## **III. RESEARCH METHOD**

#### **3.1 Research Design**

This research was intended to find out whether there is a significant difference in students' speaking ability in the application of Finding Missing Information Technique and to find out the result of the application of Finding Missing Information Technique by giving three different topics. In conducting this research, quantitative design was employed that was *one-group pretest-posttest, quasi experimental with repeated measure design*. In this research the students were given pre-test before treatment, and after one treatment the students were given post-test. The pre-test is used to find out the students' preliminary ability and the post-test was used to look how the difference is after the treatment.

And then, the students were given the two other topics. Each topic had been evaluated to know whether there were any differences among those three different topics. The criteria whether there were any differences in students' achievement at a speaking class in the application of Finding Missing Information Technique were determined by comparing the result of the evaluation of each topic. The research design is formulated as follow:

T1 X<sub>1,2,3</sub> T2

Where,

T1 : Pre Test

X : Treatments (finding missing information technique)
T2 : Post Test
(Setiyadi, 2006:131)

In this research, the researcher used one treatment( $1^{st}$  topic) after pre-test. And then, the researcher used two treatments ( $2^{nd}$  and  $3^{rd}$  topics) after post-test for the first treatment.

## 3.2 Sample

The sample of this research was the first year students (7<sup>th</sup> grade) of SMP Satya Dharma Sudjana, PT. GMP, Lampung Tengah. The writer used one class, as the sample of this research class 7.6 was chosen as an experimental class, which is chosen randomly by lottery.

In this research, the researcher used three post-tests that were given after three treatments.

## 3.3 Data

The data of this research was in form of the students' speaking ability in performing transactional dialogue in terms of fluency, pronunciation, grammar and vocabulary under three topics, that are: Asking for Information, Like and Dislike and Describing Thing.

# **3.4 Procedure of Collecting Data**

In order to collect the data, the researcher followed the following steps:

a. Selecting speaking materials

In selecting the speaking material, the researcher saw the syllabus of the first year of SMP based on KTSP (unit-based curriculum). One of the objectives of KTSP (unit-based curriculum) for the first year students of SMP is that the students are intended to convey transactional dialogue in pair. So, the topic chosen were: Asking for Information, Likes and Dislikes and Describing Thing (preposition of places).

b. Determining instrument of the research

Since students' speaking ability would be evaluated, speaking test had been the instrument of this research. The researcher conducted pre-test and post-test of students' speaking ability in form of transactional dialogue in order to find the data before and after the treatment conducted. There were two raters to reduce the subjectivity in judging the students' speaking ability. In the intention of increasing reliability scoring, the raters judged the students' oral test twice, directly in the classroom and by listening the students' performance recorded. The final score was considered more based on recording since the raters analyze it thoroughly. The validity of the pre and post speaking tests of this research were related to face, content and construct validity.

c. Determining subject

There were seven classes of the seventh year in SMP "Satya Dharma Sudjana" Lampung Tengah in which this school already implemented KTSP. In determining the subject, the researcher used simple random probability sampling. That was through lottery. One class was chosen as the subject of this research. Each class of the seventh year had the same opportunity to be the subject of the research. One class taken as a subject was 7.6.

*d.* Conducting pretest

The researcher conducted the pre test before treatment of Finding Missing Information technique, which was at least 80 minutes. The pre test was conducted to investigate the students' present speaking ability. Pre test was given to know the students' speaking ability before the treatment. The topic chosen was asking for information (syllabus KTSP). The teacher provided some information that would be completed by them, let the students made group of two. Then the students were called to perform their dialogue in front of the class. In performing the task, the students were asked to speak clearly since the students' voice would be recorded. The researcher herself and another English teacher judged the students' performance.

e. Giving treatment (Finding Missing Information technique)

The treatment of Finding Missing Information technique was conducted in the class for three meetings with three different topics. The first treatment (1<sup>st</sup> topic: asking for information), second treatment (2<sup>nd</sup> topic: like and dislike) and the third treatment (3<sup>rd</sup> topic: describing thing). Each topic was evaluated at the end of the treatment. And the result of the evaluation of each treatment had been compared to find the differences of students' achievement in speaking ability. As many as three treatments, they were paired with different pair. Then, the teacher provided the topic that would be taught to the students. The topics are kinds of daily conversation. The students were asked to ask and answer to complete the missing information that he or she lacks.

### f. Conducting post test

The post test was gained at knowing the difference of students' speaking ability after being given treatment. The materials tested for the post test were similar in term of difficulties with the material for post test. In conducting the post test the researcher provided topic: asking for information. The students were, then called to perform their dialogue in front of the class. Again, the students were asked to speak clearly since the students' voice would be recorded. The researcher herself and another English teacher judged the students' performance.

# g. Analyzing, interpreting and concluding the data gained

After collecting the data in reference to the rating scales namely fluency, pronunciation, grammar and vocabulary, then analyzing, interpreting and concluding the data gained were done. First, the data, in form of score, gained from pre test and post test were tabulated and calculated to inter-rater reliability. Then it calculated minimal score, maximal score and mean of the pre test and the post test and its standard deviation. Repeated Measures T-Test or Paired Sample T-Test was used to draw the conclusion. The comparison of two means counted using Repeated Measures T-Test would tell us whether there was a difference in students' speaking ability significantly. The data were computed manually that shown two tail significance for equal variances as the value of significance. The hypothesis was analyzed at the significant level of 0,05 in which the hypothesis is approved if sig <  $\alpha$ .

## 3.5 Instrument of the research

Considering the newest curriculum i.e. KTSP (unit-based curriculum), some materials related to the topic were provided for speaking test. The material topics chosen were asking for information, like and dislike and descriptive text. In the pre test and the post test the students were given topic asking for information. The speaking test for both pre test and post test was done directly and orally, the teacher called the students in pair to perform transactional dialogue. They had five minutes for preparing their performance, and three minutes for performing the dialogue. In performing the task, the students were asked to speak clearly since the students' voice would be recorded. For the evaluation of each topic the researcher did the same procedure as well as the teaching procedure of Finding Missing Information technique.

In fulfilling the criteria of good test, reliability and validity of the test were clearly explained. The validity of the test of this research related to face, content and construct validity. To get face validity, the instructions of speaking test were previously examined by advisors and colleagues until the test which was in form of instructions looked right and understandable. The content validity means that the test is good reflection of what hass been taught and of the knowledge that the researcher wants her students to know. Here, the researcher correlated the test with syllabus and curriculum for Junior High School. If the table represents the material that the researcher wants to test, it can be said that it has content validity in line with the theory of what it means to know the language. It means that the test will measure certain aspect based on the indicator. The researcher examined it by referring the aspects that would be measure with the theories of those aspects (fluency, grammar, pronunciation and vocabulary).

In this research, reliability is defined as the stability or consistency of the test. One of the reliabilities purposed by Harris (1974:14) is reliability of the scoring of the test. Since the speaking test was a subjective test meaning the scoring process

dominantly influenced by the scorer, there were two raters to reduce the subjectivity in judging the students' speaking ability. The raters were the researcher herself and another English teacher. The raters worked collaboratively to judge students' performance. In the intention of increasing reliability of scoring the test, the raters judged the students' oral test twice. The first judgment was done directly in the classroom when the students were performing the task, while the second judgment was done by listening the students' performance recorded. The final scores were considered more based on recording since the raters analyzed it thoroughly.

The reliability of the test can easily be checked by comparing the scores they gave for the same students performance. The score given by the raters ideally should be nearly the same for the same performance. If the score gained is clearly excessive or highly different, the two raters should score the performance/composition again. If the score is still highly different, there should be the third rater. The score which is nearer to that of the third rater will be accepted (Harris, 1969:15,16,79,92).

In evaluating the students' speaking scores, the writer and another rater listened to the students' record and implemented the Analytic Rating Scale for speaking proposed by Shohamy (1985:180). Based on the Analytic Rating Scale, there are four aspects tested those were Pronunciation, Grammar, Vocabulary, Comprehension and Fluency.

#### Pronunciation

25-21 Excellent to very good: Has few traces of foreign accent.

- 20 16 Good to average: Always intellegible tough one is conscious of define accent.
- 15 11 Fair to Poor: Pronunciation problems necessitate concentrated listening and occasionally lead to misunderstanding.
- 10-6 Poor: Very hard to understand because of pronunciation problems must frequently be asked to repeat.
- 5-1 Very Poor: Pronunciation problems too serve as to make speech virtually unintelligible.

## Grammar

- 25–21 Excellent to Good: Make few (if any) noticeable errors of grammar or word order.
- 20 16 Good to Average: Occasionally makes grammatical and/or word errors, which do not, however, obscure meaning.
- 15 11 Fair to Poor: Make frequent errors of grammar and word order, which obscure meaning.
- 10-6 Poor: Grammar and word orders make comprehension difficult must often rephrase sentences and/or restrict him to basic patterns.
- 5-1 Very Poor: Errors in grammar and word order to serve as to make speech virtually unintelligible.

#### Vocabulary

25-21 Excellent to Good: Use of vocabulary and idioms is virtually that of native speaker.

- 20 16 Good to Average: Sometimes uses inappropriate terms and/or must rephrase ideas because of lexical inadequacies.
- 15-11 Fair to Poor: Frequently uses the wrong words, conversation somewhat limited because of inadequate vocabulary.
  10-6 Poor: Misuses of words and very limited vocabulary make comprehension quite difficult.
- 5-1 Very Poor: Vocabulary limitation to extreme as to make conversation virtually impossible.

## Fluency

- 25 21 Excellent to Good: Speech as fluent and effortless as that of native speaker problems.
- 20 16 Good to Average: Speed of speech seems to be slightly affected by language problems.
- 15 11 Fair to Poor: Speed and fluency are rather strongly affected by language problems.
- 10-6 Poor: Usually hesitant, often forced into silence by language problems.
- 5-1 Very Poor: Speech is so halting and fragmentary as to make conversation virtually impossible.

### **Table of Rating Sheet Score**

S's	Pronunciation	Fluency	Grammar	Vocabulary	Total
Code	(1-25)	(1-25)	(1-25)	(1-25)	(1-100)
1					
2					
3					

#### 3.6 Reliability

Reliability is a measure of accuracy, consistency, dependability or fairness of score resulting from administration of particular examination. To ensure the reliability of scores and to avoid the subjectivity of the researcher used Inter Rater Reliability. Inter Rater Reliability is used when two or more judges or raters independently estimate score on the test. In this case, the first rater of the research is the researcher herself and she asked the teacher of English as the second rater. To know how reliable the scoring is the researcher used Spearman Rank

Correlation.

The statistical formula is:

$$\mathbf{r} = \mathbf{1} - \frac{6 \sum \left(d^2\right)}{N\left(N^2 - 1\right)}$$

**r** : Coefficient of rank correlation.

**d** : Difference of rank correlation.

**1 and 6** : Constant number.

N : Number of Students. (Shohamy, 1985)

In this case, the researcher then analyzed the coefficient of rank correlation with the standard of reliability below:

0.8 - 1.00: very high reliability

0.6 - 0.79 : high reliability

0.4 - 0.59 : medium reliability

0.2 - 0.39: low reliability

0 - 0.19 : very low reliability

Slameto (1998:147) in Susan (2001:10)

In ensuring the reliability of the scores, the writer used inter rater reliability, that was by taking the scores from two scorers. So, there are two scores on each student's draft.

For example:

	Voc.	Gram.	Fluent.	Pron.	Total
Scorer 1	20	15	20	20	80
Scorer 2	20	15	25	20	80
					160:2
	So, the	e is	80		

## 3.7 Data Analysis

The data analysis has been done for the learning product, the researcher used speaking test to collect the data. There were some steps used to analyze the data got from the test:

a. Transcribing the students' utterance

After the teacher recorded the students' utterance, the researcher transcribed the record into the written form. This is very useful in order to give scores to the students and also to know the error mostly made by the students during speaking.

b. Scoring the students' speaking ability

Based on the transcription, the researcher and the teacher could decide the scores for the students' speaking test. The researcher used the Analytic Rating Scale proposed by Shohamy (1985).

c. Tabulating the result of the test and finding the difference mean of each post-test.

The mean was calculated by applying Repeated Measure t-test, with the following formula:

$$t_{obs} = \frac{X - \mu}{S_{\bar{x}}}$$

$$S_{\overline{x}} = \frac{S_x}{\sqrt{N}}$$

So, the *t* observed formula could be also be written as

$$t_{observed} = \frac{\overline{X}_1 - \overline{X}_2}{S_{x/\sqrt{N}}}$$

## Notes:



# d. Testing the Hypothesis

The Hypothesis of this research is:

There is a significant difference in students' speaking ability especially in terms of fluency, pronunciation, grammar and vocabulary for the application of Finding Missing Information technique.

The hypothesis was statistically analyzed by using Repeated Measure T-Test. By seeing the probability level (p) which is shown by two tail significance as the value of significance, we can draw the conclusion (Setiyadi, 2006:172). The researcher used the significant level of 0.05. It means that the probability of error in the hypothesis is only 5% from 100%, and the hypothesis was approved if p < 0.05.