III. RESEARCH METHOD

3.1 Research Design

This quantitative research was conducted in order to know whether there was increase of students' reading comprehension achievement of narrative text or not.*One-Group Pre-test and Post test design*was used in this research. The writer used one class as the experimental class. The research was conducted to see whether there was improvement of students' narrative text reading comprehension after being taught using CTL. The treatment was conducted three times by using narrative text. The researcher conducted pretest, treatment, and posttest. The design is as follow:

T1 X T2

Where,

T1is Pretest

X is Treatment

T2is Posttest

(Setiyadi, 2004: 40)

3.2 Population and Sample

The population of this research was the students of the second year of SMA Persada Bandarlampung in 2010. There were six classes, and two of them were IPA classes, namely XI IPA 1and XI IPA 2. Through lottery drawing, the researcher wrote the name of the classes then folded the paper then took randomly, the researcher had chosen students of XI IPA 1 as the sample.. The number of students in the class was 35. Class XI IPA 2 was chosen as the try out class.

3.3 Data Collecting Technique

This research used reading test as the instrument in collecting the data. The reading test consisted of 40 multiple choice items which was tried out to see its reliability. Then the researcher took 20 items by analyzing the index difficulty and index discrimination for pre-test and post-test. The try out, pre-test and post-test items consisted ofobjective multiple choices. The questions had four alternative answers for each number (A, B, C and D), one as the correct answer and the rest were the distracters. If one participant answered all the items correctly he or she would got 100 points. The treatment also used reading text.

1.Pretest

This was done before presenting the special treatment to know how far the student's ability in comprehending the reading text given by the writer. The researcher administered pretest for 40 minutes.

2. Posttest

This was done after presenting the special treatment to know how far the students master the material of narrative reading text by using Contextual teaching and Learning. The posttest was

conducted for 40 minutes. The test was designed to measure the increase of the students' reading comprehension achievement after the treatments were given.

3.4 Research Procedure

This research was based on the following procedures:

1. Selecting the instrument materials: the instrument materials (reading test) were chosen from Students' book and internet.
2. Determining the population and sample of the research: the sample of the research was determined through simple random probability sampling.It meant that the sample was selected randomly by using lottery.
3. Administering try out test : The researcher administered the try out test by using reading text and 40 items of multiple choices. The maximal points was 100, each correct answer had 2, 5 points. The test was given to find the quality of the test before it was used in order to get the data on the research. It was to find out whether the test items were good or not in validity, reliability, level of difficulty and the discrimination power. The researcher used split-half method to measure the reliability in which required him to provide the items into two same group, first half and second half.Some items were dropped to be administered for pretest and posttest. The difficulty level of try out test consisted of 11 easy items (1, 2, 3, 4, 12, 14, 16, 23, 33, 34 and 35), 20 average items (5, 7, 8, 9, 10, 11, 13, 15, 17, 18, 19, 22, 24, 27, 28, 29, 30, 36, 38 and 40), 9 difficult items (6, 20, 21, 25, 26, 31, 32, 37 and 39). Meanwhile for discrimination power, consisted of 11 bad items (1, 9, 15, 16, 18, 21, 25, 26, 34, 35, and 37), 9 poor items (2, 13, 20, 24, 31, 33 and 39), and 22 items satisfactory (3, 4, 5, 6, 7, 8, 10, 11, 12, 14, 17, 19, 22, 23, 27, 28, 29, 30, 32, 36, 38 and 40). Items with negative and zero

discrimination were dropped, meanwhile for the items with satisfactory difficulty level and discrimination power were administered for pretest and post test.

4. Determining final test of the instrument. In this step, the researcher dropped some items based on the results of try out.
5. Administering the pre-test: pre-test was conducted before the treatments. The pretest was aimed at finding out the students' basic reading comprehension achievement. The researcher administered pretest by using reading text and 20 items multiple-choice test. The pretest was 40 minutes.
6. Giving treatment: The reseacher gave three times treatments by using Contextual Teaching and Learning.
7. Conducting post-test: The posttest was aimed to find out the students' reading comprehension achievement after giving the treatments which is used as their reading technique. The test using reading text and 20 items multiple-choice texts. The post test was 40 minutes.

3.5 Analyzing the Data

The researcher analyzes the data by comparing the average score (mean) of the pretest and posttest to know whether there is improvement of students' reading ability through Contextual Teaching and Learning.

$$\overline{x} = \frac{\sum x}{N}$$

Where:

$\bar{x}$ = mean

$\sum x$= total scores

N = number of students

3.6 Criteria of Good Test

A test is said to have a good quality if it has good validity, reliability, level of difficulty and discrimination power (Heaton, 1991:5).

3.6.1 Validity

Validity of a test is the extent to which it measures what it is supposed to measure and nothing else (J.B Heaton 1975; 159). According to Heaton 1975; 159), there are four types of validity; face validity, content validity and construct validity and empirical validity. In this research, the writer use Face

validity, Content validity and Construct validity.

a. Face Validity

If a test items looks right to other testers, teachers and testees, it can be described as having at least face validity. Only if the test is examined by other people can some of the absurdities and ambiguities then be discovered. For example, the teacher asks to the students “can the tests be understood clearly? or are the items test clearly understood?”. The students’ motivation is maintained if the test looks sound. If, on the other hand, the test appears to have little of relevance in the eyes of the student, it will clearly lack face validity.

b. Content validity

Content validitymeans that the test is a good reflection of what has been taught andof the knowledge that the teacher wants his students to know. The focus of the content validity is adequacy of the sample and not simply on the appearance of the test. Content validity is intended to know how whether the test items are good reflection of what will be covered. The composition of the test items is presented in table 1 : Table of specification below.

Table 1. **Specification of the items test.**

NO	Reading Aspects	Item Numbers	Percentage of Reading Aspects
1	Determining main idea	2., 6., 10., 14., 16., 19., 24., 33., 34., 35..	25%
2	Finding specific information	1., 3., 7., 12., 17., 25., 28., 29., 38..	22,5%
3	Inference	5., 15., 20., 32., 37., 40..	17,5%
4	Reference	4., 9., 11., 13., 22., 26., 31..	15%
5	Vocabulary	8., 18., 21., 23., 27., 30., 36., 39..	20%

c. Construct validity

Construct validity means whether the test is actually in line with the theory of what it means to know the language, whether the test is actually a reflection of what it means to know a language (Shohamy, 1985:74-75). Here, the materials are arranged based on the curriculum and are adopted from the students handbook for the second year of SMU students. Nuttal (1985) states that the relation validity of the instrument refers to construct validity in which the questions represent five of sort reading skills, i.e. determining main idea, finding the detail information, inference,reference andvocabulary. Skills of reading in the test are a part of the construct validity and the item numbers are a part of the content validity.

3.6.2 Reliability

Reliability is a measure of accuracy, consistency, dependability, or fairness of scores resulting from administration of particular examination. To test the reliability of the instruments, the researcher will use*split-half* method.

Split half method is used by the researcher to estimate the reliability of the test. This formula is simple to use since : (1) it avoids troublesome correlation and (2) in addition to the number of items in the test, it involves only the test, mean and standard deviation. Both of which are normally calculated anyhow as a matter of routine, (Heaton, 1991:164). To measure coefficient of the reliablity the first and second half group, the researcher will use the following formula :

$$r_1 = \frac{\sum xy}{\sqrt{\sum X^2 \sum Y^2}}$$

Where :

r_1 = The coefficient reliability between first and second half group

X = The total numbers of first half group

Y = The total numbers of second half group

X^2 = The Square of X

Y^2 = The square of Y (Lado in Hughes,1991:3)

Then to know the coefficient correlation of the whole items, Spearman Brown formula is used:

$$rk = \frac{2r1}{1 + r1}$$

Where :

rk : The reliability of the test

rl : the reliability of the half test

The criteria of reliability are:

0.90 – 1.00 = high

0.50 – 0.89 = moderate

0.0 – 0.49 = low (Hatch and Farhady, 1982:268)

3.6.3 Level of Difficulty

To see the level of difficulty, the researcher uses the following formula :

$$LD = \frac{R}{N}$$

Where :

LD : Level of difficulty

R : The number of the students who answer correctly

N : The total number of the students following the test

The criteria are :

< 0.30 : difficult

< 0.30-0.70 : average

> 0.70 : easy (Shohamy, 1985:79)

According to the try out test there were 11 easy items (1, 2, 3, 4, 12, 14, 16, 23, 33, 34 and 35), 20 average items (5, 7, 8, 9, 10, 11, 13, 15, 17, 18, 19, 22, 24, 27, 28, 29, 30, 36, 38 and 40), 9 difficult items (6, 20, 21, 25, 26, 31, 32, 37 and 39).

3.6.4 Discrimination Power

To see the discrimination power, the researcher use the following formula :

$$DP = \frac{U - L}{½N}$$

Where:

DP : Discrimination Power

U : The proportion of the upper group students

L : The proportion of the lower group students

N : The total number of the students

The criteria are:

a. If the value is positive discrimination, it means that more high level students than low level students. If the value is zero, no discrimination.

b. If the value is negative, means that more low level students than the high level students get the item correct.

c. In general, the higher the discrimination index, the better. In classroom situation most items should be higher than 0.20 indexes.

(Shohamy, 985:81)

Based on the try out test there were 11 bad items (1, 9, 15, 16, 18, 21, 25, 26, 34, 35, and 37), 9 poor items (2, 13, 20, 24, 31, 33 and 39), and 22 items satisfactory (3, 4, 5, 6, 7, 8, 10, 11, 12, 14, 17, 19, 22, 23, 27, 28, 29, 30, 32, 36, 38 and 40).

3.6.5 Scoring System

In scoring students result of the test, the researcher uses Arikunto formula. The highest score was 100. The score of pretest and posttest was calculated by using following formula :

$$S = \frac{R}{N} 100$$

Where:

S : The score of the test

R : The total of the correct answer

N : The total items

3.6.5.1 Calculating of Mean

After obtaining the result of the students' test, the writer lists the scores and calculates their means through mean formula as follows:

$$\overline{x} = \frac{\sum x}{N}$$

Where:

$\overline{x}$ = mean

$\sum x$ = total scores

N = number of students

The mean tells us how difficult or easy a test is. According to Heaton (1991, p.175), the mean score of any test is the arithmetical average i.e. the sum of the separate scores divided by the total

number of students. A mean of 90 means that the test is easy; while an average of 40 means that it is difficult.

3.7 Data Analysis

In order to see whether there is an improvement of students' reading comprehension achievement, the researcher examines the students' score using the following step:

1. Scoring the pre test and post test.
2. After getting the raw score, the researcher tabulatedthe results of the test and calculated the score of the pre test and post test. Then the resercher used SPSS to calculatethe mean, reliability and coefficient correlation of pretest and posttest to see whether there was a significant increase or not after the students were taught by using Contextual Teaching and Learning.
3. Drawing conclusion from the tabulated result of the pre test and post test. The researcher used statistical computerization i.e. *repeated measures T-test of Statistical Package for Social Science (SPSS)for windowsversion 15* to test whether the increase of students' gain is significant or not. The researcher used this following formula:

$$t = \frac{\bar{\tilde{x}}_1 - \bar{x}_2}{S_{\tilde{D}}}$$

$$S_{\tilde{D}} = \frac{SD}{\sqrt{n}} \text{SD} = \sqrt{\frac{\sum D^2 - \left(\frac{1}{n}\right)\ (\sum D)^2}{n-1}}$$

X1 = Mean of pretest

X2 = Mean of posttest

$S_{\tilde{D}}$ = Standard Deviation of the differences

SD = Standard Deviation

n = Number of pairs (Hatch and Farhady, 1982:116)

$SD = \sqrt{\frac{4125 - \left(\frac{1}{35}\right)(355)^2}{35-1}}$ $SD = 3.92$

$S_{\tilde{D}} = \frac{3.92}{\sqrt{35}} S_{\tilde{D}} = 0.66$

$t = \frac{\bar{\tilde{x}}_1 - \bar{x}_2}{S_{\tilde{D}}}$ $t = \frac{60.28 - 7\ 042}{0.66}$

t = -15.28

The data above based on the Appendix 12. So a value of -15.28 with d.f. 34 is significant at the level of 0.000 ($p<0.005$). Thus, there is significant increase of students' reading comprehension achievement of narrative text in intensive reading after the students were taught by using Contextual Teaching and Learning.

3.8 Hypothesis Testing

Hypothesis of this research is:

There is significant increase of students' reading comprehension achievement of narrative text in intensive reading after students are taught by using Contextual Teaching and Learning.

The hypothesis testing was used to prove whether the hypothesis proposed in this research was accepted or not. The hypothesis was analyzed by using Repeated measure T-Test through computing with *Statistical Package for Social Science (SPSS)for windowsversion 15*. The

researcher used the level of significance 0,05 in which the hypothesis was approved if sign < p. It means that the probability of error in the hypothesis is only 5%.