from the Academic Word List. For each vocabulary level, words are classified into 10 clusters. Each cluster contains six words and three definitions, so there are 30 items in every level and 150 items overall. The test uses a matching format as in the following example, where three words are being tested (*horse, pencil, and wall*):

1 business	
2 clock	part of a house
3 horse	animal with four legs
4 pencil	something used for writing
5 shoe	
6 wall	

This test requires test takers to match a word with its definition by writing only the number of words on the blank space. The test-takers are given 30 minutes to complete this test. The test scores provide an indicator of the number of words at each frequency level that the students know. Each word chosen correctly is awarded one point. Since the Vocabulary Levels Test in this study consists of five levels, the highest possible score is 150 points. It is also necessary to determine a cut-off point for mastering a specific vocabulary level in the VLT. Schmitt, Schmitt and Clapham's validation of the test uses a mastery criterion of 26 out of 30. The same criterion is applied in the analysis of this study. Therefore, students who obtain a score of 26 or above at the 2000, 3000, 5000, 10.000 and academic vocabulary levels are thus regarded as having mastery of that particular vocabulary level.¹¹

2. Listening Comprehension Test

The listening comprehension test is a test to assess the students' spoken word recognition in listening. It consists of three listening worksheets. These listening worksheets are developed by the researcher based on the lecturers' syllabus. All of the worksheets are designed in intermediate level

¹¹ Norbert Schmitt, et.al, "Developing and Exploring the Behavior of Two New Versions of the Vocabulary Levels Test". *SAGE Publications*. Vol. 18 No. 1, 2001, 145

and in open questions following presentation of a listening text (questions scored on a scale of correctness and completeness). The worksheets are about:

a. Future Plans

There are two worksheets for this topic. The first worksheet is about the conversation between Crystal and George who are discussing about jobs and life after college (see Appendix 2). The second worksheet is about six speakers who are talking about their plans for tonight (see Appendix 3). Each of those worksheets consists of 10 blank spaces which have to be completed by the students. The students are instructed to fill in the blank with one or two words only. In scoring, the right answer is awarded 10 points, so the maximum score for each worksheet is 100. The recordings are taken from the English Listening Lesson Library Online (ELLLO) website. Each of the recordings is played twice.

b. Advice

This worksheet is about the conversation between Mike and Sonia who are talking about healthy eating (see Appendix 4). There are two activities in this worksheet; first the students have to mark the sentences with true or false. If they answer it with "false", they have to write the right sentence. There are 12 questions for this activity. Second, the students have to write 3 healthy and unhealthy eating based on Mike and Sonia's conversation. The maximum score for this worksheet is 100. The recording is taken from the book "Listening Extra" and it is played twice for answering those two activities.

From those worksheets, each student's mean score is taken to be analyzed with the student's total score of vocabulary size. Future Plans and Advice are chosen as the topic of the worksheets because these two topics have not been discussed by two different lecturers in Literal Listening classes.

E. Data Collection Technique

The data collection for this study is in the form of paper-and-pen testing which can be classified into three steps, they are: preparation, research process, and tabulation.

1. Preparation

First, the researcher prepared the printed Vocabulary Levels Tests to be given to the students. Second, the researcher gave the listening worksheets to the lecturers to be uploaded in the Literal Listening *Edmodo* groups, so the students brought their own printed worksheets to the class.

2. Research Process

The Vocabulary Levels Tests was given to all of the Literal Listening students in normal class hours. The time allotted to complete the test was 30 minutes. At the beginning of the test, clear instructions were given both orally and in written form. The oral instructions were given in the students' mother tongue to clarify what they were asked to do. The students were told that they were being asked to answer a test which measured how many words they knew in English. It was believed that the learners will be more motivated to do the test if they are also benefited from it, thus they were promised to get their test results. Yet, for the data collection through listening worksheets, it was helped by the lecturers so it will not disturb the teaching and learning process.

3. Tabulation

This process includes two stages: first, scoring the students' Vocabulary Levels Test (VLT) and listening comprehension test. The second stage is tabulating the data into Microsoft Excel before analyzing it with SPSS version 16.0.

F. Data Analysis Technique

The data is analyzed using Statistical Package for the Social Sciences (SPSS) version 16.0. The statistical devices used in this study are as follows:

1. Pearson Correlation (Product Moment)

Pearson correlation is used to answer the first research questions about the correlation between 2nd semester students' vocabulary size and their spoken word recognition in Literal Listening. This statistical device is widely used for interval or ratio data such as test scores.¹²

¹² Roger Sapsford – Victor Jupp, Data Collection and Analysis...225

2. Hypothesis Testing

The hypothesis testing is done by comparing the significance level (written by Sig. in SPSS) and the margin of error. The margin of error of this study is 5% or 0.05. The criteria of the hypothesis testing are described below:

- a. If the significance level is higher than 0.05 (> 0.05), then the Ho (null hypothesis) is accepted and Ha (alternative hypothesis) is rejected, which means there is no correlation between two variables.
- b. If the significance level is lower than 0.05 (< 0.05), then the Ho (null hypothesis) is rejected and Ha (alternative hypothesis) is accepted, which means there is a correlation between two variables.

3. Simple Linear Regression

Simple linear regression is used to examine the extent of association between vocabulary size and spoken word recognition in Literal Listening. There are some steps of the analysis: first, analyzing the coefficient of determination (the proportion of the variance in the dependent variable which is predictable from the independent variable) by converting the squared correlation coefficient to a percentage, and second, analyzing the simple linear regression equation.