

Chapter Three

Research Methodology

This chapter presents a research methodology used in this research. The points provided here are the research design used in this research, research setting, population and the sample of the research, and instruments of the research. The end of this chapter explains about data collection methods and data analysis used in this research.

Research Design

This research is a quantitative research. There are six quantitative research characteristics mentioned by Creswell (2012) while four of those characteristics are in line with this research. First, describing a research problem is through a description of trends or a need for an explanation of the relationship among variables. This research explained the relationship between students' receptive vocabulary size level and their paraphrasing ability. Second, quantitative research is creating purpose statements, research questions, and hypothesis that are specific, narrow, measurable, and observable. This research also has the purpose statements, research questions, and hypothesis which are specific, narrow, measurable and observable. Third, quantitative research is collecting numeric data from a large number of people using instruments with preset questions and responses. This research also collected data through instrument into a large number of people. Fourth, quantitative research is analyzing trends, comparing groups, or relating variables using statistical analysis, and interpreting results by comparing them with prior predictions and past research. This research was

related to students' receptive vocabulary size level and their paraphrase ability that were statically analyzed. Those characteristics are involved in this research which make quantitative research is suitable with this research.

This study used correlation research as a research design. Correlational research design is a statistical test to determine the tendency or pattern for two variables or two sets of data to vary consistently (Creswell, 2012). The type of correlation research design used in this research is explanatory research design. It is a correlation design in which the researcher is interested in the extent to which two variables (or more) co-vary, where changes in one variable are reflected in changes in the other (Creswell, 2012). The four common characteristics of explanatory research design are mentioned by Creswell; First, the investigator correlates two or more variables. Second, the researcher collects data at one point in time. Third, the investigator analyzes all participants as a single group. Fourth, the researcher obtains at least two scores of each individual in the group (one for each variable). Those characteristics are equal with this research setting in which it is correlated with the receptive vocabulary size and paraphrasing ability, collected the data only once, standardized all participants, and provided two data of each participant. The aim of explanatory correlation design is to explain the association between or among variables (Creswell, 2012), while this research's aim was to find out whether or not there is any correlation between students' receptive vocabulary size level and their paraphrasing ability. Thus, the explanatory correlation research design matched this research.

Research Setting

This research took place in English Education Department of Universitas Muhammadiyah Yogyakarta. The first reason why researcher chooses EED of UMY was EED of UMY lecturers often give paper assignments to the students which make students have to be familiar and master the paraphrasing ability to avoid plagiarism. The second reason was that EED of UMY provided the data needed by the researcher since there is a subject in which the lecturers scoring student's paraphrasing in EED of UMY. After all, EED of UMY was suitable with this research and provides the data needed, which is receptive vocabulary size test and paraphrasing score. Moreover, the researcher is one of the EED of UMY students. It makes the researcher easier to get the data of this research because the researcher knew the technical administration of EED of UMY.

Population and Sample of the Research

There are four populations' criteria of this research. First, they have to be students of EED of UMY in batch 2016. The main reason why the researcher choose batch 2016 was because the researcher would like to know the average of the freshest students in EED of UMY. Second, they have to join an Interpretive Reading and Argumentative Writing course class A or B. The reason why the researcher only chooses class A and B was because students' paraphrasing scores only available in both classes. Third, they must have a receptive vocabulary size score. Fourth, they must have a paraphrasing score. The data taken from administrative office of EED of UMY stated that there were 121 active students of

EED of UMY batch of 2016 who enrolled into three classes, 40 students in class A, 35 students in class B, and 46 students in class C.

The target of this research participants consisted of 75 students from class A and B. Unfortunately; there are 37 students who did not fulfill the criteria of this research participants. Those 27 students did not fulfill the criteria since they did not have paraphrasing score due to plagiarism and did not read the selected book as the source of paraphrase in the course Interpretive Reading and Argumentative Writing. Moreover, the other 10 students did not meet the criteria because they were absent and unable to complete receptive vocabulary size test score carried out by the researcher. It is therefore only 38 participants who fulfilled the criteria of this research participant. Even though the number of participants did not reach the target, 38 participants were enough on correlation design research. It is the same as Cohen, Manion and Morrison's (2011) statement who said that the minimal participant needed in statistical analysis research is 30 participants.

The sampling technique used in this research is total population sampling. Sugiyono (2007) stated that total population sampling is a sampling collection data technique in which the number of the sample is equal with the number of population. The reason why the researcher chooses this sampling technique was because the total of population consisted of less than 100 participants. Sugiyono (2007) mentioned that if the number of population is less than 100, all of participant should be a sample of the research. Since the total of participant of this research was 38 students, it can be said that total population sampling matches

this research. It also can be said that the population and sample of this research was 38 students of EED of UMY in batch 2016.

Instruments of the Research

The researcher used two instruments of this research including receptive vocabulary size test and paraphrasing score. The receptive vocabulary size test was done to measure students' receptive vocabulary size in English. The receptive vocabulary size test was taken from Nation and Beglar's (2007) receptive vocabulary size test at <http://www.victoria.ac.nz/lals/staff/paul-nation.aspx>.

There are two kinds of receptive vocabulary size test based on Nation and Beglar (2007), the 20,000 version and 14,000 versions. The 20,000 version is provided to the native speakers, while the 14,000 version is provided to the non-native speakers. That is why this research used Nation and Beglar's (2007) 14,000 receptive words vocabulary size test version which provided to the second language learners. For practical reasons, it is impossible to test all the words that the language learner might know, so the researcher has typically started with a large dictionary and then drawn a sample of words representing, say, 1 per cent (1 in 100) of the total dictionary entries (Read, 2000). That is why this test only consists of 140 multiple choice questions which represent 14,000 words. Each 10 words represent 1,000 words. Nation (1990) mentioned the steps of selecting words when making a test. First, include the nouns, verbs, adjectives, and adverb words and exclude words that cannot easily be tested like *a, the, of, be, etc* from the words list. Second, number the words and choose the words with arrangement number 10 then 110 and the multiple numbers. The instruction of receptive

vocabulary size test was the students should choose the closest meaning of the question words without seeing any dictionary or ask to other students. The four categories the researcher made for the vocabulary size test such as Very Good, Good, Poor, and Very Poor.

The paraphrasing score was taken in order to know in which category EED of UMY students' average paraphrasing score belong to. The paraphrasing score was taken from students' paraphrasing score at Interpretive Reading and Argumentative Writing course. There are also four categories the researcher made for the level of student's paraphrasing score, including Very Good, Good, Poor, and Very Poor.

Data Collection Methods

The receptive vocabulary size test was distributed to the students in their classrooms by the researcher. The first thing that researcher did before distributing the receptive vocabulary size test was searching for a lecturer who was willing to give their teaching time to the researcher. The researcher asked for approximately 30 minutes from the lecturer to distribute the test. After getting the permission, the researcher distributed the test to the students, gave instructions of the test and gave the students time to finish the test while the researcher waited until all participants were done the test.

To get the paraphrasing score, the researcher made a permission letter to the Interpretive Reading and Argumentative Writing lecturer to take the paraphrasing score data. The permission letter was signed by the researcher, the research supervisor, and the chief of English Education Department. After the

receptive vocabulary size data and the paraphrasing score data were collected, the researcher started to analyze the data.

Validity and Reliability

Before the receptive vocabulary size test was distributed to the students, the validity and reliability should be checked. Validity is an important key in order to get the effective instrument of the research since if a piece of research is invalid then it is worthless (Cohen et al, 2011). Nation and Beglar (2007) mentioned the information of the receptive vocabulary size test validity which was already through three steps. Those steps represent the validity of receptive vocabulary size test. Nation and Beglar (2007) wrote that their receptive vocabulary size test has been already through some piloting tests and expert judgments including; First, it has been applied by linguists who are the native speakers of English to individually read and give a suggestion of the test. Second, this test is already tested through the Range program in order to check the frequency levels of words used in the context and choice. Third, it is already through Rasch-based analysis under 200 students in Japan in which the result of this test was valid.

Moreover, before analyzing the data, the reliability of instrument should be checked. Cohen et al (2011) stated that reliability is “essentially a synonym for dependability, consistency and replicable over time, over instruments and over groups of respondents” (p.199). In order to recheck the reliability of the receptive vocabulary size test, SPSS 17 is used to calculate the reliability using Cronbach’s

Alpha coefficient. The category of reliability using Cronbach's Alpha coefficient mentioned by Cohen et al (2011) are:

Cronbach's Alpha Score	Category
>0.90	Very Highly Reliable
0.80-0.90	Highly Reliable
0.70-0.79	Reliable
0.60-0.69	Marginally/Minimally Reliable
<0.60	Unacceptably Low Reliability

Table 1: The Criteria of Reliability

After checking the reliability of receptive vocabulary size test in SPSS 17, the researcher found the result. The table below is the result of reliability test:

Cronbach's Alpha	N of Items
.873	140

Table 2: The Reliability Statistics

From the table above, it shows that the Cronbach's Alpha number of the receptive vocabulary size test was 0.873 which means this test is in Highly Reliable category.

Data Analysis

In order to answer the first question of the research, firstly the researcher input the receptive vocabulary size data through Microsoft Excel first. Then, the data were copied and analyzed through SPSS 17 to seek the mean, maximum,

minimum, range, and frequencies of the receptive vocabulary size data. The answer of the first question showed from the frequencies of the data. Then, the researcher matched the frequencies number in the range number of each category to know the student's receptive vocabulary size level. Since the score of each item is 100, so the maximum score of vocabulary size test is 14000, and the minimum score is 0. In order to make the interval of vocabulary size level category, the researcher used an interval formula from Supranto (2006):

$$c = \frac{\text{Maxvalue} - \text{Minvalue}}{N}$$

$$c = \frac{14000-0}{4} = \frac{14000}{4} = 3500$$

c = class width

Maxvalue = Maximal value

Minvalue = Minimal value

N = Number of classes

The formula shows the class width of the first research question is 3500.

Hence, the classification can be categorized as follows:

Score	Category
10.501 – 14.000	Very Good
7.001 – 10.500	Good
3.500 – 7000	Poor
0 – 3.500	Very Poor

Table 3. The Receptive Vocabulary Size Category

The steps of second research question analysis were the same as the steps above. The difference was in the interval of second research question. At Interpretive Reading and Argumentative Writing course, there is an assignment which asked the students to summarize a novel. This assignment gave maximum 10% scores of the total score of this course with the score criteria:

Aspects	Grade
1. Only the most relevant information is included.	2
2. Key points and minor details	2
3. Plot is presented in a comprehensible manner.	2
4. Logistics (grammar, vocabulary, punctuation, etc.)	2
5. Paraphrasing quality	2
Total	10%

Table 4. The Criteria of Summary Assignment Score

Thus, this research only used the paraphrasing quality aspects in which the maximum score is 2. It means the maximum score of paraphrasing is 2, and the minimum score is 0. It makes the interval class width of the second research question is different from the first one. The interval class width of second research question was:

$$c = \frac{\text{Maxvalue} - \text{Minvalue}}{N}$$

$$c = \frac{2-0}{4} = \frac{2}{4} = 0,5$$

c = class width

Maxvalue = Maximal value

Minvalue = Minimal value

N = Number of classes

The formula shows the class width of the first research question is 1,5.

Hence, the classification is as follows:

Score	Category
1,6 – 2	Very Good
1,1 – 1,5	Good
0,6 – 1,0	Poor
0 – 0,5	Very Poor

Table 5. The Paraphrasing Score Category

In order to answer the third research question, the researcher gathered the data of receptive vocabulary size test and paraphrasing score in SPSS 17, then see the normality, the linearity and the coefficient correlation between both variables. The coefficient correlation number matched the Cohen's et al (2011) effect size number to know whether there is any correlation between student's receptive vocabulary size level and their paraphrasing score.