

III. RESEARCH METHODS

This part discusses the design of this research and how to collect the data from those samples. The writer encloses the data collecting technique and the procedure of this research. The writer also gives the scoring system and how the data are analyzed.

3.1 Research Design

This research was intended to find out whether there was any significant difference of students' vocabulary achievement between those who were taught total physical response and those who were taught through audio lingual method. The quantitative research was conducted to gain the objective of this research. In this research, there are three classes and two of them are given a treatment and one class as a tryout class. The two group pre-test post-test design is employed. The two classes, experimental class I and experimental class II are compared in order to find out the significant difference on vocabulary achievement. In this case, the experimental class I is students who are taught through Total Physical Response method and the experimental class II is the students who are taught through Audio Lingual method.

The design of the research was presented as follow:

G1 = T1X1T2

G2 = T1X2T2

Notes:

G1 : The experimental class I

G2: The experimental class II

X 1: Treatment 1 (Total physical Response)

X 2: Treatment 2 (Audio Lingual Method)

T1: Pre-test

T2: Post-test

(Hatch and Farhady, 1982:20)

The research was conducted in six meetings with presentation as follows. The first meeting was for tryout test. The second meeting was for pre-test. The third, fourth, and fifth meetings were for the treatment. The last meeting was for post-test.

3.2 The Population and Sample

The population of this research is the second grade students of SLTP N 22 Bandar Lampung in even semester in the year 2010/2011. One class was a tryout class.

There were two classes that were used as a sample in this research. These two classes were determined as the experimental class 1 and experimental class 2.

They were purposively sampling as a consideration from the teacher in the school that they had almost similar ability. The experimental class 1 was taught through

total physical response and the experimental class 2 was taught through audio lingual method.

3.3 Variables

Based on the problem of the research, the variables can be defined as follows:

1. the dependent variable is the students' achievement on vocabulary
2. independent variable is the treatments used in the research (ALM and TPR)

3.4 The Data Collecting Technique

In collecting the data, the researcher used test as the instrument. There are two kinds of test, pre-test and post-test. Pre-test are given in order to measure the students' achievement in vocabulary before the treatment, and post-test are given after presenting the treatment in order to know the achievement of vocabulary. The test will be created based on KTSP curriculum for the Junior High School students. The test is multiple choices.

3.5 The Procedure of Collecting the Data

There are some procedures that are applied for the research, they are:

1. Determining the population and the sample.

The researcher took two classes to determine the experimental class I and experimental class II.

2. Administering try-out.

The try out administered to know the quality of the test.

3. Administering the pre-test.

The researcher and the teacher administered the pre-test on both class experimental class I and experimental class II.

4. Conducting treatment.

The researcher conducted the treatment.

5. Administering the post-test.

The researcher and the teacher administered the post-test to class, experimental class I and experimental class II.

6. Scoring the students' work.

The researcher scored the learners' work in order to get the data.

7. Analyzing the data.

After collecting the data, the researcher analyzed the data.

8. Testing hypothesis.

After analyzing the data, the researcher tested the hypothesis.

3.5 Instrument of the Research

This research instrument was tested in the form of objective test. In this research, the researcher administered three tests: try out test, pre test and post test. Try out test is given to know how quality of the test which is used as the instrument of the research. The pre test is given in order to know the students' vocabulary achievement before treatment. The post test is given in order to know the students' vocabulary achievement after treatment.

The form of try out test, pre test and the post test is objective. The total number of the items of the try out test is 50 items, and the total number or the items of the pre test and post test is determined from the result of the try out test. The validity of the instrument concerned with the content and construct validity. Test items are multiple choices.

3.6 Criteria of Good Test

In this research, to prove whether the test had good quality, it is tried out first. The test is said to have a good quality if it has a good validity, reliability, level of difficulty and discrimination power. The try out test has given to the students' to know how the quality of the test which is used as the instrument of the research. The try out test will be given to another class that is not included in the sample. The data which has been gained will be analyzed to judge the level of difficulty, discrimination power, validity, and the reliability of the test.

3.6.1 Validity

A test can be said to be valid if the test measures what is intended to measure (Shoahamy: 1985, 74). It means that the test can be said valid if the test measures to measure whether the test has good validity; the researcher analyzed the test from content validity and construct validity.

Content validity is concerned with whether or not the content of the test is sufficiently representative and comprehensive for the best to be valid measure it is supposed to measure. In the content validity, the material is given is suitable with

the curriculum used. The materials of this research in this case are preposition of place, imperative sentences, and daily activities.

The researcher will use the table of specification to check content validity of the test items. Gronlund (1981:101) stated that table of specification are to illustrate how such a table is used to check on content validity. The percentage in the table indicates the relative degree of emphasis each content area and each instructional objective is to be given in the test. In selecting the test, a table of specification could help us to determine which test is most relevant to our particular situation and also necessary to check whether test items have a good content validity.

Test is a good reflection of what has been taught and of knowledge, which the teacher wants his students to know. Content validity can be examined from the table of specification. If the table represents the material that the tester wanted to test, then it is a valid test from point a view (Shohamy, 1985: 74). The table of specification as follows:

Table 1. Table of Specification of Try Out Test

Content	Aspect	Items	Total	Percent
Vocabulary	Noun	1.,2.,3.,6.,8.,10.,14.,15.,18.,20.,33., 35.,36.,	13	26%
	Verb	5.,7.,9.,12.,16.,17.,23.,24.,25.,27., 28.,29.,30.,31.,32.,34.,38.,39.,41., 42.,43.,44.,45.,46.,47.,48.,49.,50.	28	56%
	Adjective	19.,22.,26.,37.,40.	5	10%
	Preposition	4,11,13,21	4	8%
Total		50	50	100%

3.6.2 The Reliability

Shohamy (1985: 70) states that “reliability refers to the extent to which is consistent in its score, and it gives one an indication of how accurate the test score is”. To find the reliability of this test the writer used Split Half Method, which has two steps, they are:

1. First, using Pearson Product Moment Correlation, we should find the correlation between odd and even number of the items.

$$r_{xy} = \frac{\sum xy}{\sqrt{(\sum x^2)(\sum y^2)}}$$

Where:

r_{xy} : coefficient of reliability between odd number and even number items

$\sum x^2$: total square of odd number items

$\sum y^2$: total square of even number items.

$\sum xy$: total score of odd and even number items.

2. Second, after we get the coefficient Correlation between odd and even number, we continue to put them into the reliability formula.

The reliability formula is below:

$$R_{11} = \frac{2 \times r_{xy}}{1 + r_{xy}}$$

Where:

R_{11} : reliability coefficient

R_{xy} : coefficient between odd and even number

The criteria of reliability

0.80 – 1.00 : very high reliability

0.60 – 0.79 : high reliability

0.40 – 0.59 : medium reliability

0.21– 0.39 : low reliability

0.00 – 0.20 : very low reliability

(Arikunto, 1998:260)

In this research, the result of reliability was 0.95 (see appendix 3). It can be concluded that the test has very high reliability in which the criteria is in the range 0.80-1.00. It indicated that the instrument would produce consistent result when administered under similar condition, to the same participant and in different time (Hatch and Farhady, 1982:286). So, it can be stated that the test has fulfilled the criteria of reliability. In other word, the test was reliable.

3. Level of Difficulty

In determining the level of difficulty of the tests, the writer uses the following formula:

$$LD = \frac{U + L}{N}$$

Note:

LD : level of difficulty

U : Upper group students who answer correctly

L : Lower group students who answer correctly

N : the total number of the students

The criteria are:

0.00-0.30 = difficult

0.31-0.70 = average

0.71-1.00 = easy

(Arikunto, 1997 :121)

The writer will use the item with if LD = average and others will be revised.

Based on the tryout test result related to the criteria there were 24 average items, the rest (26 items) was difficult items and there were no easy items. Some items which were difficult were dropped or revised, meanwhile for the average items were administered for pretest and posttest. The results of difficulty level of try out test are shown in Appendix 2.

4. Discrimination Power

To estimate the discrimination of power of the tests, the writer will use this following formula:

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

DP = discrimination power

U = the number of upper group students who answer correctly

L = the number of lower group students who answer correctly

N = the total number of students

The criteria are:

0.00 – 0.20 = poor

0.21-0.40 = satisfactory

0.41 – 0.70 = good

0.71 – 1.00 = excellent

(Arikunto, 1997 :121)

Based on the try out test result related to those criteria there were 16 poor items, eight items had negative discrimination power, 21 items were satisfactory, and the rest (5 items) were good items. Negative discrimination items were dropped while the good items and satisfactory were administered. Some of items, which have poor discrimination, were revised and used as pretest and posttest. The total items that were administered for pretest and posttest were 30 items (1,4,5,6,7,8,10,14, 16,18,19,20,21,22,23,24,25,27,28,30,32,33,35,36,37,38,39,44,45,50).

F. Data Analysis

After collecting the data, it was computed through drawing conclusion from tabulated result of the pretest and posttest by using SPSS 17. In doing so, the researcher had analyzed the data statistically by administering random test, normal distribution test, homogeneity test of variance and hypothesis test.

1. Random Test

To see whether the data in experimental class 1 and experimental class 2 were random, random test was computed. Here, the writer used Runs test.

The hypothesis for random test was:

H = the data is random

The data were random if it filled the criterion. In this research, the criterion for the hypothesis was:

The hypothesis was accepted if $\text{sign} > \alpha$. In this case, researcher used level of significant of 0.05.

Random test from the pretest in the experimental class 1 showed the significance value was 0.862 (see appendix 10). Seeing the result, it can be concluded that the data was random since $\text{Sign} > \alpha$ ($0.862 > 0.05$). Meanwhile, the result of random test of posttest in the experimental class 1 showed the number of 0.94 (see appendix 10). The value also $> \alpha$, in which $0.94 > 0.05$. It could be stated that the data was random.

The analysis of random test of pretest in the experimental class 2 showed the value of 0.27 (see appendix 11). Since it more than α , it can be concluded that the data was random. The random test of posttest showed the value was 0.74 (see appendix 11). Because the value was higher than α ($0.74 > 0.05$), the data from this class was also determined random.

In short, the data from the experimental class 1 and the experimental class 2 showed the value higher than alpha ($\text{Sign} > \alpha$), which meant that the data from both class were random.

2. Normal Distribution Test

The normality testing was held twice. These tests were employed to know whether the data of pretest and posttest were normally distributed or not. The normality of pretest was assumed if the significance was greater than 0.05. The result of the normality testing can be seen in table 2 below:

Table 2. Normality Testing

	Kolmogrov- Smirnov Z	
	N	Sig. (2-tailed)
Pretest VII A	30	0.562
Posttest VIIA	30	0.256
Pretest VIIC	32	0.243
Posttest VIIC	32	0.673

Table 2 inferred that the significance of pretest in the experimental class I (VIIA) was 0.562. Since the significance was higher than 0.05, it could be concluded that the data of the pretest in the experimental class I was normally distributed. While in the experimental class II (VIIC), the significance was 0.243. The significance was more than 0.05 and it meant that the data of pretest in the experimental class II was also normally distributed. These indicated that the data of pretest to both classes were normally distributed.

Moreover, Table 2 also showed that the data of the posttest in the experimental class I was normally distributed since the significance was 0.256. The significance was higher than 0.05, it could be concluded that the data was normally distributed. In the experimental class II, the significance was 0.673, which more than 0.05. So, the data in the experimental class II was also distributed normally.

3. Homogeneity Test of Variance

To find out whether the data from the two groups were met the criteria of the equality of variance, the researcher used homogeneity test.

In this research, the hypothesis for homogeneity test was:

H = the data is homogenous

Criterion for the hypothesis was:

The hypothesis was accepted if $\text{Sign} > \alpha$. In this case, researcher used level of significant of 0.05.

From the result of homogeneity test for the pretest scores both in the experimental class 1 and the experimental class 2 was 0.696 (see appendix 15). It showed that $\text{Sign} > \alpha$ ($0.696 > 0.05$). Therefore the hypothesis was accepted. In short, both classes were equal since the Sign was same and the research could be conducted to both classes.

G. Hypothesis Testing

This test was calculated to find out whether the hypothesis proposed by the researcher is accepted or not. The hypotheses of research were as follow:

1. There is a significant difference of vocabulary achievement of students who are taught through audio lingual method and those who are taught through total physical response.

2. The method of total physical response is better than audio lingual method in teaching vocabulary in helping the students increase their vocabulary achievement.

The hypothesis was statistically analyzed by using Independent Group T-Test. It was used to draw conclusion at the level of 0.05 (p,0.05).