

## **CHAPTER III RESEARCH METHOD**

In order to make the points of method clear, this chapter presents research design, population and sample, data collecting technique, validity and reliability, data treatment, research procedure, criteria for evaluating students' speaking achievement, data analysis, and hypothesis testing.

### **3.1 Research Design**

This research was quantitative in nature because the major data were quantitative information, the students' scores of speaking skill. This research was done by using one group pretest-posttest design. The research were investigated (1) whether there is a difference of the students' speaking achievement before and after being taught through Information Gap and (2) there is significant increase of the students' speaking achievement after being taught through Information Gap technique at the second grade of SMA Negeri 3 Metro. Then, the means (average scores) of both pre-test and post-test were compared to find out the difference and significant increase before and after the treatments. This design is easy and useful way of getting preliminary information on the research question (Hatch and Farhady, 1982: 20).

This study uses one class as experimental class using simple random probability sampling, which is selected randomly by using lottery. This class had both pretest-posttest and three treatments.

The research design is represented as follow:

T1 X T2

Where:

- T1 : Pre-test: it is a test aims to see the students' ability (in this research, the researcher focus on speaking ability) before they are given such kind of treatment (Information Gap).
- T2 : Post-test: it is a test to see the students' ability of a certain skill (speaking ability) after they have given treatment (Information Gap).
- X : Treatment: in this research, the teacher treats the students by applying Information Gap in teaching speaking.

(Hatch and Farhady, 1982:20 in Setiyadi 2001:44)

Firstly, the researcher administered a pre-test to the students to identify their achievement of speaking skill before applying the technique. Then, the students were given three treatments by using Information Gap.

Eventually, a post-test was administered to identify students' speaking achievement after being taught by using Information Gap. If the average score of the pre-test (mean) was higher than the average score (mean) of the post-test, it indicated that Information Gap can not be used to increase students' speaking achievement. However, if the average score (mean) of the post-test was higher than the average score of the pre-test (mean), it showed that Information Gap can be used to increase students' speaking achievement.

### **3.2 Population and Sample**

The population of this research was the 2<sup>nd</sup> grade students of SMA Negeri 3 Metro period of 2011/2012. There were 8 classes in the 2<sup>nd</sup> grade of SMA Negeri 3 Metro and 36 students for each class (i.e. XI IPA 1-XI IPA 4 and XI IPS 1-XI IPS 4). Their ages range from 16-17 year old. The sample was one class as experimental class, which is selected by using simple random probability sampling.

The class was selected randomly by using lottery, lottery was the technique to choose the sample by using folded paper which is put into glass. Each paper represented one class, so there were six folded paper, and the researcher took one folded paper with closed eyes. Since the 2<sup>nd</sup> grade in SMA Negeri 3 Metro was not stratified class, there was no priority class and it was assumed that all class has similar English proficiency. It was applied based on the consideration that every class at the second grade had the same chance to be chosen and in order to avoid the subjectivity in the research (Setiyadi, 2006: 39). The experimental class has pre-test, post-test, and three treatments.

### **3.3 Data Collecting Technique**

In collecting the data, this research used speaking tests as instrument (the same tests for both pre-test and post-test) in collecting the data.

The researcher administrated pre-test in 90 minutes. The purpose of this test was to know the students' achievement in speaking before the treatment. In this test, the students were divided into pairs or groups. The materials were about invitation, opinion, and permission. Then, students made conversation based on the topic provided and their own knowledge. The test would be done in pair. This kind of activity stimulated the interaction among students. Then, the researcher recorded the students' conversation in cassette and checked the result by using J.B. Heaton table criteria of speaking ability.

The post-test done in 90 minutes. This was done in order to see the students' speaking achievement after three times of treatments (Information Gap). Similar with pre-test above, the students were divided into pairs or groups. The materials were about invitation, opinion, and permission. Then, students made conversation based on the topic provided and their own

knowledge. The test would be done in pair. This kind of activity stimulated the interaction among students. This kind of activity stimulated the interaction among students. Then, the researcher recorded the students' conversation in cassette and checked the result by using J.B. Heaton table criteria of speaking ability.

### **3.4 Validity and Reliability**

#### **3.4.1 Validity**

A test is said to be valid if it measures accurately what is intended to measure. There are some types of validity; content validity, constructs validity, and face validity (Hughes, 1989:22).

The validity of the test of this research related to:

##### **3.4.1.1 Face Validity**

To get face validity, the instructions and the directions of the speaking test examined by advisors and English teachers until the test looks right and understandable. Furthermore, a test is said to have face validity if it looks as if it measures what it is supposed to measure (Hughes, 1989:27).

##### **3.4.1.2 Construct Validity**

A test, part of test, or a testing technique is said to have construct validity it can be demonstrated that it measures just the ability which it is supposed to measure (Hughes, 1989:26). In this research, the researcher measured speaking skill referring to the aspects of speaking (pronunciation, fluency, and comprehensible).

##### **3.4.1.3 Content Validity**

While the content validity means that the test may be said to have content validity if the test reflect such an analysis of mastery of a specific skills or the content of a particular course of study (Harris, 1969:19). It means that the test has to reflect what has been taught and of the

knowledge that the researcher wants her students to know. Here, the researcher correlated the test with the newest syllabus and curriculum for senior high school.

### 3.4.2 Reliability

Reliability is a measure of accuracy, consistency, dependability, or fairness of scores resulting from administration of particular examination. To ensure the reliability of the speaking score and to avoid subjectivity of the researcher, inter-rater reliability will be used in this research. This reliability is used when test score were independently estimated by two or more judges or raters.

The first rater was the researcher and the second rater was the English class teacher. All of them discussed and put in mind of the speaking criteria in order to obtain the reliable result of the test.

$$IR1 \longrightarrow IR2$$

Where:

IR1 : First rater

IR2 : Second rater

To have the reliability of the test, this research used Rank-Difference Method whose formula is:

$$\rho = 1 - \frac{6 \cdot \sum D^2}{N(N^2 - 1)}$$

Where:

$\rho$  : Rank-Difference

Rank : Each student's performance on each of the two tests  
(In case of tied ranks, average the ranks)

$D^2$  : Square of Differences

$\sum D$  : The sum difference between each pair of ranks

(Hatch and Farhady, 1982: 222)

In this case, to interpret the reliability of the tests, the coefficient of rank correlation was considered through the standard criteria bellows:

- 0.00 – 0.20 is *very low*; if the result of the reliability is about 0.00 to 0.20, it is categorized to very low reliability;
- 0.20 – 0.40 is *low*; if the result of the reliability is about 0.20 to 0.40, it is categorized to low reliability;
- 0.40 – 0.70 is *medium*; if the result of the reliability is about 0.40 to 0.70, it is categorized to medium reliability;
- 0.70 – 0.90 is *high*; if the result of the reliability is about 0.70 to 0.90, it is categorized to high reliability;
- 0.90 – 1.00 is *very high*; if the result of the reliability is about 0.90 to 1.00, it is categorized to very high reliability;

(Sudijono, 2007:193)

After calculating the data (see Appendix 18 1nd 19), the result of the two reliabilities can be seen in the following tables:

Table 3.1 Inter-rater Reliability

Reliability	Pretest	Posttest	Criteria
	0,8326	0.8880	High Reliability

In Table 1, the criterion of reliability of the pretest and posttest is high reliability. It means that both of the raters were consistent in scoring the students' speaking.

### 3.5 Data Treatment

Before the writer analyzed the data, it was necessary that the writer had to find out whether the data was random or not or whether the data had normal distribution or not. There were two kinds of the data treatment in this research:

#### 1. Random Test

Random Test is used to ensure whether the data is random or not. One of the assumption should be fulfilled in using t-test is the data should be taken from random sample in a population. In this research,  $H_0$  was accepted if  $\text{sign} > \alpha$ , and the researcher used the level significant 0.05.

#### 2. Normality Test

This test was used to measure whether the data is normally distributed or not.

The hypothesis is accepted if  $\text{sign} > \alpha$ . In this case, the research uses the level of significance of 0.05.

### **3.6 Research Procedure**

In collecting the data, the researcher used the following steps:

#### **3.6.1 Selecting The Instrument Materials**

The instrument materials (speaking test) were chosen from authentic materials (English students' book for Senior High School, English magazine and internet). The selecting process considers materials that had been taught to the students and students' interest.

#### **3.6.2 Determining Research Instrument**

For both speaking tests (pre-test and post-test), the materials were taken from students' authentic materials (invitation, opinion, and permission), i.e. Students' English book (two passages) and internet (two passages). It aimed at making an equal proportion and level of difficulty of both pre-test and post-test.

#### **3.6.3 Determining The Population and Sample of the Research**

The sample of the research was determined through simple random probability sampling. It means that the sample was selected randomly by using lottery, since the 2<sup>nd</sup> grade in SMA Negeri 3 Metro was not stratified class, there was no priority class. There were eight classes of second grade at SMA Negeri 3 Metro. Then, the researcher determined one class to be the experimental class randomly.

#### **3.6.4 Conducting Pre-test**

The researcher conducted pre-test before treatment of Information Gap technique. This test took at least 90 minutes. The pre-test would be conducted to investigate the students' present speaking achievement before treatment. The topic would be chosen were invitation, opinion, and permission. The students were divided into pairs. Then, students made conversation based on the topic provided and their own knowledge. The teacher judged their speaking based on the three components of speaking; pronunciation, fluency, and comprehensibility.

#### **3.6.5 Giving Treatments**

Three treatments by using Information Gap technique were given in two weeks. The treatments were conducted in three meetings and 90 minutes for each. The treatments were classroom activity, which is used and applied Information Gap technique in speaking.

#### **3.6.6 Conducting Post-test**

The post-test aimed to know the progress of students' speaking ability after was given the treatment. This test was administered in 90 minutes. Similar with pre-test above, the students were divided into pairs or groups. The materials were about invitation, opinion, and permission. Then, students made conversation based on the topic provided and their own knowledge. This kind of activity stimulated the interaction among students. Then, the researcher recorded the students' conversation in cassette and checked the result by using J.B. Heaton table criteria of speaking ability.



### 3.6.7 Analyzing, Interpreting and Concluding the Data Gained

The data, in the form of score gained from pre-test and post-test were tabulated and calculated its inter-rater reliability. Then, calculate minimal score, maximal score, and mean of the pre-test and the post-test and its standard deviation. The comparison of two means would be counted using Repeated Measures t-test that told us whether there is any difference between the students' speaking achievement before being taught through Information Gap technique and the students' speaking achievement after being taught through Information Gap technique.

### 3.7 Criteria for Evaluating Students' Speaking Achievement

In evaluating the students' speaking scores, the researcher and another rater listened to the students' record and used the oral ability scale proposed by Heaton (1991). In scoring the test, the researcher implemented analytically scoring which covers pronunciation, fluency and comprehensibility. So the researcher scored the three aspects such as, pronunciation, fluency, and comprehensibility separately.

The following table was the oral ability scale proposed by Heaton (1991) that would be used as the scoring standard for the students' speaking ability.

Table 3.2 The Oral Ability Scale Proposed by Heaton

Range of score	Pronunciation	Fluency	Comprehensibility
90-100	Pronunciation only very slightly influenced by mother-tongue. Two or three minor grammatical and lexical errors.	Speaks without too great an effort with a fairly wide range of expression. Searches for words occasionally but only one or two unnatural pauses.	Easy for listener to understand the speaker's intention and general meaning. Very few interruptions or clarifications required.
70-89	Pronunciation is slightly influenced by the mother tongue. A few minor grammatical and lexical errors, but most utterances are correct.	Has to make an effort at times to search for words. Nevertheless, smooth delivery on the whole and only a few unnatural pauses.	The speaker's intention and general meaning are fairly clear. A few interruptions by listener for the sake of clarification are necessary.

60-69	Pronunciation is still moderately influenced by the mother tongue but no serious phonological errors. A few grammatical and lexical errors, but only one or two major errors causing confusion.	Although he has made an effort and search for words, there are not too many unnatural pauses. Fairly smooth delivery mostly. Occasionally fragmentary but succeeds in conveying the general meaning. Fair range expression.	Most of the speakers say is easy to follow. His intention is always are clear but several interruptions are necessary to help him to convey the message or to seek clarification.
40-59	Pronunciation is influenced by the mother tongue but only a few serious phonological errors. Several grammatical and lexical errors, some of which cause confusion.	Has to make an effort for much of the time. Often has to search for the desired meaning. Rather halting delivery and fragmentary. Range of expression often limited	The listener can understand a lot of what is said, but he must constantly seek clarification. Cannot understand many of the speaker's more complex or longer sentences.
30-39	Pronunciation seriously influenced by the mother tongue with errors causing a breakdown in communication. Many grammatical and lexical errors.	Long pauses while he searches for the desired meaning. Frequently halting delivery and fragmentary. Almost gives up for making the effort at times. Limited range of expression.	Only small bits (usually short sentences and phrases) can be understood – and then with considerable effort by someone who is used to listening to the speaker.
10-29	Serious pronunciation errors as well as many grammatical and lexical errors. No evidence of having mastered any of the language skills and areas practiced in course.	Full of long and unnatural pauses. Very halting and fragmentary delivery. At times gives up making the effort. Very limited range of expression.	Hardly anything of what is said can be understood. Even when the listener makes a great effort or interrupts, the speaker is unable to clarify anything he seems to have said.

### 3.8 Data Analysis

The data, in the form of score gained from pre-test and post-test were tabulated and calculated its inter-rater reliability. Then, calculate minimal score, maximal score, and mean of the pre-test and the post-test and its standard deviation. The comparison of two means would be counted using Repeated Measures t-test that will tell us whether there is a difference of the students' speaking achievement before and after being taught through

Information Gap technique and there is significant increase of the students' speaking achievement after being taught through Information Gap technique at the second grade of SMA Negeri 3 Metro.

### **3.9 Hypothesis Testing**

In order to test the hypothesis, the researcher compared the pre-test and post-test score by using Repeated Measure t-test of SPSS since the data comes from the same sample. This test is also known as paired t-test or dependent sample t-test (Setiyadi, 2006:170).

With t-table 0.05 (see Appendix 13), the hypothesis that has been mentioned is accepted with the following criteria:

1.  $H_1$ : If  $t\text{-value} > t\text{-table}$ , the alternative hypothesis ( $H_1$ ) is accepted and the null hypothesis ( $H_0$ ) is rejected. It means that there are a difference and significant increase of the students' speaking achievement before and after being taught through Information Gap technique at the second grade of SMA Negeri 3 Metro.
2.  $H_0$ : If  $t\text{-value} < t\text{-table}$ , the alternative hypothesis ( $H_1$ ) is rejected and the null hypothesis ( $H_0$ ) is accepted. It means that there are not any difference and significant increase of the students' speaking achievement before and after being taught through Information Gap technique at the second grade of SMA Negeri 3 Metro.