CHAPTER III

RESEARCH METHODOLOGY

A. Design of the Research

This research was an experimental research. Based on what Cresswell said that we use experimental research when we want to establish possible cause and effect between our independent and dependent variables (Cresswell, 2008:331). The design of this research was a Quasi experimental research, to know the effect of using World Cafe Strategy on speaking ability.

This research design used two groups. The first group was as experimental group treated by using World Cafe Strategy. The second group was as control group which was treated without using World Cafe Strategy. Before doing treatment, the writer administered pre-test to all samples. After doing the treatment, writer administered post-test to obtain the students' speaking ability in classroom discussion.

Table III.1
The Research Design

Group	Pre-test	Treatment	Post-test
A	X_1	X	Y_2
B	X_1	-	Y_2

Note:

a. A = Experimental Class
b. B = Control Group
c. X_1 = Pre-test for experimental and control class
d. Y_2 = Post-test for experimental and control class
e. X = Treatment for experimental group by using World Cafe
f. - = The using of traditional strategy

34

B. The Location and the time of the Research

The research was conducted to the eleventh grade students at State senior High School 3 Bagan Sinembah of Rokan Hilir Regency. This research was conducted o from July to August 2016. The study of experimental group was conducted for eight meetings in which one meeting were pre-test, six meetings as treatment and one meeting was post-test. Then, the study of control group was conducted in two meetings in which one meeting was pre-test and one meeting was post-test.

C. The Subject and the Object of Research

1. The subject of the research

   The subject of the research was the eleventh grade students at State Senior High School 3 Bagan Sinembah of Rokan Hilir Regency in (2016-2017) academic year.

2. The object of the research

   The object of the research was the effect of using World Cafe Strategy on students' speaking ability.

D. The Population and the Sample of the Research

1. The Population of the Research

The population of this research was the eleventh grade students at State Senior High School 3 Bagan Sinembah of Rokan Hilir Regency. The spesification of population can bee seen in the table III.2 below:

Table III.2
The Total Population of the Eleventh Grade Students
at SMAN 3 Bagan Sinembah

No.	Class	Total Students
1	XI IPA 1	30
2	X1 IPA 2	30
3	XI IPA 3	33
4	XI IPS 1	32
	Total	125

2. The Sample of The Research

Based on the total population above, the writer took sample by
using cluster sampling. According to Gay, *et al* (2000: 129), cluster
sampling randomly selects groups, not individuals. Therefore, the
writer selected two groups of students to be sample in this research. It
was the students of XI IPA 1 as an experimental class and XI IPA 2 as
a control class.

Table III.3
The Total Sample of the Eleventh Grade Students
at SMAN 3 Bagan Sinembah

No.	Class	Total Students
1	XI IPA 1	30
2	XI IPA 2	30
	Total	60

E. Validity and Reliability of The Instrument

1. The Validity of the Test

Validity refers to the appropriatness of a given test or any of its
components parts as a measure of what it is purposed to measure. A test is
said to be valid if it measures what to be measured. According to Porte

(2002: 232-233), there are several types of validity namely; face validity, content validity, external validity, internal validity, and construct validity.

Face validity relates to content validity but assesses informally and/or intuitively whether the instrument appears to measure what it purposed to measure. Content validity considers formally the extent to which a particular instrument measures accurately what it is claimed to measure. A group of expert would normally decide on this, focusing on the instrument's representativeness and comprehensiveness. External validity is of little value unless it has been preceded by adequate address of internal validity concern, which give us confidence in the basic descriptive conclusion drawn from the data themselves. Internal validity is the extent to which the result of the study can be put down to the treatment applied rather than to the design of the study. It also reflects on the degree to which sound conclusion can be drawn about the result of the study. Construct validity describes the extent to which a particular instrument measures accurately construct of interest that have been obtained theoretically.

The writer concluded that this research belongs to the content validity, because the test reflects to what the student have learned the content of the curriculum. And also Gay (2000: 164) stated that there is no formula used to calculate the content validity and there is no way how to express it quantitatively. So, it means tests of content validity were given

based on material they had learned. The material of the test took from the textbook used by eleventh grade students at state Senior High School 3 Bagan Sinembah of Rokan Hilir Regency.

2. The Reliability of the Test

According to Gay, reliability is the degree to which a test consistently measures whatever it is measuring (Gay, 2000: 169). It is reflected in obtaining how far the test or instrument test is able to measure the same subject on different occasions indicating the similar result. In short, the characteristic of reliability is sometimes termed consistency. It is clear that reliability is used to measure the quality of the test scores and the consistency of the test.

According to Shohamy (1985), there are five types of reliability. They are test retest, parallel forms, internal consistency, inter rater and intra rater. In this research, to know the reliability of the speaking test, the writer used inter rater reliability because the writer had two raters in order to score the students' speaking ability. Gay says that inter judge reliability can be obtained by having two (more) judges independently score to be compared to the score of both judges. Then, the scores of rater 1 were correlated with the scores of rater 2. The higher correlation, the higher the inter judge reliability will be. The following table describes the correlation

between score of rater 1 and rater 2 by using pearson product moment correlation formula through SPSS 16.0 version.

Table III.4
Correlations

		Rater1	Rater2
Rater 1	Pearson Correlation	1	.640[**]
	Sig. (2-tailed)		0
	N	30	30
Rater 2	Pearson Correlation	.640[**]	1
	Sig. (2-tailed)	0	
	N	30	30

**. Correlation is significant at the 0.01 level (2-tailed).

From the output above, it can be seen that r_o ($r_{obtained}$) is 0.640 will be correlated to r_t (r_{table}). It is necessary to find the df (degree of freedom).

Formula: (df=N-nr)   df= 60-2=58
df: degree of freedom
N: Number of cases
Nr: number of correlated variable

(Sudijono(2012:194))

The writer took df= 58 to be correlated either at level 5% or 1%. At level 5%, r_{table} is 0.250, while at level 1% r_{table} is 0.325. Thus, the $r_{obtained}$ is obtained higher than r_{table}, either at level 5% or 1%. So, the researcher

concluded that there is a significant correlation between score given by rater 1 and score given by rater 2. In the other words, the speaking test is reliable.

F. Technique of the Data Collection

1. Observation

According to Arikunto (2006:156), observation is an activity that is concerned on some objects by using the five senses. Observation is the way to get some data, by observing the object of the research. In this research, observation was used to collect data on the application world cafe strategy in teaching process. While the writer did the treatment, one person (teacher) contributed as observer. She followed the treatment process and check whether the observation list ran or not. The observation list was collected and analyzed in order to consider how far the implementation of the treatment in the classroom.

Table III.5

Observation Checklist

No	Indicators of World Cafe Strategy	Yes	No
1.	Teacher teaches the students about today's topic.		
2.	The teacher asks the students form random groups of five.		
3.	Teacher sets the classroom such on world cafe model.		
4.	Teacher asks the students to name their group.		
5.	The teacher gives the students a topic to discuss.		
6.	Every 5 minutes, the teacher asks the students to randomly move to new coffee shop.		
7.	The teacher monitors the process that will be continued for several movements.		
8.	The teacher asks the students to performance individually in front of the class.		
	TOTAL		

2. Oral Presentation Test

In order to get the data needed by writer, the writer used a test. Brown (2003) stated that test is a method of measuring of students' ability, knowledge, and performance. Based on the statement above, the writer used an oral presentation test. This test was given to collect the data to find out students' speaking ability based on speaking components, such as pronunciation, grammar, vocabulary, fluency, and comprehension.

In this research, writer gave pre-test and post-test to every group. The test is explained as follows:

a. Pre-test

Pre-test was given to measure students' speaking ability before they were taught by using world cafe strategy. This test was given for both experimental and control class.

b. Post-test

Post-test was given to measure students' speaking ability for both experimental and control class after the experimental class had already been treated by using world cafe strategy.

G. Technique of the Data Analysis

1. Normality of the Data

Before analyzing the data by using T-test formula, the writer had to find out the normality test of the data. In this research, the researcher used lilliefors through SPSS.16 version. In analyzing the normality of the data, the writer used the students' post test score of experimental and control class. Analysis:

If the probability > 0.05 Ho is accepted

If the probability < 0.05 Ho is rejected

Table III.6
Test of Normality

	Kolmogorov-Smirnov[a]		
	Statistic	Df	Sig.
Post test experiment	0.127	30	.200[*]
Post test control	0.111	30	.200[*]

a. Lilliefors Significance Correction

*. This is a lower bound of the true significance.

Hypothesis:

H_0 (Null Hypothesis) : Data are normally distributed

H_a (Alternative Hypothesis) : Data are abnormally distributed

Testing Criteria:

If probably (sig) > 0.05, H_0 is Accepted

If probably (sig) < 0.05, H_0 is Rejected

According to Priyatno (2012: 36), If the "Sig" column of either test is higher than 0.05, the data are normally distributed. From the table IV.15 above, the significant value of post-test experimental and control class were 0.200 and 0.200. Because of *sig*> 0.05 (0.200 >0.05) and (0.200), the initial data of experimental and control class were normally distributed. Therefore, the researcher used independent sample T-test.

2. Homogeneity of the Data

According to Siregar (2013:167), the purpose of homogeneity test is to know whether the object of the research has the same variance or not. The method used in this test was comparing the biggest variance with the smallest one. The writer used students' pre-test scores of experimental and control class to analyze the homogeneity of the data.

Furthermore, in order to know whether the object researcher had the same variance or not, the writer previously described the homogeneity analysis as follows:

Table III.7
Test Homogeneity of Variances

		Levene Statistic	df1	df2	Sig.
Pre experiment And Pre control	Based on Mean	2.393	1	58	0.127
	Based on Median	1.822	1	58	0.182
	Based on Median and with adjusted df	1.822	1	51.398	0.183
	Based on trimmed mean	2.229	1	58	0.141

Based on the table above, the probability (sig) based on trimmed mean was 0.141. It was higher than 0.05 (0.141 > 0.05). It can be concluded that data were homogenous.

3. Analysis of the Data

In analyzing the data, the writer used students' post test score in experimental and control class. This score was analyzed statistically. In this research the writer used these formulas;

a. Independent sample T-test

To find out whether or not there is a significant difference between two or more variables that can be analyzed by using independent sample t-test (Hartono, 2008:146). The different mean was analyzed by using T-test (independent sample t-test). It was calculated by using SPSS.16 version.

The T-table was employed to see whether or not there is a significant difference among the mean scores both of experimental and control class. Statistical hypothesis:

1. H_0 : $t_0 < t - table$
2. H_a : $t_0 > t - table$

b. Effect Size

According to Pallant (2005:199), effect size is the strength of the difference between groups or the influence of independent variable. There are a number of different effect size statistic, the most commonly used being eta squared. Eta squared can range from 0 to 1 and represents the proportion of variance in the

independent variable that is explained by the independent (group) variable. The formula for eta squared is as follows:

$$\text{Eta squared}^2 = \frac{t^2}{t2 + (N1 + N2 - 2)}$$

Where :
$t^2 = t_o$
N_1 = Number of students

c. **Assessment**

In this research, the speaking was evaluated by concerning five components that should be considered in giving students' speaking ability score. According to Hughes (2003:131), there are some components that should be considered in giving students' speaking ability score: they are accent, grammatical, vocabulary, fluency, and comprehension. The students' speaking ability was measured by using oral language scoring rubric. So, Hughes desribed the rating as follows:

Table III.8
Speaking Assessment

a. **Accent**

Score	Requirement
1	Pronunciation frequently unintelligible.
2	Frequent gross error and a very heavy accent make understanding difficult, require frequently repetition.
3	"Foreign accent" requires concentrated listening, and miss pronunciation lead to occasional misunderstanding and apparent errors in grammar of vocabulary.
4	Marked "foreign accent" and occasional mispronunciations which do not interfere with understanding.
5	No conspicuous, miss pronunciation, but would not be taken for a native speaker.
6	Native pronunciation, with no trace of "foreign accents'".

b. **Grammar**

Score	Requirement
1	Grammar almost entirely inaccurate except in stock phrase.
2	Constant errors showing control of view major patterns and frequently preventing communication.
3	Frequent errors showing some major patterns uncontrolled and causing occasional irritation and misunderstanding.
4	Occasional errors showing imperfect control of some patterns but no weakness that causes misunderstanding.
5	Few errors, with no patterns of failure.
6	No more than two errors during the interview.

c. **Vocabulary**

Score	Requirement
1	Vocabulary inadequate for even the simple conversation.
2	Vocabulary limited to basic personal and survival areas (time, food, transportation, family, etc.).
3	Choice of words sometimes inaccurate, limitations of vocabulary prevent discussion of some common professional and social topics.
4	Professional vocabulary adequate to discuss special interest; general vocabulary permits discussion of any non-technical subject with some circumlocutions.
5	Professional vocabulary broad and precise; general vocabulary adequate to cope with complex practical problems and varied social situations.
6	Vocabulary apparently as accurate and extensive as that of an educated native speaker.

d. **Fluency**

Score	Requirement
1	Speech is so halting and fragmentary that conversation is virtually impossible.
2	Speech is very slow and uneven except for short or routine sentences.
3	Speech is frequently hesitant and jerky; sentences may be left uncompleted.

4	Speech is occasionally hesitant, with some unevenness caused by rephrasing and groping for words.
5	Speech is effortless and smooth, but perceptively non-native in speed and evenness.
6	Speech on all professional and general topics as effortless and smooth as a native speaker.

e. **Comprehension**

Score	Requirement
1	Understand too little for the simplest type of conversation.
2	Understand only slow, very simple speech on common social and touristic topics; requires constant repetition and rephrasing.
3	Understand careful, somewhat simplifies speech when engaged in a dialogue, but may require considerable repetition and rephrasing.
4	Understand quite well normal educated speech when engaged in a dialogue, but requires occasional repetition and rephrasing.
5	Understand everything in normal educated conversation except for very colloquial or low frequency items, or exceptionally rapid or slurred speech.
6	Understand everything in both formal and colloquial speech to be expected of an educated native speaker.

Each component had score or level. Each component had 20 the highest score and the total of all components has 100. The specification of the test is as follow:

Table III.9
The Specification of the Test

No	Speaking skill	The highest score
1	Accent	20
2	Grammatical	20
3	Vocabulary	20
4	Fluency	20
5	Comprehension	20
	Total	100

Table III.10

Classification of Students' Scores

No	Score	Categories
1	80-100	Excellent
2	66-79	Good
3	56-65	Average
4	40-55	Poor
5	30-39	Fail
Total		

(Arikunto, 2008:245)

Next, to know the students' score of speaking ability in group activity higher or lower, the writer used the score 70 based on the minimum of passing grade score (KKM) at Senior High School 3 Bagan Sinembah. Thus, the students who got score ≥ 70 they passed the minimum of passing grade score (KKM), while the students got < 70, they did not pass the minimum of passing grade score (KKM).