

CHAPTER III

RESEARCH METHODOLOGY

3.1 Introduction

This chapter describes the research methodology which has been briefly introduced in chapter I. It also describes the procedures and process involved in the research. In detail, this chapter explains research design, data collection, research procedure, and data analysis.

3.2 Research Design

Experimental research was used in the research. According to Geoffrey (2006), “Experimental research is the only type of research that can test hypothesis to establish cause-effect relationship. It represents the strongest chain of reasoning about the links between variables”. The design used in the research was quasi-experimental design by using nonequivalent control group design since it is sometimes hard to randomly assign individual participants to groups. In nonequivalent control group design, the entire classroom, not individual students, are assigned to treatments (Geoffrey, 2006).

The formula is represented as follows:

$$\begin{array}{cccc} \mathbf{G1} & \mathbf{T1} & \mathbf{X} & \mathbf{T2} \\ \hline \mathbf{G2} & \mathbf{T1} & & \mathbf{T2} \end{array}$$

Notes:

G1 = Experimental Group

T2 = Posttest

G2 = Control Group

X = Treatment (*MUF* framework)

T1 = pretest

Two classes were taken as the investigated groups in this research. The first class is for the experimental group (G1) that will be treated (X) by using *M-U-F* framework. The second class is for control group (G2) which was not given treatment. The pretest was given before applying *M-U-F* framework as a treatment. Then, the posttest was given to both groups after the treatment to see whether or not the students make progress in their grammar ability.

The experimental study has at least one hypothesis to show an expected causal relationship between two variables. The experiment is employed to support or refute the hypothesis (Geoffrey, 2006). The hypothesis used in this research was null hypothesis and alternative hypothesis. The null is that there is no difference in mean adjustment level between those who use who receive *M-U-F* framework and those who do not. It means the experiment works if the null hypothesis is rejected. Whereas, the alternative hypothesis states that there is a difference between those who receive *M-U-F* framework and those who do not.

The hypotheses above were used to show an expected causal relationship between two variables that were independent variable and dependent variable. The characteristic of experimental research is the manipulation of independent variable.

According to Hatch and Farhady (1982) the selected, manipulated, and measured variable is called independent variable. In the research, the use of *M-U-F* framework served as independent variable. While, students' grammatical competence served as the dependent variable or the variable that was used to see the effect of independent variable.

3.2.1 Data Collection

3.2.1.1 Population and Sample

Johnson (1992:110) states that population is the entire group of entities or person to which the result of a study are intended to apply. The population of the research is 68 students of fifth grade of one of elementary school in Western Bandung. Two classes of fifth grade were assigned as a sample. Class 5A was an experimental group and class 5B was a control group. Each of the class consists of 24 students.

3.2.1.2 Data Collection Instruments

According to Fraenkel and Wallen (1990), an instrument is any kind of device that a researcher uses in collecting data. Test instrument and non test instrument were used in the research. Test instrument is in the form of multiple choice items which consist of 20 items. The multiple choice test was in the form of story; it tells a story from number one until number 20. It is based on the principle of *M-U-F* framework that students have to know the context of using the language. Therefore, in building the context, the test was arranged as a story. The test was constructed by the researcher

herself based on the syllabus of the fifth grade. It covers several topics namely *New Friends; My School; My House; Fruit & Vegetable; and Food and Drink* which covered grammar aspects such as *this is...; she is in (grade); preposition; singular/plural; measurements; can I help you?; verb do*. The test instrument was used in the pretest and posttest for both experimental and control groups.

In addition, video recording was the non test instrument that was used in the research. The researcher used video recording as an instrument for capturing clear images in classroom. Video recording was employed to find the strengths and challenges that the teacher faced in implementing *M-U-F* framework. Bottorff (1994 cited in Yu, 2007) mentioned that video recording has a function to record both verbal and non-verbal communication between the teacher and the students in details. It was used in every treatments of experimental and control group. The video recorder was placed in the open space in the corner of the back of the classroom in order to cover the teacher and the biggest number of the students at the widest range. The teachers' voices could be clearly captured. However, due to practical circumstances i.e. the size of the class, students' voices could not be adequately recorded especially when they were not directly involved with the teacher.

3.2.2 Research Procedure

3.2.2.1 Organizing Teaching Procedure

In conducting the research, the researcher acted as a teacher of experimental group. There were some steps taken in preparing the teaching process for experimental group. The first step was preparing the appropriate materials based on the syllabus of the fifth grade. Then, preparing the media that support the materials. The second step was organizing teaching procedure based on *M-U-F* framework. Finally, the last step was carried out the materials. *M-U-F* framework was applied both in making the materials and in organizing the procedure of the treatments. While in the control group, conventional teaching method was applied.

3.2.2.2 Administering Try-out Test

In order to produce a good test, the instrument should be tried out first. Try-out test was conducted to find out the validity, level of difficulty, discrimination, and reliability of the instrument. The try-out test was conducted to 33 students of fifth grade from different elementary school. The test consisted of 20 multiple choice items.

3.2.2.3 Conducting Treatment

The treatment was conducted in six meetings. The teacher carried out the materials related to grammar that had been arranged in the lesson plan. The materials and the procedure were arranged based on the principles of *M-U-F* framework. *Meaning*, *Use* and *Form* were the main principles in *M-U-F* framework. *Meaning* was

built by introducing the new language in a context through story, role play or puppet. It will help them to understand the meaning. *Use* was introduced by trying the language in purposive way through information gap or games. Students' attention to *Form* was introduced in the activity such as fill in the gap, worksheet or games. The details of the treatment activity were shown in the following table.

Table 3.1
Schedule of the study

No	Experimental Group (5A)	Control Group (5B)	Grammar Focus
	Date	Date	
1.	21 st July 2009	July 21 st 2009	-
2.	28 th July 2009	28 th July 2009	Hello, this is... She is in grade...
3.	4 th August 2009	4 th August 2009	- Preposition (near, between, beside, across) - Where is (....) ?
4.	11 th August 2009	11 th August 2009	- Preposition (near, between, beside, across) - It is (downstairs/upstairs)
5.	25 th August 2009	25 th August 2009	Singular & Plural - there are... - there is....
6.	1 st Sept 2009	1 st Sept 2009	- Measurements (a bottle of, a piece of, etc) - Can I help you? - I want
7.	8 th Sept 2009	8 th Sept 2009	Verb do Do you like....? Yes, I do. No, I don't
8.	6 th October 2009	6 th October 2009	-

3.2.2.4 Administering Pre-test and Post-test

A pre-test was administered in experimental group and control group in order to find out the students' mastery in grammar before the treatment. The pretest was in the form of multiple choice that consisted of 20 items. Then, the posttest was given after students received the treatment to find out whether or not the students make progress in their grammar ability. It was given to both experimental and control group in the end of program. The procedure and the instrument of post-test were similar to the pretest.

3.2.3 Data Analysis

After collecting the data, the data from pilot test, pretest, and posttest of experimental and control group were analyzed. The data from pilot test were analyzed to measure the validity, reliability, level of difficulty, and discrimination level of the instrument. While the data from pretest and posttest were used to analyze the independent and dependent *t*-test.

3.2.3.1 Data Analysis on the Pilot test

After conducting the pilot test, the data were analyzed to measure the validity, reliability, level of difficulty, and discrimination level of the instrument.

Validity Test

A valid instrument will have a high validity. Test validity is defined as the degree to which a test measures what it claims to be measuring (Brown, 1990: 101). The Pearson

Product Moment in SPSS 16 for windows was used to calculate the validity of each item of the test.

Table 3.2

R coefficient Correlation (Validity)

Raw score	Interpretation
0.800 – 1.00	Very High
0.600 – 0.800	High
0.400 – 0.600	Moderate
0.200 – 0.400	Low
0.00 – 0.200	Very Low

(Arikunto, 2007: 147)

Difficulty Test

Fulcher and Davidson (2007) defined the difficulty or item facility as the proportion of the test takers who answer an item correctly. The items should neither too difficult nor too easy for the population for whom the test has been designed. According to Henning (1987) as cited from Fulcher and Davidson (2007), the items considered to be ideal if the facility value is around 0.5, with an acceptable range being from around 0.3 to 0.7.

Discrimination

The index of discrimination is used to indicate how far a single test item can differentiate the upper group from the lower group of the class (Arikunto, 2003). According to Fulcher and Davidson (2007), the importance of discrimination is related

to an approach of scoring that assumes that the more correct answer is the more of the ability in question and the fewer correct answer is less of the ability in question. Point biserial correlation is the most commonly used method to calculate the item discrimination. It is used to measure the association between responses to any specific item and the score on the whole test. The data were analyzed with the formula as follows:

$$r_{pbi} = \frac{\bar{X}_p - \bar{X}_q}{S_x} \sqrt{pq}$$

Where

r_{pbi} = point biserial correlation

\bar{X}_p = mean score on the test for those who get the item correct

\bar{X}_q = mean score on the test for those who get the item incorrect

S_x = standard deviation of test scores

p = the proportion of test takers who get the item correct (facility value)

q = the proportion of test takers who get the item incorrect

Items with an r_{pbi} of 0.25 or greater are considered acceptable, while those with lower value would be rewritten or excluded from the test (Henning, 1987 cited from Fulcher and Davidson, 2007)

Reliability Test

A test must produce consistent result whenever it is administered again. It is the concept of reliability. According to Hatch &, Farhady 1982: p.244, reliability is the extent to which a test produces consistent result when administered under similar conditions.

KR-21 method was used to define reliability of the test in order to find out the stability of the test. The material was twenty items and tested to twenty four students of the fifth grade. Reliability was searched with the formula:

$$r_{pbi} = \frac{k}{k-1} \left[1 - \frac{x(k-\bar{x})}{ks^2} \right]$$

Where:

K = the number of items in the test

\bar{x} = the mean of the sample

s^2 = the variance of the sample

3.2.3.2 Data Analysis on the Pretest and Posttest

The normal distribution test and homogeneity of variance test were conducted before the data from pretest and posttest were analyzed to calculate the independent and dependent t-test.

Normal Distribution Test

Scores can be examined by t-test only if they are normally distributed. Kolmogorov-Smirnov test in SPSS 16 was used to calculate normal distribution. By using this test, the sample scores are compared with normally distributed set of scores with the same mean and standard deviation (Field, 2005).

According to Field (2005, p.93), if the test is non significant ($p > .05$) it tells us that the distribution of the sample is not significantly different from a normal distribution. However, if the test is significant ($p < .05$) then the distribution of the sample is significantly different from a normal distribution.

The Homogeneity of Variance test

Scores can be examined by t-test only if variance of the scores in the population is equal. Lavene test formula in SPSS 16 for windows was used to test the homogeneity of variance. If the significant value is more than the level of significance (0.05) the null hypothesis is accepted; the variance of control group and experimental group are homogeneous.

The Independent t-test

According to (Kranzler and Moursund, 1999) the independent t-test was used to determine whether the means of two groups of scores differ to a statistically significant degree. There are three assumptions that must be considered before conducting independent t-test. First, the participants must be different in each group. Second, the data has normal distribution. Third, the variance of the two groups is equal (Coolidge, 2000, p.143).

Then, to find out the differences between experimental and control group, the independent t-test in SPSS 16 was used. If the significant value is less than the level of significance (0.05) the null hypothesis is rejected.

The Calculation of Effect Size

Although the t-statistic is statistically significant, it does not mean the effect is important in practical term. The effect size is used to find out whether the effect is substantive or not (Field, 2005). According to Coolidge (2000) size effect refers to the effect of the influence of independent variable upon the dependent variable. If there is a

large difference between the two groups' means, the effect size will be large. It means the treatment works well.

The formula of effect size is:

$$r = \sqrt{\frac{t^2}{t+df}}$$

Where:

r = effect size

t = t value from the calculation of independent t -test

$df = N_1 + N_2 - 2$

After getting the r value, match the score with the following scale to interpret the effect size.

Table 3.3

Effect Size Value

Effect Size	r value
Small	.100
Medium	.243
Large	.371

(Coolidge, 2000: p.151)

The Dependent t-test

The assumptions of dependent t -test were similar with the independent t -test. First, the data has normal distribution. Second, the variance of the two groups is equal. Third, the participants are from the same groups (Coolidge, 2000). The dependent t -test was used

to find out whether or not the difference of pretest and posttest means of experimental group is significant. If the significant value is less than the level of significance (0.05) the null hypothesis is rejected. On the other hand, if the significant value is more than the level of significance (0.05), so the null hypothesis is accepted.

3.2.3.4 Data Analysis of Video Recording

The data from video recording were transcribed. The data from experimental group were analyzed from the first treatment until the six treatments. The strengths and challenges that the teacher faced in implementing *M-U-F* framework were noted from the video. Then, those strengths and challenges were categorized into *Meaning*, *Use* and *Form*.