III.RESEARCH METHOD

This chapter describes the design of the research, how to collect the data from the sample of the research and how to analyze the data. This chapter also describes research procedure, population and sampling , data collecting technique, validity, and reliability , criteria for evaluating students' speaking ability , data analysis, and hypothesis testing.

3.1 Research Design

This research intended to find out whether drill technique could improve students' speaking ability. Therefore, the researcher conducted this quantitative research by using time series design. The researcher used one class where the students were given three times pre-test, three times treatment, and three times post-test.

The research design was described as follows:

$$T1 \ X \ T2 \ \ T3 \ X \ T4 \ T5 \ X \ T6$$

Note: T1 T2 T3   :Pre-test

        X   :Treatment   (Using Drill Technique)

        T4 T5 T6   :Post-test   (Setiyadi,2006: 131)

Pre-test was administered before the treatment of teaching speaking through drill technique was implemented, to see the students' basic speaking ability. Then, the treatment of teaching speaking through drilltechinque. The post-test was adminstered afterward, to analyze how the improvement of their speaking ability through drill technique.

3.2 Population and Sampling.

The population of the research was the second grade students of MAN 1 Bandar Lampung, in which consisted of ten (10) classes that were chosen by Purposive Random Sampling. The class chosen was class XI A_2, consisting of 32 students. The researcher chose XI A_2 as the sample because the students still had low ability in their speaking and afraid of making mistakes in expressing themselves.

3.3Data Collecting Techique

The data of this research was the form the students' speaking ability in performing transactional dialogue in terms of pronunciation, fluency, and comprehensibility. In collecting the data, the researcher used the following steps:

1. Administrating the Pre-test

The pre-test was administered to the students before the treatment of teaching speaking through drill technique was implemented. Pre-test was conducted to know the students' basic speaking ability. Meanwhile, before administered pre-test, the researcher explained the topic that would be tested. The tests focused on dialogue form of oral test.

2. **Administering the Post-test**

The post-test was administered to the students after the treatment of teaching speaking technique through drill technique would be implemented. It was a subjective test and focused in oral test..

3.4 Research Procedure

In the collecting data, the resercher follows the following steps:

1. **Determining the subject**

   There wereeight classes at second grade of MAN 1 Bandar Lampung which consisted of about 30- 38 students for each class. The sample of this reseacher choose by using probability sampling technique as the control and experimental class.

2. **Selecting the Materials**

   The researcher choose the materials from the students' book based on the syllabus. The material was about agreement and disagreement. The researcher conducted three times in improving students' speaking ability through drill technique.

3. **Conducting Pre-test**

   The pre-test was adminstered to the students before the treatment of teaching speaking through drill technique, to know the students' basic speaking ability. Meanwhile, before administer pre-test, the researcher explained the topic that would be tested. The tests focused on oral test. The researcher was conducted three times test by giving the same

topics in every tests. It is a subjective test and focused in oral test. The reseacher explained generally the test and asks the students to make a group that consist two persons. The researcher give the students situational dialogue andthey will perform it in front of class. In performing the test, the students ask to speak up clearly since the students' voice would be recorded. Futhermore, the researcher and another English teacher judge the students' performance.

4. **Treatments**

In this research, the treatments were administered in three meetings.At the first treatment the resercher deliver the indicators and objectives of treatment. She also explains the material by using drill technique. And the next, the researcher asks the students to make conversation dialogue with their pair group. The procedure of teaching speaking through drill technique as follows:

   a. Pre Activities

   b. While Activities

   c. Post Activities

5. **Administering Post-test**

Post-test was conducted after the treatment. Post-test was used in order to know the progress of speaking ability after using drill technique. Based on the design of this research that was time series design, post-tests was conducted in three times, after each meeting or after each treatment. The reseacher used a subjective test in oral test. The

researcher gave a text of situational dialogue for the student and they had to perform the dialogue in front of the class. In performing the dialogue, the students wasasked to speak up clearly since the students' voice would be recorded. Moreover, the researcher and the English teacherjudge the students' performance.

6. **Analyzing Data**

After conducting the final test, the reseacher analyzed the data. After collecting the data, the students' worksheet were analyzed subjectively by both reseacher and teacher. Then, The reseacher analyzed the mean of every test by compared from the two raters based on the test. The mean of pre-test and post-test were used to know the improvement of students' speaking ability through drill technique.

3.5 Criteria for Evaluating Students' Speaking Ability

The form of the test was subjective test since there was no exact answer. In this test, the researcher used *inter-rater reliability* to assessed the students' performance, in which the performance then were given score and were recorded together by the researcher as the first rater and the English teacher of MAN 1 Bandar Lampung as the second rater. The rater gave the score by recording the students' speech performance. The researcher recorded the students' utterances because it helped the raters to evaluate the data more objectively.

In fulfilling the criteria of a good test, validity and reliability of the test should be considered. Theywere as follows:

3.5.1   **Validity of the Test**

The test can be said valid if the test measures the objective to be measured and

suitable with the criteria (Hatch and Farhady, 1982: 250). To measure whether the

test in this research had a good quality, the types of validity was used in this

research were *content validity* and *construct* validity.

1. Content Validity

Content validity is the extent to which a test measures a representative sample of

the subject matter content, the focus of content validity is adequacy of the sample

and simply on the appearance of the test (Hatch and Farhady, 1982: 251). It is

correlated the test with the educational goal sated on 2006 English curriculum and

the syllabus for the first year of SMA's students.

In content validity, the test is a good reflection of what has been taught and the

knowledge which the teacher wants his students to know. It can best be examined

by the table of specification (Shohamy, 1957: 74). Below are the table of

specification:

No	Aspects of Speaking	Theories
1.	Pronunciation	Pronunciation refers to the ability to produce easily comprehensible articulation (Syakur, 1987). It refers to the intonation patterns (harris, 1984).
2.	Grammar	Grammar is needed for the students to arrange a correct sentence in creating a conversation. It is one of the ability of the students to manipulate and to distinguish appropriate grammatical form in appropriate ones (Heaton, 1978: 5).
3.	Vocabulary	Vocabulary means the appropriate diction which is used in communication (Syakur, 1987).
4.	Fluency	Fluency refers to the ease and the speed of the flow of the speech (Harris, 1974: 81). It also can be defined as the students' ability to speak fluently and accurately. Signs of fluency include a

		reasonably fast speed of speaking and a small numbers of pauses (Brown, 1974: 4).
5.	Comprehension	It defines that comprehension for oral communication which requires a subject to respond to speech as well as to initiate it (Syakur, 1987).

2. Construct Validity

Construct validity is concerned with whether the test is actually in line with the theory of what it means to know the language (Shohamy, 1985: 74) that is being measured, it examines whether the test actually reflects what it means to know a language. It means that the test measures certain aspects of speaking: *pronunciation, grammar, vocabulary, fluency,* and *comprehension* based on the indicator.

3.5.2 Reliability of the Test

Reliability refers to extend to which the test is consistent in its score and gives us an indication of how accurate the test score are (Hatch and Farhady, 1982: 244). The concept of reliability stems from the ideas that no measurements is perfect even we go to the same scale there will always be differences.

Inter-rater reliability was applied in this research in order to ensure the reliability of the score and to avoid the subjectively of the researcher. To achieve the reliability in judging the students' speaking performance, the researcher used a speaking criteria based on Harris (1974), in which the focus of speaking skills that have been asses are; *pronunciation, grammar, vocabulary, fluency,* and *comprehension.* The second rater was the English teacher who has been experienced in rating students' speaking, in means of getting a consistent and fair judgement. The statistical formula for counting the reliability was as follow:

$$R= 1 - \frac{6.\sum d^2}{N\,(\,N^2-1\;)}$$

Where:

R   : Reliability

N   : Number of Students

D   : The Different of Rank Correlation

1-6 : Constant Number

(Nitko, 1983: 395)

After finding the cooeficient between raters, the researcher then analyzed the cooeficient of reliability with the standard of reliability according to Slameto (1998: 147) in Hayanti (2010: 38) as follow:

A very low reliability   (range from 0.00 – 0.19)

A low reliability   (range from 0.20 – 0.39)

An average reliability   (range from 0.40 – 0.59)

A high reliability   (range from 0.60 – 0.79)

A very high reliability   (range from 0.80 – 0.100)

After calculating the data, the result of the reliability can be seen as the following tables:

Rater's Reliability

Reliability	Pretest	Posttest	Criteria
	0.94	0.96	Very high reliability

From the calculating and the criteria of a reability above, it can be concluded that the reliability of the rater is high, in which it means that the way of the first's rater of scoring was similar to the researcher's. They have almost the same scoring system.

3.6 Speaking Test

The researcher conducted an speaking test lasted for 90 minutes in each meeting. In conducting the tests, the researcher provided a topic. Each pair had to make a dialog based on the topic that was given by the teacher. The test was done orally and directly. The teacher called each of the pair one by one to come in front of the class to perfomed their dialog based on the topic given by the teacher. The researcher asked the students to speak clearly since the students' performance was being recorded during the test.

The form of the test was subjective since there was no exact answer. The teacher gave the score of the students' speaking ability based on the oral rating sheet provide. The teacher assessed the students concern on five aspects that should be fulfilled by the students in speaking skill; *pronunciation, grammar, vocabulary, fluency, and comprehension*. In scoring the test, the researcher used inter-rater. The one was the researcher itself, and then the other one was the English teacher of MAN 1 Bandar Lampung. During the process of evaluating the students' speaking ability scores, the researcher and the another rater who was the class teacher, listened to the students' record carefully and used the oral English. The

researcher recorded the students' utterance because it helped the raters easy to evaluate them and made them more objectively.

The rating sheet was modified from Harris (1974). Based on the oral rating sheet, there were five (5) aspects of speaking skills that should be tested, namely *pronunciation, grammar, vocabulary, fluency,* and *comprehension*. Below are the Sample Oral - English Rating Sheet of Harris (1974):

Pronunciation

___ 5. Has few traces of foreign accent.
___ 4. Always intelligible, though one is conscious of a definite accent.
___ 3. Pronunciation problems necessitate concentrated listening and occasionally lead to misunderstanding.
___ 2. Very hard to understand because of pronunciation problems. Must frequently be asked to repeat.
___ 1. Pronunciation problems so severe as to make speech virtually unintelligible.

Grammar

___ 5. Makes few (if any) noticeable errors of grammar or word order.
___ 4. Occasionally makes grammatical and word-order errors which do not, however, obscure meaning.
___ 3. Makes frequent errors of grammar and word order which occasionally obscure meaning.
___ 2. Grammar and word-order errors make comprehension difficult. Must often rephrase sentences and/ or restrict himself to basic patterns.
___ 1. Errors in grammar and word order so severe as to make speech virtually unintelligible.

Vocabulary

___ 5. Use of vocabulary and idioms is virtually that of a native speaker.
___ 4. Sometimes uses innapropriate terms and/ or must rephrase ideas because of lexical inadequacies.
___ 3. Frequently uses the wrong words; conversation somewhat limited because of inadequate vocabulary.
___ 2. Misuse of words and very limited vocabulary make comprehension quite difficult.
___ 1. Vocabulary limitations so extreme as to make conversation virtually impossible.

Fluency

    5. Speech as fluent and effortless as that of a native speaker.
    4. Speed of speech seems to be slightly affected by language problems.
    3. Speed and fluency are rather strongly affected by language problems.
    2. Usually hesitant often forced into silence by language limitations.
    1. Speech is so halting and fragmentary as to make conversation virtually
        impossible.

Comprehension

    5. Appears to understand everything without difficulty.
    4. Understands nearly everythig at normal speed, although occasional
       repetition may be necessary.
    3. Understands most of what s said at slower-than-normal speed with
       repetitions.
    2. Has great difficulty following what is said. Can comprehend only "social
       conversation" spoken slowly and frequent repetitions.
    1. Cannot be said to understand even simple conversational English.

Note: The interpretation of grading system was as follows:

Identification of the scores.

If the student got   5, so 5 x 4 = 20

                             4, so 4 x 4 = 16

                             3, so 3 x 4 = 12

                             2, so 2 x 4 = 8

                             1, so 1 x 4 = 4

For example:

A students got 5 in pronunciation, 2 in grammar, 3 in vocabulary, 4 in fluency,

and 3 in comprehension. Therefore, the student's total score would be:

Pronunciation   5 x 4 = 20

Grammar   2 x 4 = 8

Vocabulary   2 x 4 = 8

Fluency   4 x 4 = 16

Comprehension   3 x 4 = 12

It means that the student got 64 for speaking.

The oral production score was based on the five (5) components and can be compared in the percentage.

3.7 Data Analysis

In order to find out the improvement of the students' oral production before and after being taught through drill technique by using recount text, the researcher examined the students' score by using the following steps. The first step was transcribing the students' dialog performance, in which the researcher recorded the students' spoken work and transcribed into written form. The second step was scoring the pretest and posttest of the student's scores from the two raters using rating scale of Harris (1974)

The data of pretest (T1) and posttest (T2) scores can be seen on the table below:

Student's Name	Pronunciation		Grammar		Vocabulary		Fluency		Comprehension		Total	
	R1	R2	R1	R2	R1	R2	R1	R2	R1	R2	R1	R2
A												
B												
N =												

The data of inter – rater reliability scores of pretest and posttest can be seen on the table below:

No	Student's Code	Rater 1		Total	Rater 2		Total
		Pretest	Posttest		Pretest	Posttest	
1							

2							
3							
4							

The third step was calculating the mean scores of the data such as minimum score and maximum score to know the mean of the test, in which to know the improvement by making both pretest and posttest a graphic on the data. Hatch and Farhady (1982) stated that there were three posibilities of analysis result, shown by graphic below.
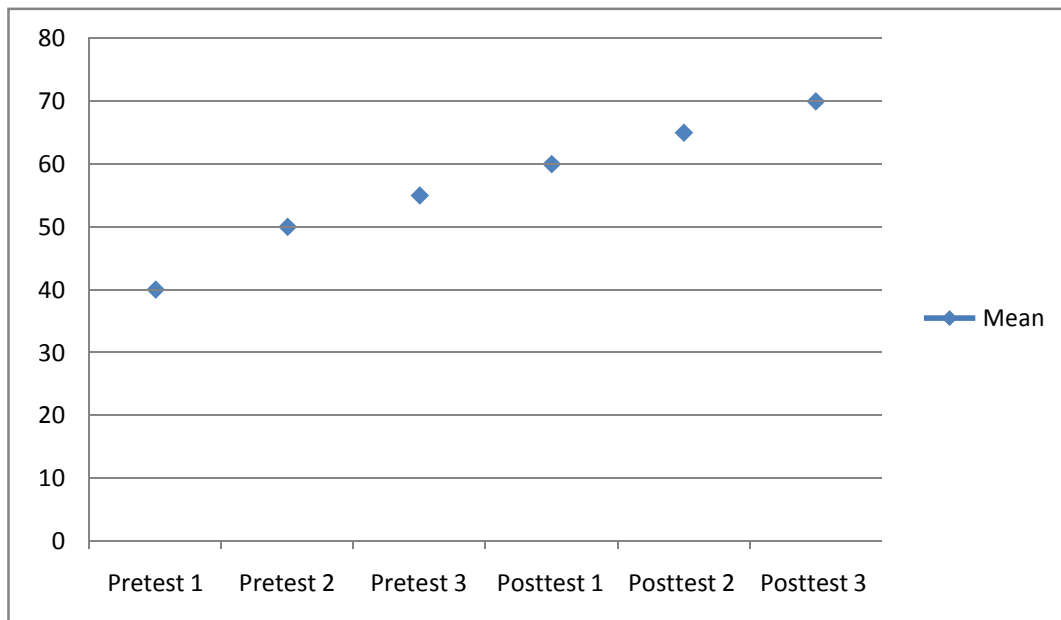
3.8 Hypothesis Testing

The hypothesis testing was used to prove whether the hypothesis planned in this research was accepted or not. SPSS was used to know the improvement of treatment effect. The hypothesis is analyzed at significance level of 0.05 in which the hypothesis is approved if Sig < . It means that probability of error in

hypothesis is only about 5%. After collecting the data, the researcher recorded and analyzed them in order to find out whether there was an increasing in students' ability in writing or not after the treatment. The researcher used Paired Sample T-test to know the level of

significance of the treatment effect.

The formulation is:

$$t = t = \frac{Md}{\sqrt{\frac{\sum x^2 d}{N(N-1)}}}$$

and

$$x^2 d = d^2 - \frac{(\sum d)^2}{N}$$

Md = mean from the differences pretest and posttest (posttest-pretest)

Xd = deviation of each subject (d – md)

$x^2 d$ = total of quadratic deviation

N = subjects on sample

(Arikunto, 2010: 349-350)

The criteria are:

H_0 : There is no improvement of the students' speaking ability after being taught by using drill technique.

The criteria is Ho (null hypothesis) is accepted if alpha level is higher than 0.05 (> 0.05)

H_1 : There is an improvement of the students' speakingability after being taught by using drill technique.

The criteria H1 is accepted if alpha level is lower than 0.05(< 0.05).